D3.4 – Dataset specifications - Second set of datasets

Authors: Ann Kristin Østrem, Roald Vang, Wouter Koch, Joël Durant, Torill Hamre, Nina Nordlund, Aave Lepland, Svein Østerhus, Øystein Godøy, Svein Kristiansen, Cindy Hornaff, Benjamin Pfeil, John Rune Selvik, Stein Tronstad, Rune Palerud

Revision: 0.5
Date: 2015-08-18

Revision history

Version	Date	Comment	Responsible
0.1	2015-03-12	Initial topic layout, copied descriptions for most partners from D3.3.	Ann Kristin Østrem
0.2	2015-03-26	Updated references, added summary.	Ann Kristin Østrem
0.3	2015-04-30	Added descriptions from NERSC, FFI, NGU, Uni Research, IMR	Torill Hamre, Nina Nordlund, Aave Lepland, Svein Østerhus, Ann Kristin Østrem
0.4	2015-06-26	Added descriptions from METNO, UiT, UiN, UiB, NIVA, NPI	Øystein Godøy, Svein Kristiansen, Cindy Hornaff, Benjamin Pfeil, John Rune Selvik, Stein Tronstad
0.5	2015-08-18	Added description from Akvaplan-niva	Rune Palerud

Contents

1 Introduction 5
1.1 Background 5
1.2 Objectives of report 5
1.3 Terminology 5
1.4 References 5
1.5 Outline of report 6
2 Data format and protocols 7
2.1 Institute of Marine Research 7
2.1.1 Data description and format 7
2.1.2 Data delivery protocols 8
2.2 Norwegian Defence Research Establishment (FFI) 9
2.2.1 Data description and format 9
2.2.2 Data delivery protocols 9
2.3 University of Nordland 10
2.3.1 Data description and format 10
2.3.2 Data delivery protocols 10
2.4 Norwegian Meteorological Institute 10
2.4.1 Data description and format 10
2.4.2 Data delivery protocols 10
2.5 Nansen Environmental and Remote Sensing Center 11
2.5.1 Data description and format 11
2.5.2 Data delivery protocols 11
2.6 Geological Survey of Norway 12
2.6.1 Data description and format 12
2.6.2 Data delivery protocols 14
2.7 Norwegian Institute of Nature Research 14
2.7.1 Data description and format 14
2.7.2 Data delivery protocols 15
2.8 Norwegian Polar Institute 16
2.8.1 Data description and format 16
2.8.2 Data delivery protocols 17
2.9 University of Bergen 17
2.9.1 Data description and format 18
2.9.2 Data delivery protocols 18
2.10 University of Oslo 18
2.10.1 Data description and format 18
2.10.2 Data delivery protocols 18
2.11 University of Tromsø 18
2.11.1 Data description and format 18

2.11.2 Data delivery protocols 19
2.12 Norwegian Institute of Water Research 19
2.12.1 Data description and format 19
2.12.2 Data delivery protocols 19
2.13 Akvaplan-NIVA 19
2.13.1 Data description and format 19
2.13.2 Data delivery protocols 19
2.14 Uni Research 20
2.14.1 Data description and format 20
2.14.2 Data delivery protocols 20
2.15 Norwegian Biodiversity Information Centre 20
2.15.1 Data description and format 20
2.15.2 Data delivery protocols 21
3 Summary 21
3.1 Data formats 21
3.2 Data delivery protocols 21

1 Introduction

1.1 Background

The Norwegian Ministry of Fisheries and Coastal affairs have identified the importance of an operational national service and competence centre for marine data. The national infrastructure, the Norwegian Marine Data Centre (NMDC), will be the focal access point for marine data in Norway. NMDC is a continuation of an earlier infrastructure project [1] and aims to serve the marine science community with seamless access to documented marine data sets covering waters of Norwegian interest [2].

Data sets for the NMDC are collected during field experiments or regular monitoring programmes, by fixed or moving instruments mounted on land or deployed in the ocean. Each (data providing) partner has his own established data acquisition, pre-processing and data quality chain that make the data available for internal use within the partner organisation. Some organisations also offer access to their data repositories through interfaces, but not necessarily in a standard format or through a standard data exchange protocol.

