

# NODCi: op NAAR een NATIONALE Infrastructuur VOOR ontsluiting VAN Oceanografische EN Mariene DATA EN informatie

Ruimte voor **Geo-Informatie**



## Dick Schaap

coördinator NODCi project, MARIS BV, Leidschendam  
schaap@maris.nl

**NODCi is één van de projecten in het programma Ruimte voor Geo-Informatie (RGI) en richt zich op het verbeteren van het overzicht en de toegankelijkheid van data van de oceanen en zeeën. Fase 1, de definitie en analyse fase is recent afgerond, en thans ligt het voorstel voor Fase 2, implementatie en operatie, ter beoordeling van het RGI programmabureau.**

### 1. Belang van mariene en oceanografische data

Wereldwijd zijn vele honderden instituten actief in de verzameling en verwerking van data van de oceanen, zeeën en kusten. Metingen worden verricht door onderzoeksinstituten, overheden en bedrijven. Dit betreft data op het gebied van o.a. fysische oceanografie, biologie, waterkwaliteit, geologie, bathymetrie en meteorologie. Metingen geschieden in-situ vanaf schepen, drijvende en vaste opstellingen en via remote sensing technieken vanaf satellieten, vliegtuigen en helicopters.

De snelheid en het gemak waarmee gebruikers oceanografische en mariene data kunnen identificeren, localiseren, opvragen, uitwisselen en gebruiken zijn belangrijke voorwaarden voor succesvol onderzoek en kennisontwikkeling op het gebied van de oceanen, zee en kust. Multi-disciplinaire data zijn ook belangrijk voor beleidsvorming en uitvoering, coastal zone en water management, ontwerp en aanleg van mariene infrastructuur en offshore installaties, be-

scherming van het land tegen stormvloed, rampenbestrijding e.d. De data ondersteunen verder de uitvoering van internationale protocollen, conventies en overeenkomsten, die door kuststaten zijn ondertekend voor het beheer en de bescherming van de zeeën en kusten, en de uitvoering van Europese richtlijnen.

De noodzaak van een sterke verbetering van de data-ontsluiting wordt bevestigd in recente grootschalige raamwerkprogramma's op wereldschaal en op Europese schaal. Voorbeelden zijn o.a. het GEO initiatief (ad-hoc Group of Earth Observations van de G8), de daarvan afgeleide Europese component GMES (Global Monitoring for Environment and Security) – een gezamenlijk programma van European Space Agency – ESA (satelliet observaties) en European Commission DG Research (in-situ data) – en INSPIRE, een initiatief van de EC voor een European Spatial Data Infrastructure. Deze programma's worden elk voorgesteld als 'system of systems', daarmee aangevende dat zo veel mogelijk gebouwd zal worden op bestaande componenten, organisaties en netwerken. Met name wordt ingezet op verbetering van onderlinge cohesie, samenwerking, uitwisseling van data, harmonisatie van kwaliteit van data, harmonisatie van protocollen voor data toegang en ontwikkeling van modellen en applicaties voor het maken van data-producten met hogere toegevoegde waarden en van nut voor diverse gebruikers groepen.

### 2. NODC als initiatiefnemer

Het NODCi project omvat de ontwikkeling in Nederland van een Nationale Infrastructuur voor ontsluiting van Oceanografische en Mariene Data en Informatie, die aansluit op de internationale netwerken en ontwikkelingen. Het project is een initiatief van de Nationale Oceanografische Data Commissie (NODC). Deze Commissie is in 1997 opgericht als Nederlands platform voor de uitwisseling van oceanografische gegevens en informatie, alsmede voor advisering op het gebied van oceanografisch data management. Via een conve- nant zijn de volgende organisaties lid van de NODC:



The screenshot displays the SEA-SEARCH CDI web interface. At the top left, there are navigation tools like 'Enlarge', 'Help', 'Position', and 'Index'. The main area shows a map of the North Atlantic with a grid. On the right, a 'Layer control' panel allows users to toggle various map layers such as 'CDI entries', 'Main Sea-areas', 'Regional Seas', 'Grid lines', 'Partner locations', 'Main sea-area labels', 'Regional sea labels', and 'Bathymetry'. Below the map is a search bar and a 'Lat/long' input section with 'Upper-left' and 'Lower-right' coordinates and 'Search' and 'Clear' buttons.

