



## **HiSEA DELIVERABLE 3.3**

### **UPDATED DATA MANAGEMENT PLAN**

**WORK PACKAGE NUMBER: 3**

**WORK PACKAGE TITLE: SERVICE SPECIFICATIONS AND USER  
REQUIREMENTS**



| HISEA Project Information              |  |
|--|--|
| <b>Project full title</b>              | High Resolution Copernicus-Based Information Services at Sea for Ports and Aquaculture |
| <b>Project acronym</b>                 | HISEA  |
| <b>Grant agreement number</b>          | 821934   |
| <b>Project coordinator</b>             | Dr. Ghada El Serafy  |
| <b>Project start date and duration</b> | 1 <sup>st</sup> January 2019, 30 months  |
| <b>Project website</b>                 | <a href="https://hiseaproject.com/">https://hiseaproject.com/</a>                      |

| Deliverable Information                                  |  |
|--|--|
| <b>Work package number</b>                               | 3  |
| <b>Work package title</b>                                | Service Specifications and User Requirements                   |
| <b>Deliverable number</b>                                | 3.3  |
| <b>Deliverable title</b>                                 | Updated Data Management Plan                                   |
| <b>Description</b>                                       | Updated version of guidelines related to HiSea data management |
| <b>Lead beneficiary</b>                                  | HIDROMOD   |
| <b>Lead Author(s)</b>                                    | Adélio Silva (Hidromod)  |
| <b>Contributor(s)</b>                                    |  |
| <b>Revision number</b>                                   | V0.3   |
| <b>Revision Date</b>                                     | 18/03/2020   |
| <b>Status (Final (F), Draft (D), Revised Draft (RV))</b> | F  |





Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))

PU

#### Document History

| Revision | Date       | Modification    | Author        |
|----------|------------|-----------------|---------------|
| 0.1      | 18/03/2020 | Text Updating   | Adélio Silva  |
| 0.2      | 24/03/2020 | Revision        | Sandra Gaytan |
| 0.3      | 30/03/2020 | Review and edit | Bracha Ehrman |

#### Approvals

|                    | Name            | Organization | Date       | Signature (initials) |
|--------------------|-----------------|--------------|------------|----------------------|
| <b>Coordinator</b> | Ghada El Serafy | Deltares     | 31/03/2020 |                      |
| <b>WP Leaders</b>  | Adelio Silva    | Hidromod     | 31/03/2020 |                      |





## Table of Contents

|  |           |
|--|-----------|
| <b>Introduction</b>  | <b>8</b>  |
| <b>1 Data standards to be adopted in HiSea</b>                   | <b>10</b> |
| 1.1 Data quality control   | 10        |
| 1.2 In situ observations quality control                         | 11        |
| 1.3 Forecasts quality control                                    | 11        |
| <b>2 Data integration and fusion</b>                             | <b>12</b> |
| <b>3 Data management</b>   | <b>13</b> |
| 3.1 Data vocabulary  | 13        |
| 3.2 Metadata   | 13        |
| <b>4 HiSea datasets</b>  | <b>16</b> |
| 4.1 Dataset-level metadata                                       | 16        |
| 4.2 Data format standards  | 17        |
| 4.2.1 Ocean Data View data model and netCDF Format               | 17        |
| <b>5 Data privacy policy</b>                                     | <b>19</b> |
| 5.1 General principles   | 19        |
| 5.2 Use of Cookies   | 20        |
| <b>6 References</b>  | <b>21</b> |
| <b>ANNEX I: Inspire mandatory and optional metadata elements</b> | <b>23</b> |
| <b>ANNEX II: Examples of data files structure</b>                | <b>30</b> |
| Example of ISO 19139 XML encoding                                | 30        |
| Example of CMEMS data file                                       | 38        |
| <b>ANNEX III: Data Privacy Procedures</b>                        | <b>40</b> |
| Data protection principles                                       | 40        |
| Rights of Data Subjects  | 42        |
| Personal Data  | 42        |
| Processing Personal Data   | 42        |





## List of Tables

*Table 1: Metadata for spatial datasets and spatial dataset series specified in the INSPIRE Metadata Regulation [REGULATION 1205/2008/EC] .....23*

*Table 2: Mandatory and conditional theme-specific metadata for the HiSea metadata profile.....24*





## Executive Summary

This deliverable will provide an updated view of the guidelines adopted in HiSea regarding the data management. A first Deliverable (D3.2) provided an overview of the most advised procedures to adopt in managing data while, at this stage, Deliverable 3.3 reflects those being included in the HiSea service chain.