1.2 Objectives of report

This report describes the format and protocols of the most important datasets that will be made available through the NMDC system. All partners contribute with a description of the datasets that they will make available.

1.3 Terminology

For terms used within this report, see the NMDC Acronyms and Concepts document [3].

1.4 References

[1] NMDC pre-project, http://talos.nodc.no/infrastruktur, 2015 (accessed 23 March 2015)

[2] Norwegian Marine Data Centre, Project plan 1. August 2012 - 31.July 2017, NFR Contract 108849.

[3] Norwegian Marine Data Centre, Acronyms and Concepts, version 0.1.

[4] Norwegian Marine Data Centre, Deliverable D3.1 Definition of data formats and metadata structure, version 0.12.

[5] Norwegian Marine Data Centre, Deliverable D3.2 Internal and external protocols, version 0.6.

1.5 Outline of report

The remainder of the report is organised as follows. Section 2 describes the formats and protocols of the most important datasets that will be made available in NMDC. Section 3 summarises the data formats and protocols used by the NMDC partners for the second set of datasets.

2 Data format and protocols

This section contains a brief description of the most important datasets that will be made available for the NMDC in a standardised format and through an open standard protocol, specifying the data formats and protocols offered by the partner owning the datasets, and including a link to the metadata record in the NMDC Metadata Catalogue.

2.1 Institute of Marine Research

2.1.1 Data description and format

2.1.1.1 Fixed oceanographic stations

The first datasets made available from IMR are CTD data from the 8 fixed oceanographic stations Ingøy, Eggum, Skrova, Bud, Sognesjøen, Utsira Indre, Utsira Ytre and Lista. The data consists of temperature and salinity from standard depths in the period 1993-2013, and is delivered as text files in ODV (Ocean Data View) format.

Measurements started between 1935 and 1946 at the different stations. The datasets will soon be updated with all standard depths data, and delivered in NetCDF format.

Links to the metadata records:

Ingøy:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=251
Eggum:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=250
Skrova:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=249
Bud:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=248
Sognesjøen:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=247
Utsira Indre:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=253
Utsira Ytre:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=252
Lista:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=245

2.1.1.2 Zooplankton

The IMR zooplankton data has been split in three datasets by sea area: Barents Sea (1981-2014), Norwegian Sea (1989-2014) and North Sea (1989-2014). The data is obtained from standard seasonal transects and a regional coverage of each area, and is delivered in Darwin Core format (http://rs.tdwg.org/dwc/).

Links to the metadata records:

Barents Sea:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=723
Norwegian Sea:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=724
North Sea:	http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=725

2.1.1.3 Mareano benthos data

3 datasets of benthos data from the Mareano project (http://www.mareano.no/en), obtained with beamtrawl, grab and RP-sledge. The data is from the period 2006-2011, and is delivered in Darwin Core format.

Links to the metadata records:

Beamtrawl: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=667
Grab: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=684
RP-sledge: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=668

2.1.1.4 Nutrients

The nutrients dataset contains nitrite, nitrate, silicate, phosphate, dissolved oxygen, chlorophyll and phaeopigment data from the Barents Sea, Norwegian Sea, Greenland Sea, North Sea and North Atlantic Ocean. Data from 1980-2010 will be delivered first, and 2011-2014 within a year. The data will be delivered in ODV format, and will probably be split in two or more datasets.

Link to metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=164

2.1.1.5 Acoustics

Analysed and interpreted data for calculation of fish density and stock size. Vertical resolution is 10-50 m, horizontal resolution is 0.1 - 5.0 nm. Data from the period 1983-2007 will be delivered in 2015, and 2008-2014 in 2016. The data will be available in LUF20 (XML) format.

Link to metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=162

2.1.1.6 Hydrography

The hydrography data will be split in two datasets: CTD data from 1995-2014, and Nansen bottle data from 1912-1946. These datasets will be available in NetCDF format. Data from 1947-1994 will be delivered later. The data is collected in the Norwegian Sea, Greenland Sea, Barents Sea and North Sea/Skagerrak.