The search results section shows a list of data sets. The selected entry is 'Fixed Oceanographic Station - Lista (1942-)' from Norway, established in 2005. The 'Series Details' panel provides comprehensive information:

- Data set name:** Fixed Oceanographic Station - Lista (1942-)
- Discipline:** Physics
- Category:** Water column temperature and salinity
- Variables measured:** Salinity of the water column, Temperature of the water column
- Abstract:** Long time series of oceanographic stations at fixed positions, and standard depths, carried out by local fishermen. Lista is one of currently 8 stations in activity. All oceanographic data are processed, and partly quality controlled. Simple meteorological observations are normally carried out on each station (no QA).
- Data topic:** Oceans
- Language:** English
- Data set creation date:** 20051107
- Longitude 1:** 13.63333
- Latitude 1:** 68.36333
- Datum:** WGS 84
- Measuring area type:** Point
- Minimum instrument depth (m):** 1
- Maximum instrument depth (m):** 175
- Map:** A small map showing the station location.
- Start date:** 19350223
- End date:** 19350223
- Sampling interval:** From weekly to less than monthly
- Instrument / gear type:** CTD profiler
- Platform type:** Small shore-based craft
- Originator:** Institute of Marine Research - Norwegian Marine
- Data holding organisation:** Institute of Marine Research - Norwegian Marine
- E-mail data contact:** j20917@smr.no
- Access/ordering of data:** Direct web access without registration
- Internet access/ordering:** Click here
- CDI-METADATA:** 3619
- CDI-record id:** 19350223
- CDI-record creation date:** Institute of Marine Research - Norwegian Marine

Below the details, there are two depth profile graphs for 'Eggum: 23-feb-1935'. The first graph shows 'Temperature' with measured values of 3, 5, and 7, and a normal value of 4.5. The second graph shows 'Salinity' with measured values of 33, 34, and 35, and a normal value of 33.9. A green shaded area in the graphs indicates the normal values calculated from all available data for the period 1935-1993, interpolated to this date.

## Common Data Index - CDI

### www.sea-search.net

Web-dialogo van Common Data Index (CDI), ISO19115 compliant (zie [www.sea-search.net/cdi/welcome.htm](http://www.sea-search.net/cdi/welcome.htm))

- Koninklijk NIOZ
- Rijkswaterstaat, RIKZ
- Rijkswaterstaat, Directie Noordzee
- TNO- Bouw & Ondergrond
- KNMI
- NIOO-CEME
- WL/Delft Hydraulics
- Dienst der Hydrografie
- MARIS BV (Adviseur)

Samen beheren de partners ca 90% van de oceanografische en mariene data, die door Nederland wordt ingewonnen op de Noordzee, maar ook breder wereldwijd via oceaanonderzoek. Daarnaast beheren de partners informatie van vastgoed op de Noordzee: offshore platforms, pijpleidingen, putten, concessiegebieden, e.d.

Tot op heden heeft de NODC zich vooral gericht op de samenstelling van metadata van beschikbare data sets, monitoring locaties en projecten. Hiervoor wordt aangesloten op metadata formats, die internationaal in de oceanografische data wereld worden toegepast en ontwikkeld, zoals o.m.:

- European Directory of Marine Environmental Datasets - EDMED
  - European Directory of Ocean-observing System - EDIOS.
- Deze directories zijn te raadplegen op de Europese portal website: <http://www.sea-search.net>

Via het RGI project breidt de aandacht van het NODC zich uit naar het ontsluiten van de data sets, die de partners lokaal beheren in hun eigen data management systemen. Uiteindelijk doel is te komen tot een situatie, waarbij gebruikers volledig transparant en uniform kunnen zoeken

via internet naar gegevens, die door de partners worden beheerd, en via een portal toegang kunnen krijgen tot de geselecteerde gegevens, zonder zich verder te verdiepen in de werking en interfaces van de verschillende partnersystemen.

### 3. Internationale samenwerking en afstemming

Nationale Oceanografische Data Centra (NODC's) zijn actief in 62 landen wereldwijd als nationale representanten in het **International Oceanographic Data and Information Exchange (IODE)** programma van de Intergovernmental Oceanographic Commission (IOC) van UNESCO. De NODC's richten zich op professioneel data management, ontwikkeling van standaarden, ondersteuning van instituten en gebruikers en productie van toegevoegde waarde producten.