Some of the main differences between the first thought guidelines and the ones that are required to be in fact adopted result from the fact that services, and not data, are the main focus of HiSea. For this reason, some critical aspects related to data management (apart from the data privacy issues), although still existing, do not play as central a role as they would if distributing data was among the HiSea target objectives. Of course HiSea will produce data that requires managing and preserving (either resulting from modelling or locally acquired) but the target users will be mostly interested in information derived from these data sets and not in the data sets themselves.

This said, HiSea will still follow as close as possible the requirements of the H2020 Programme (H2020, 2016) regarding Data Management Plans (DMPs) which have as main objectives to *make research data findable, accessible, interoperable, and reusable (FAIR)*. In this context a DMP should include information on:

- *the handling of research data during and after the end of the project;*
- *what data will be collected, processed and generated;*
- *which methodology and standards will be applied;*
- *whether data will be shared/made open access and;*
- *how data will be curated and preserved (including after the end of the project).*

In practice the HiSea services will collect a limited amount of personal data regarding the platform users (to manage the user's rights to the platform access) and it will use data sets from existing networks and platforms (CMEMS, SEADATANET, EMODNET, NOAA, etc.) and locally produced data (operational numerical models' data, in-situ data, etc.).

This document summarizes a set of guidelines that is being adopted in HiSea to deal with these issues. These guidelines focus on the best practices in themes such as data privacy, metadata, data vocabulary, data standards, and data quality control procedures. The adoption of these guidelines will imply taking actions at different levels:

- Adopt proper data management procedures to implement metadata, provide integrated access to data in order to facilitate the integration in existing systems and assure the adoption of adequate data quality control;
- Enable integration of more data,
- improve the enhancement of the services (viewing, downloading, traceability and monitoring) to users and providers,
- provide OGC (Open Geospatial Consortium) services (WMS, WFS, etc.) to facilitate development; and the visibility of existing data and the identification of gaps.





Other relevant issues are:

- how personal data will be curated and preserved,
- what data will be collected, processed and generated,
- whether data will be shared/made open access





## Introduction

The goal of the HiSea project is to provide a set of services focused on different coastal user's needs (navigation safety, ports operations, aquaculture, etc.). These services will exploit the added value of integrated Earth Observation (EO) technologies (satellite, airborne, and ground-based), Copernicus Marine Service, and ICT to deliver customized and ready to use information. HiSea will provide an easy way to get in-situ data, local high-resolution forecasts, and products and services (e.g., meteo-oceanographic conditions at specific locations, identification of optimum or critical working windows, support to sea pollution response actions, etc.) to a broad range of different users.

This report describes the project strategies to deal with the data management, data privacy, and data quality control. According to the EU recommendations, these strategies will take into consideration the data that will be collected, processed and generated, how to handle this data during and after the project, and the standards to adopt following the FAIR (*findable, accessible, interoperable and re-usable*) data concept. For data to be findable and interoperable, the project will use;

- Data integration and fusion;
- Data quality control;
- Data privacy policy.

The issues of metadata and vocabulary are relevant to assure the interoperability and easy discovery of data. Following widely accepted standards contributes to reducing the data management effort and to improving the data quality (both at HiSea platform level and a broad data share level) while helping to make data more accessible and available.

Aiming to contribute to these objectives, HiSea is following the procedures already proposed in the most relevant EU initiatives such as CMEMS, EMODNet, and SeaDataNet, especially in regarding the standards related to vocabularies, metadata and data formats.

In practice, the gridded data sets addressing either dynamic data sets (like CMEMS) or static data sets (like EMODnet) follow the procedures adopted by these two services. For time-series data, HiSea assumes the SeaDataNet guideline, and NetCDF-CF format is the standard.

The data integration and fusion policies are also relevant issue for the project. Data integration and fusion describe the best strategies to merge datasets obtained from different data sources, to build the best available datasets, or fuse various data sources to produce aggregated data. Properly addressing this issue represents a valuable contribution to improve data accuracy and the robustness of models' initial and boundary conditions. It can also provide the user with comprehensive data that merge different data sets based on reliable criteria.

The data quality control related to the quality of observed in-situ data (e.g., tidal gauges, wave buoys, weather stations, etc.) or the modeled forecasts is another relevant aspect that will be addressed by HiSea. In the case of locally acquired data, automatic procedures will run regularly to detect and remove abnormal values from observed







datasets. In the case of the models, the system automatically compares results with observations (e.g., buoys and CMEMS grid observation products). The statistical analysis will be provided daily to the end-users.