Link to metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=179

2.1.2 Data delivery protocols

The first datasets are available for downloading using ftp. The datasets can be downloaded from: ftp://ftp.nmdc.no/nmdc/IMR/fastestasjoner/ .

The zooplankton and benthos datasets can be downloaded from our IPT server, and are also available from the ftp server:

http://gbif.imr.no/ipt/resource.do?r=imr_zooplankton_barents_sea
http://gbif.imr.no/ipt/resource.do?r=imr_zooplankton_norwegian_sea
http://gbif.imr.no/ipt/resource.do?r=imr_zooplankton_north_sea

http://gbif.imr.no/ipt/resource.do?r=imr_mareano_beamtrawl
http://gbif.imr.no/ipt/resource.do?r=imr_mareano_grab
http://gbif.imr.no/ipt/resource.do?r=imr_mareano_rpsledge

ftp://ftp.nmdc.no/nmdc/IMR/zooplankton/
ftp://ftp.nmdc.no/nmdc/IMR/mareano/

Nutrients, acoustics and hydrography data will be available for downloading using HTTP. NetCDF files will also be available from our THREDDS server.

2.2 Norwegian Defence Research Establishment (FFI)

2.2.1 Data description and format

Datasets made available from FFI are CTD/MVP data from two areas in Nordland/Troms and Oslofjorden. The data are delivered as text files in ODV (Ocean Data View) format.

Links to the metadata records:

Nordland/Troms: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=692
Oslofjorden: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=688

Datasets to be delivered are CTD/MVP data from three more areas:

- Bergen area, planned delivery was March 2015, postponed to July 2015.
- Stavanger area, planned delivery March 2016
- Trøndelag area, planned delivery March 2017

Links to metadata records:

Bergen area: http://webprod1.nodc.no:8080/infrastruktur/viewdataset.html?dataset_id=691
Stavanger area: http://webprod1.nodc.no:8080/infrastruktur/viewdataset.html?dataset_id=689
Trøndelag area: http://webprod1.nodc.no:8080/infrastruktur/viewdataset.html?dataset_id=690

2.2.2 Data delivery protocols

The data already delivered are available for downloading using ftp, from:
ftp://ftp.nmdc.no/nmdc/FFI/Hydrography/

2.3 University of Nordland

2.3.1 Data description and format

The first datasets made available from the Faculty of Biosciences and Aquaculture at UIN are CTD data from various stations in the Salten area from the period 1997-2005. The data consists of temperature, salinity, oxygen level, density and conductivity, and are delivered as HEADER and ASCII files.

Link to the metadata record:
http://webprod1.nodc.no:8080/infrastruktur/viewdataset.html?dataset_id=100

2.3.2 Data delivery protocols

The first dataset is available on:
ftp://ftp.nmdc.no/nmdc/UIN/

2.4 Norwegian Meteorological Institute

2.4.1 Data description and format

The first datasets that were made available for METNO were numerical ocean simulations from the ROMS modelling system operated operationally by METNO. Two models are published, this is the 20 km Arctic and the 4 km Nordic model. Data are available as NetCDF.

These models have now been amended with a numerical ocean model hindcast archive/ SVIM, also presented in NetCDF following the Climate and Forecast convention. The total size of this archive is 1.98 Tb.

Metadata for NMDC are currently becoming available through a dedicated OAI-PMH harvest at http://arcticdata.met.no/metamod/oai using set specification NMDC.

2.4.2 Data delivery protocols

The first datasets are available through a THREDDS Data Server using HTTP for direct file download or OPeNDAP for data streams (e.g. used in MATLAB or R).

The 20 km model is available at:

http://thredds.met.no/thredds/fou-hi/arctic20km.html

The 4 km model is available at:

http://thredds.met.no/thredds/fou-hi/nordic4km.html

The recent dataset is available through the NorStore archive pending application servers for this dataset. Direct link to the NorStore archive is provided at https://archive.norstore.no/pages/public/datasetDetail.jsf?id=10.11582/2015.00014.