Daarnaast is de NODC als Nederlands vertegenwoordiger ook actief in het EU project Sea-Search en haar recente opvolger SeaDataNet, die leidend zijn in de opbouw en harmonisering van de Europese infrastructuur voor mariene datamanagement. Sea-Search telt 33 partners (NODC's) uit 30 landen rond Europese zeeën. In SeaDataNet is dit uitgebreid naar 49 partners uit 35 landen.

De ontwikkeling van de NODCi infrastructuur in Nederland gaat in praktijk derhalve gelijk op aan de ontwikkeling van de Europese infrastructuur tussen de diverse nationale netwerken. Dit betekent, dat er een brede afstemming is op het gebied van standaarden (metadata format, data format, QC methodes, etcetera) en tevens dat nationale ontwikkelingen internationaal kunnen worden ingebracht.

### 4. Uitvoering van NODCi Fase 1

Vooraf gaande aan Fase 1 is een NODCi concept opgesteld, dat het afgelopen jaar is getoetst bij elk van de deelnemende gegevenshouders en afgestemd met het concept, zoals dat in het RGI project "Geoloketten" wordt voorgesteld. Dit laatste project zal de 'ruggegraat' gaan bieden voor de Nationale Geo Informatie Infrastructuur (NGII), die via RGI wordt nagestreefd. Opgave voor NODCi is derhalve aan te sluiten op de internationale afspraken binnen het mariene en oceanografische domein en op de nationale afspraken, die uitwisseling tussen domeinen in Nederland mogelijk moeten maken. De adoptie van ISO normen biedt hiervoor de geschikte weg.

Met alle partners (zowel managers als technici) is overleg gevoerd om nader inzicht te krijgen in hun locale database systemen en om het originele NODCi concept te toetsen op haalbaarheid. Uitgangspunt is namelijk, dat de locale systemen gehandhaafd blijven en verantwoordelijk zijn voor het data beheer. De NODCi infrastructuur gaat rusten op standaard protocollen voor communicatie, (meta) data formats en op termijn voor Quality Control methodieken en op de aanleg van een centrale portal voor identificatie van data en geregelde toegang tot data.

In hetzelfde tijdsvenster is actief deelgenomen aan meetings en activiteiten van het EU-Sea-Search project, waarin met name is bijgedragen aan de realisatie van een pilot toepassing voor een centrale index tot mariene data sets, die bij

partners worden beheerd. Er is een Common Data Index (CDI) metaformat geformuleerd, als subset op het ISO19115 model, en volledig gedocumenteerd met ondersteunende libraries (gecontroleerde trefwoordenlijsten) en tools voor generatie van CDI-XML records uit locale systemen. In 2005 is vervolgens een pilot CDI service gelanceerd, gebouwd door MARIS en met bijdragen van NODCi partners NIOZ en RIKZ, die via [www.sea-search.net](http://www.sea-search.net) kan worden bezocht. (zie figuur voor dialoog).

De analyse en overleg activiteiten hebben geresulteerd in een gemeenschappelijk concept voor de NODCi infrastructuur. Dit is vervolgens ook besproken met een brede groep van potentiële gebruikers uit de scheepvaart, scheepsbouw, offshore, bagger en survey industrie, wetenschappelijk zeeonderzoek sector en van de diverse partner organisaties.

### 5. Voorgesteld NODCi concept

Een vitaal onderdeel is een centrale en up-to-date metadirectory op de NODCi portal voor indexering van individuele data sets bij partners. Deze NODCi catalogus biedt gebruikers de functionaliteit om beschikbare data te kunnen identificeren en om door te kunnen schakelen naar mogelijke data toegang. In Fase 1 is met succes gezien dat hiervoor aangesloten kan worden op het Europese Common Data Index format (CDI). Ook is met succes de toepasbaarheid verkend van de Parameter Discovery Vocabulary (PDV), die binnen Sea-Search is ontwikkeld onder leiding van het British Oceanographic Data Centre. Het geadopteerde CDI index format is met hulp van RAVI vergeleken met de recent verschenen NL Metadatastandaard voor Geografische Datasets. Hieruit blijken een aantal kleine verschillen, vooral ingegeven door het feit, dat in NODCi sprake is van dynamische datasets (tijdseries en verticale profielen) en minder van kaartdatasets. Dit zal in Fase 2 nader worden opgelost.

