Checklist en Overdrachtsdocument data Informatiehuis Marien, Datahuis Wadden en Waterinfo Extra.

*Toelichting:* onderstaand document is alleen bedoeld voor het aanleveren van data aan het Informatiehuis Marien, de Waddenviewer en Waterinfo-extra.

Ingevuld door:

1. Basiseisen dataset(s)

Toets (en vink aan) of onderstaande van toepassing is:

[ ]  het betreft data die openbaar zijn.

Toelichting: Dit is vrijwel altijd het geval. Mogelijke uitzonderingen kunt u [hier](https://data.overheid.nl/ondersteuning/data-gebruiken/niet-beschikbare-datasets) lezen.

Het gaat om data die (meerdere vinkjes mogelijk):

[ ]  voor/door Rijkswaterstaat zijn verzameld of berekend,

[ ]  van toepassing zijn op de Noordzee,

[ ]  van toepassing zijn op de Waddenzee

[x] Voor RWS data: De data kunnen niet of niet op korte termijn in een van de standaard bronsystemen van Rijkswaterstaat opgenomen worden (geldt alleen voor opname in viewer waterinfo extra)

Toelichting: Als de data bijvoorbeeld in WADAR (RWS-standaardopslag voor waterkwaliteit/waterkwantiteit) of aquadesk (Biologische data) kunnen worden opgenomen, is dit de voorkeursoptie. Dit voor de meeste chemische parameters en waterkwantiteit (waterstanden, debieten enz.), maar ook voor gegevens over macrofauna, plankton, vissen, vogels het geval. Opname in deze bronsystemen is meestal niet mogelijk voor opgewerkte data (bijvoorbeeld jaargemiddelden) of als de data niet aan de AQUO-standaard kunnen voldoen. Overleg dit met je contactpersoon bij WIE/DHW/IHM.

1. Basisgegevens dataset

 Vul de tabel zoveel mogelijk in. Kopieer de tabel en gebruik meerdere malen als er meerdere datasets aangeleverd worden.

|  |  |  |
| --- | --- | --- |
| Item | Antwoord:  | Toelichting |
| Naam dataset(s)  |  | Bijvoorbeeld: Vispopulatiegegevens Haringvliet 2012-2020 |
| Opdrachtgever metingen/verzameling data |  | Voorbeelden: RWS-WVL, of RWS-ON, provincie Zeeland |
| Beheerder dataset |  | Voorbeeld ‘Wageningen Marine Reseach’, ‘RWS-CIV’ |
| contactpersoon |  | - voor vragen over inhoud van de dataset. - wordt alleen gebruikt voor intern contact, dus niet openbaar gedeeld  |
| Korte beschrijving van het project (een of twee alinea’s) |  | - waar? kaartmateriaal is welkom- doel van metingen (bijv. effect van baggeren op vissen)- Methodiek (bijv ‘electrovisserij’)- relatie met andere data (wellicht zijn er ook andere datasets binnen zelfde project) |
| 3 a 4 trefwoorden |  | Bijvoorbeeld ‘oeverinrichting’ , ‘macro-fauna’ |
| Link naar (data)rapport(en) |  | deze worden bij voorkeur gepubliceerd op het publicatieplatform van Rijkswaterstaat (dit kan je [zelf](https://rijksportaal.overheid-i.nl/onderwerpen/werkplek/artikelen/bibliotheek--en-archiefdiensten/bibliotheekdiensten/bibliotheekdiensten-rws/rapporten-van-rws-publiceren_rws.html) doen, link naar rijksportaal), of via een andere openbare digitale bibliotheek (bijvoorbeeld van een onderzoeksinstituut |
| Looptijd  | Van: ddmmyyyy tot ddmmyyyy |  |
| Wordt de dataset nog beheerd/gecureerd, en zoja, door wie? |  | Wanneer deze vragen met nee worden beantwoord, wordt de dataset als een finaal niet meer beheerd product beschouwd.  |
| Juridische restricties |  | Zie tabel A.10 op deze [pagina](https://geonovum.github.io/Metadata-ISO19115/#codelist-restrictioncode) voor opties (gaat meestal om ‘otherRestrictions’) |
| Gebruiksbeperkingen |  | Bijvoorbeeld: niet te gebruiken voor navigatie. Vaak: ‘geen gebruikslimitaties’ |
| coordinatensysteem |  | Bijvoorbeeld: Rijksdriehoekscoördinaten of ‘n.v.t.’ |
| Welke bemonsterings- en labprotocollen zijn van toepassing? |  | Bijvoorbeeld: link of verwijzing naar protocollen, of: ‘n.v.t.’ |
| Heeft er databewerking plaatsgevonden (weglaten van nulwaarnemingen, middelen van metingen enz.), eventueel volgens een protocol? Zoja deze graag bijvoegen/beschrijven in een bijlage  |  | Bijvoorbeeld: Ja (+bijlage)/Nee |

1. Aanlevering data en toelichting

Toets (en vink aan) of onderstaande van toepassing is:

[ ]  Bij de dataset wordt een toelichting op het datamodel aangeleverd (mag als losse bijlage).

Toelichting: Om de waarde van een dataset te behouden moet voor de gebruiker de betekenis van kolommen/kolomnamen/parameters duidelijk zijn. Hierbij moet ook duidelijk zijn welke grootheden/eenheden gebruikt worden.

Voorbeelden:

<https://waterinfo-extra.rws.nl/publish/pages/155698/wmr_toelichtingen_visdata.xlsx>

<https://www.informatiehuismarien.nl/publish/pages/109348/ihm_database_velden_2020_03_20.xlsx>

[ ]  Het gaat om een opgeschoonde dataset met consequent gebruik van datums en tekens. Zie toelichting in bijlage voor tips en richtlijnen.

Bijlage: richtlijnen aanleveren datasets voor WIE/IHM/DHW

Voor de verwerking van gegevens is het vereist om de gegevens op een consistente manier binnen te krijgen waarbij er gestreefd dient te worden naar ondubbelzinnige gegevens. Let hierbij op de volgende aspecten:

* Datasets het liefst aanleveren in een csv bestand. Andere formats zijn mogelijk na overleg.
* Liefst geen lege regels in sheets
* Pas op met komma’s in XLSX en andere data typen (datum getal, tekst) in spreadsheets (zo staat excel het gebruik van tekst in een datum veld toe, of bijvoorbeeld een + teken in een veld wat percentages zouden moeten zijn)
* Dezelfde kolomnamen gebruiken over alle bestanden (als sprake is van meerdere bestanden)
* Aantal kolommen moeten in elk bestand gelijk zijn, ofwel, 1 csv bevat bijvoorbeeld 10 kolommen en niet soms 9 en soms 10.
* Inhoud van de kolommen zijn consistent, een datum is bijvoorbeeld altijd volgens een bepaald vastgelegd format, bijvoorbeeld yyyy-mm-dd (jaar – maand – dag, bijvoorbeeld 2022-12-31 voor 31 december 2022).
* Als er gebruik word gemaakt van verschillende coördinatensystemen dan is het raadzaam om het Coördinaat referentie Systeem (CRS) erbij te vermelden. Rijksdriehoeksstelsel heeft bijvoorbeeld als CRS EPSG:28992. Voor het veel gebruikte GPS systeem WGS84 is dat bijvoorbeeld EPSG:4326.