2.5 Nansen Environmental and Remote Sensing Center

2.5.1 Data description and format

The first dataset made available from NERSC is a series of XBT profiles of temperature, pressure and sound speed obtained in the Fram Strait from KV Svalbard in September 2010. There are a total of 41 profiles and 71422 observations in the dataset, which is delivered in NetCDF/CF format.

The second dataset made available from NERSC is a series of XBT profiles of temperature, pressure and sound speed obtained in the Fram Strait from KV Svalbard in September 2011. There are a total of 45 profiles and 114176 observations.

NERSC will make available a third XBT dataset from the Fram Strait later in 2015, as well as a database of oceanographic data from the Nordic Seas. These datasets will be described in the next report on datasets from WP 3 (Deliverable D3.5 Dataset specifications - Final set of datasets).

Links to the metadata records:

Fram Strait/ACOBAR September 2010:
http://webprod1.nodc.no:8080/geonetwork/srv/eng/metadata.show?id=949&currTab=advanced

Fram Strait/ACOBAR September 2011:
http://webprod1.nodc.no:8080/geonetwork/srv/eng/metadata.show?id=977&currTab=advanced

2.5.2 Data delivery protocols

The two first datasets are available for downloading using OPeNDAP and HTTP. The datasets can be downloaded from:

http://thredds.nersc.no/thredds/arcticData/fram-strait.html?dataset=NERSC_ARC_PHYS_OBS_XBT_2010_v1

http://thredds.nersc.no/thredds/arcticData/fram-strait.html?dataset=NERSC_ARC_PHYS_OBS_XBT_2011_v1

2.6 Geological Survey of Norway

2.6.1 Data description and format

2.6.1.1 Seabed sediments (Grain size), Detailed M 1:5 000 - 1:50 000

The first dataset made available from NGU is seabed sediments from Norwegian fjords and coastal areas. This is detailed full-coverage (polygon) data showing the sediment grain size composition in the upper part of the seabed (uppermost 0-20 cm of the seabed). Data is only available in some particular areas where seabed mapping projects have been conducted in time period of 1999-2014, resulting in digital maps. The data coverage is indicated in metadata. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=326

2.6.1.2 2D-seismic profiles

NGU has made available a set of 2D-seismic profiles from Norwegian waters. The data is available for download as SEG-Y-files.
2D-seismic profiles were acquired using different analog and digital equipment, and therefore the penetration depth and data quality can be variable. Acquisition method, along with other line-specific metadata, is available for each profile.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=348

This metadata provides general information about 2D-seismic data, but does not give profile-specific information. For discovery and download of individual profiles it is crucial to have access to profile-specific metadata.

2.6.1.3 Seabed sediments (Grain size), Regional M 1:100 000 - 1:250 000

The data set shows the sediment grain size composition in the upper part of the seabed (uppermost 0-20 cm of the seabed). This full-coverage data (polygon features) covers parts of Norwegian Continental Shelf from Barents Sea to Skagerrak. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=327

2.6.1.4 Seabed sediments (Grain size), Overview M 1:500 000 - 1:4 000 000

The data set shows the sediment grain size composition in the upper part of the seabed. This full-coverage data (polygon features) covers Barents Sea and parts of Norwegian Continental Shelf. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=328

2.6.1.5 Seabed sediments (Genesis), Detailed M 1:10 000 - 1:50 000

The data set shows the genetical types of sediments in the upper part of the seabed. This full-coverage sediment data (polygon features) is available from a few coastal areas, where seabed mapping projects have been conducted, resulting in digital maps. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=329

2.6.1.6 Seabed sediments (Genesis), Regional M 1:100 000 - 1:250 000

The data set shows the genetical types of sediments in the upper part of the seabed. This full-coverage data (polygon features) covers parts of Norwegian Continental Shelf from Barents Sea to Skagerrak. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=331

2.6.1.7 Seabed sediments (Genesis), Overview M 1:500 000 - 1:4 000 000

The data set shows the genetical types of sediments in the upper part of the seabed. This full-coverage data (polygon features) covers parts of Mid-Norwegian Continental Shelf. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=334