De NODC partners beheren deels overlappende data typen en deels verschillende data typen op het gebied van hydrodynamica, mariene biologie, waterkwaliteit, hydrografie, mariene geologie, vastgoed informatie, etcetera). Het oorspronkelijke NODCi concept had de ambitie om data toegankelijkheid te realiseren via dataviews (virtuele tabellen), waarbij gelijksoortige datasets van meerdere partners via datatype-gemeenschappelijke formats zouden worden toegepast. Via de dataview techniek behoeven gebruikers datasets niet te downloaden als files, maar kunnen ze direct selecties maken op de virtuele tabellen, gegevens inzien en applicaties erop aansluiten.

Op basis van de nadere analyse van de huidige database systemen bij partners en de lopende internationale ontwikkeling van gemeenschappelijke datatype formats is besloten om de data toegankelijkheid in NODCi vooralsnog te regelen via de procedure, dat gebruikers de NODCi catalogus raadplegen en vandaar data files van partners kunnen downloaden. Dit zal geschieden via een Authentication, Authorisation en Accounting (AAA) module, omdat enige controle op het gedrag van gebruikers noodzakelijk is vanuit veiligheid en vanuit verschillen in data policies tussen partners. De downloading / export zal gaan geschieden in een voor alle datatypen gemeenschappelijk uitwisselings format

(container model) in ASCII en NetCDF. Dit laatste is een binair format, dat ook de uitwisseling van grid data (zoals model uitkomsten) ondersteunt. Het wordt internationaal in toenemende mate voor oceanografische data-uitwisseling toegepast. Voor het NetCDF format zijn reeds diverse web applicaties beschikbaar, zodat gebruikers daar hun voordeel mee kunnen doen.

De dataview techniek zal (nog) niet op volle schaal worden ingevoerd, omdat het vergelijken en definiëren van gemeenschappelijke inhoudelijke data-formats voor diverse datatypes (bijv biologische data) veel tijd vergt (mede ivm internationale afstemming) en ook omdat de dataview techniek nog niet door alle partners gerealiseerd kan worden vanwege hun huidige systeem opzet en data policies. Omdat dataview techniek aantrekkelijke voordelen kan bieden, zal in Fase 2 een pilot worden uitgevoerd, gericht op 1 type data: fysisch oceanografische tijdseries op punt locaties.

Bij de verschillende partners is de haalbaarheid van NODCi infrastructuur als Web Services onderzocht. Hierbij is mede gebruik gemaakt van de kennis en ontwikkelingen op dit gebied in het project 'Geoloketten'. Dit heeft geleid tot adoptie van een Web Services architectuur, waarbij de communicatie tussen lokale systemen en de portal gaat geschieden via SOAP.

De NODCi portal wordt uitgerust met een dynamische geografische – alphanumerieke user interface. Voor de geografische component is gekozen voor de open source Minnesota Open Map Server, die OGC compliant is.

Voor Fase 2 gelden aldus de volgende specifieke doelen:

- Het realiseren van NODCi in een Web Services architectuur met een NODCi portal en zelfstandige partner systemen.
- Het realiseren van een centrale NODCi catalogus, gebaseerd op het Sea-Search Common Data Index format (CDI = ISO19115 compliant), dynamisch gevoed vanuit partner systemen en uitgerust met een gecombineerde geografische – alphanumerieke user interface
- Het realiseren van toegang tot data sets voor gebruikers via Catalogus, Registratie en downloaden van data sets in een standaard format (ASCII en mogelijk NetCDF).
- Het realiseren en evaluëren van een pilot voor directe data toegang voor partners via dataview techniek voor 1 type data: fysisch oceanografische tijdseries op punt locaties.
- Het uitwerken en vastleggen van een regeling met partners voor consolidatie van NODCi qua beheer, onderhoud en operatie.
- Het promoten van de NODCi services naar gebruikers bij partners intern en extern en het dissemineren van de kennis, verkregen in het RGI-project, naar andere data en infrastructuur beheerders.

Bij spoedige goedkeuring van het Fase 2 voorstel door het RGI programma zal de uitwerking plaatsvinden in de komende 2.5 jaar.