2.6.1.8 Sedimentary environment

The data set is an interpretation of the depositional environment at the seabed, and shows which depositional and/or erosional processes that influence the seabed at the present. Suggested scale for usage is 1:100 000 - 1:500 000. This full-coverage data (polygon features) covers parts of Norwegian Continental Shelf in Barents Sea and Norwegian Sea. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=344

2.6.1.9 Marine landscapes

The dataset shows the subdivision of the seabed into marine landscapes defined as larger geographical areas with homogenous character. The classification is based on Naturtyper i Norge (NiN). This full-coverage data (polygon features) covers all Norwegian marine areas from Svalbard to Skagerrak. Data can be delivered as ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI.

Link to the metadata record: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=330

2.6.2 Data delivery protocols

2D seismic data can be made available through a REST-based API, a geograpical search (NGU Download Services) or hardcoded HTTP-links to each dataset (requires individual metadata).

Seabed sediments (Grain size), Detailed M 1:5 000 - 1:50 000	HTTP REST API
Seabed sediments (Grain size), Regional M 1:100 000 - 1:250 000	HTTP REST API
Seabed sediments (Grain size), Overview M 1:500 000 - 1:4 000 000	HTTP REST API
Seabed sediments (Genesis), Detailed M 1:10 000 - 1:50 000	HTTP REST API
Seabed sediments (Genesis), Regional M 1:100 000 - 1:250 000	HTTP REST API
Seabed sediments (Genesis), Overview M 1:500 000 - 1:4 000 000	HTTP REST API
Sedimentary environment	HTTP REST API
Marine landscapes	HTTP REST API

2.7 Norwegian Institute of Nature Research

2.7.1 Data description and format

2.7.1.1 Fish tag database

Data is extracted from the fish trap databases at NINA. The database includes recapture of tagged and released fish in different projects since 1938.

Dataset will be made available in Darwin Core format, but also as csv files.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=74

2.7.1.2 Otter in coastal areas - Carcasses

Registrations of otter carcasses from 1971 until today. Data will be made available in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=741

2.7.1.3 Otter in coastal areas – Distribution

Short-term studies of otter distribution. Data will be made available in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=742

2.7.1.4 Otter in coastal areas – environmental samples

Short-term studies of environmental samples on otter. Data will be made available in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=743

2.7.1.5 Otter in coastal areas - Diet

Short-term studies of otter diet. Data will be made available in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=744

2.7.1.6 Seabirds in Norway - Distribution

Observation data of seabirds on locality level. Most recent observation per Locality. Data will be published in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=99

2.7.1.7 Seabirds in Norway - Estimated population sizes

Estimated population sizes for breeding locations based on monitoring series in closest colony during breeding season. Data will be made available as WMS and in Darwin Core format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=746

2.7.1.8 Seabirds in Norway - Open sea data

The data shows a modelled distribution of seabirds in open sea, based on a two-step model on transect counts of seabirds in Norwegian waters in 10x10 square kilometre resolution. Data will be made available as WMS and ESRI shape-file.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=745

2.7.2 Data delivery protocols

2.7.2.1 Fish tag database

Dataset will be available via ftp on ftp://data.nina.no and as Darwin Core Archive on http://data.nina.no/

2.7.2.2 Otter in coastal areas - Carcasses

Dataset will be available via ftp on ftp://data.nina.no and as Darwin Core Archive on http://data.nina.no/

2.7.2.3 Otter in coastal areas - Distribution

Dataset will be available via ftp on ftp://data.nina.no and as Darwin Core Archive on http://data.nina.no/

2.7.2.4 Otter in coastal areas - Environmental samples

Dataset will be available via ftp on ftp://data.nina.no and as Darwin Core Archive on http://data.nina.no/

2.7.2.5 Otter in coastal areas - Diet

Dataset will be available via ftp on ftp://data.nina.no and as Darwin Core Archive on http://data.nina.no/

2.7.2.6 Seabirds in Norway - Distribution

Dataset will be available as Darwin Core Archive on http://data.nina.no/

2.7.2.7 Seabirds in Norway - Estimated population sizes

Dataset will be available as Darwin Core Archive on http://data.nina.no/, and as WMS on http://wms.nina.no

2.7.2.8 Seabirds in Norway - Open sea data

Dataset will be available as ESRI shapefile on http://data.nina.no, and as WMS on http://wms.nina.no

2.8 Norwegian Polar Institute

2.8.1 Data description and format

The first datasets to be delivered from NPI are physical oceanography data from the Fram Strait, including cast and mooring data gathered between 1981 and 2009. More recent observations will be appended to the dataset as the subsets come in. The data are available on Climate and Forecast (CF) conformant JSON format at http://api.npolar.no/oceanography/?q=&format=json (as well as other text formats).

Metadata is available at https://data.npolar.no/dataset/e3d4f892-2ccc-5b9a-8a6b-8330bc1ec9ee.

A subsequent dataset will be buoy data from the Arctic Ocean, north of Svalbard. Initially this dataset will contain readings from more than 40 buoys deployed during the N-ICE 2015 cruise, with

data from later cruises to be appended. This dataset will also be made available on CF conformant JSON format.

Metadata is available at https://data.npolar.no/dataset/6ed9a8ca-95b0-43be-bedf-8176bf56da80.

2.8.2 Data delivery protocols

All the initial datasets are available on CF conformant JSON format over the HTTP protocol, via a restful API, and with a GeoJSON option available. The physical oceanography data are accessible at http://api.npolar.no/oceanography/ and the buoy data at http://api.npolar.no/oceanography/buoy/?q=. Description of the API features is ongoing work.

Metadata from the NPI are available over a similar API, on DIF, ISO 19115 XML, and other formats, at http://api.npolar.no/dataset/?q=. An OAI-PMH metadata provider is available at http://api.npolar.no/dataset/oai?verb=Identify, and a subset of marine datasets can be harvested from http://api.npolar.no/dataset/oai?verb=ListIdentifiers&set=marine&metadataPrefix=dif.

2.9 University of Bergen

More than 1800 datasets (time range 1968-2010) from various moorings are available to the project office without any access restrictions. This effort is being finalized in Spring/Summer 2015.

CTD data - second set of data sets:
There are several sources in our disposition: database we used for the Atlas, digitized CTD data supplied and paper archive. The main aim is the creation of a complete GFI hydrography dataset. Data from Haakon Mosby, Helland Hansen and Armauer Hansen have been uploaded (approx. 20.000 CTD stations so far).

Description of the database and data sources can be found in Atlas description:

Korablev, A., A. Smirnov, and O. K. Baranova, 2014. Climatological Atlas of the Nordic Seas and Northern North Atlantic. D. Seidov, A. R. Parsons, Eds., NOAA Atlas NESDIS 77, 122 pp., dataset doi: 10.7289/V54B2Z78

Not all stations were identified and have vessel names and the investigation of those takes time.

The third source is the paper archive located at several places in GFI. As we decided it will be systemized and moved into one location. Preliminary, number of stations for three mentioned vessels in the established inverntory and on the list (Hydrografiske data fra Geofysisk institutt) were compared. Serious offsets in numbers were encountered.

UiB would like to know what IMR already has and we need a list (with metadata) what GFI data were digitized and what data needs to be in order to avoid duplicated efforts since the most time consumption procedure is data punching.

2.9.1 Data description and format

Data from various moorings. Data was second level quality controlled and is available as text files and NetCDF files.

Link to metadata:
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=722

2.9.2 Data delivery protocols

This time data can be picked up at our ftp site: http://people.uib.no/ako091/NMDC/Currents/

2.10 University of Oslo

2.10.1 Data description and format

The first dataset made available from the University of Oslo is hydrography and plankton data in the Oslo Fjord from 1954 to present. The data was delivered in an excel sheet.

Link to metadata:
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=753

2.10.2 Data delivery protocols

The data file was sent to NMDC, and is available at: ftp://ftp.nmdc.no/nmdc/UIO/

2.11 University of Tromsø

2.11.1 Data description and format

The Hydrographic data from TUNU/East Greenland (2002 - present) project is the first dataset available from University of Tromsø. The data was delivered in csv format.

Link to the metadata:
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=637

The second dataset is Rocky bottom fauna picture study (1976 - present.) from Svalbard and Tromsø area. Data will be available in JPEG and TIFF format.

Link to the metadata: http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=28

2.11.2 Data delivery protocols

The data file for the first dataset was sent to NMDC, and is available for downloading using ftp, at: ftp://ftp.nmdc.no/nmdc/UIT/Hydrography/TUNU/

The second dataset can be downloaded from: http://seadata.uit.no/bildedatabase/

2.12 Norwegian Institute of Water Research

2.12.1 Data description and format

Datasets from NIVA are related to the environmental monitoring programmes for the Inner Oslofjord, Outer Oslofjord and the coast of southern Norway (Swedish border to Fedje) and selected data from NIVAs “Ferrybox-programme”. Datasets contain CTD data or nutrients. Data from the Inner Oslofjord will be available first. Outer Oslofjord, Coastal monitoring and Ferrybox will follow.

Metadata is available at http://www.aquamonitor.no/geonetwork .

2.12.2 Data delivery protocols

Data will be available as csv and later as NetCDF via http://www.aquamonitor.no/thredds .

2.13 Akvaplan-NIVA

2.13.1 Data description and format

The first dataset made available from Akvaplan-niva is CTD data from Gisundet, Lenvik municipality from 2002 to 2005. The data was delivered in an excel sheet.

Link to metadata:
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=713

The second dataset from Akvaplan-niva is CTD data from Vadsø, its monthly data from 4 stations in Vadsø municipality from March 2001 to February 2002.

2.13.2 Data delivery protocols

It is planned to make datasets available via the NIVA node with metadata in http://www.aquamonitor.no/geonetwork .

2.14 Uni Research

The first datasets made available from UNI Research are current meter (ADCP) data from fixed point observatories in the Faroe Bank Channel, Faroe Shetland Channel and north of the Faroes. The data consists of current profiles from the period 1995-2012, and is delivered as text files in modified WOCE current meter data format. In addition to this data UNI Research also contribute with common data made available by University of Bergen.

2.14.1 Data description and format

Data from long-term Eulerian Acoustic Doppler Current Profilers (ADCP) data. Quality control have been carried out and documented in the Nordic WOCE ADCP Deployments report series (http:// - to be added).

Link to metadata:
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=749

2.14.2 Data delivery protocols

The ADCP data will be available from NMDC.

2.15 Norwegian Biodiversity Information Centre

2.15.1 Data description and format

All datasets are available in the Darwin Core Archive format.

Links to the metadata records:

Marine macro algae (Agder NM):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=731
Marine macro algae (UiT):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=732
Salmon (NEA):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=734
Marine molluscs (Molltax):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=733
Marine invertebrates (NTNU-VM):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=735
Marine gastropods and crustacea (UiB):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=736
Marine fish species (AO):
http://talos.nodc.no/infrastruktur/viewdataset.html?dataset_id=737

2.15.2 Data delivery protocols

The data is available in Excel/CSV format through the Species Map Service at http://artskart.artsdatabanken.no, and through the NBIC API.

3 Summary

This report contains description, format and protocols for the most important datasets that will be made available for NMDC. It includes links to the metadata at http://talos.nodc.no/infrastruktur for most partners, or at http://webprod1.nodc.no:8080/geonetwork/ for harvested metadata.

3.1 Data formats

The following formats are used to deliver the second set of datasets: ODV (Ocean Data View), NetCDF/CF, ESRI shape files, ESRI file geodatabase, ESRI personal geodatabase, SOSI, Darwin Core, text files, Excel, CSV, NetCDF, LUF20 (XML), SEG-Y, WMS, CF conformant JSON, GeoJSON, JPEG, TIFF.

The formats are described in Deliverable D3.1 - Definition of data formats and metadata structure [4].

3.2 Data delivery protocols

The following data delivery protocols are used for the second set of datasets: FTP, OPeNDAP, HTTP, HTTP REST API.

The protocols are described in Deliverable D3.2 - Internal and external protocols [5].