Regarding data privacy (data protection and the rights of platform end-users, customers and business contacts), HiSea will assure the respect of personal data under the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) which replaced Directive 95/46/EC on May 25, 2018. 'Personal data' means any information, private or professional, which relates or can be related to an identified or identifiable natural person (for the full definition, see Article 2(a) of EU Directive 95/46/EC).

The following sections provide a more detailed overview of the state of the art and the procedures to be adopted in HiSea. Note that at this stage, this document mostly focuses on defining the guidelines to be followed throughout HiSea platform development. The practical implementation of these guidelines which will be the subject of a later document.





# 1 Data standards to be adopted in HiSea

## 1.1 Data quality control

HiSea addresses the issue of data quality control by following the state of the art recommendations of different projects such as SeaDataNet or AtlantOS. SeaDataNet produced a comprehensive document with a set of guidelines for marine data quality control. According to this document, data quality control should:

*“To ensure the data consistency within a single data set and within a collection of data sets and to ensure that the quality and errors of the data are apparent to the user who has sufficient information to assess its suitability for a task.”* If done correctly, quality control brings several vital advantages (SeaDataNet, 2010):

- **Maintaining Common Standards:** There is a minimum level to which all oceanographic data should be quality controlled. There is little point banking data just because it was collected; additional information must qualify the data with methods of measurement and subsequent data processing. Imposing standards on the quality and long-term value of the data that are accepted (Rickards, 1989). If there are guidelines available to this end, the result is that data are at least maintained to this degree, keeping common standards to a higher level.
- **Acquiring Consistency:** Data stored in different data centers should be as consistent with each other as possible, making it more accessible to the external user. Searches for data sets are more successful if users can identify the specific data they require quickly, even if the origins of the data are very different on a national or even international level.
- **Ensuring Reliability:** Data centers, like other organizations, build reputations based on the quality of the services they provide. Controlling the data to a ‘universal’ standard is the best way to ensure reliability. Many national and international programs or projects carry out investigations across a broad field of marine science that require complex information on the marine environment. Large scale commercial projects, such as those involved with oil and gas and fishing industries, also perform this type of quality control. Users make significant decisions, and form theories on the assumption that data is reliable and compatible, even when they come from many different sources.

The HiSea services data flux is automatically managed by the HiSea platform. The data quality control starts with the execution of automatic procedures (independently of the adoption of more complex processes). The data quality control methodology focuses on in situ observations and modeled forecasts using two perspectives: the data Quality Assurance and the data Quality Control.

Quality Assurance (QA) is a set of review and audit procedures implemented by personnel or an organization (ideally) not involved with normal project activities. It monitors and evaluates the project maximizing the probability of obtaining minimum standards of quality by verifying that data quality objectives are met within the given constraints. In the limit QA procedures may result in corrections to data which may either be applied in bulk (i.e., all data from an instrument during a deployment period) or to particular data points. The application of QA corrections will automatically result in the data reflagging as ‘corrected.’





Quality Control (QC) is a process of routine technical operations to measure, annotate (i.e., flag), and control the quality of the produced data. These operations may include spike checks, out-of-range checks, missing data checks, as well as others. QC should:

- Provide routine and consistent checks to ensure data integrity, correctness, and completeness;
- Identify and address possible errors and omissions;
- Document all QC activities.

QC operations include automated checks on data acquisition and calculations using approved standardized procedures. Higher-tier QC activities can consist of additional technical review and correction of the data by human inspection. QC procedures are essential for:

- Detecting missing mandatory information;
- Detecting errors made during the transfer and reformatting processes;
- Detecting duplicates;
- Detecting remaining outliers (spikes, out of scale data, vertical instabilities, etc.).

Project SeaDataNet has compiled a guideline of recommended QC procedures after reviewing NODC schemes and other known schemes (e.g., WGMDM guidelines, World Ocean Database, GTSP, Argo, WOCE, QARTOD, ESEAS, SIMORC, etc.). At present, it follows the QC methods proposed by SeaDataNet for CTD (temperature and salinity profiles), current meter data (including ADCP), wave data, and sea level data. SeaDataNet is also developing efforts for extending the guideline with QC methods for surface underway data, nutrients, geophysical data, and biological data.

ANNEX I provides a detailed description of the implementation process procedure to be followed for QA/QC in HiSea.

## 1.2 In situ observations quality control

For observations, quality control has two phases. Automatic checks execute during the download of in-situ observations, such as those proposed by SeaDataNet (2010) (e.g., global range test, date, and time). Only valid data that passes this control is stored. In the second phase, a tool executes periodically to perform a scientific quality control check (SeaDataNet, 2010). This quality control aims to detect spikes, high-frequency filter noise (e.g., moving average or P50), data with abnormal variability in time, etc. Specific tools will be running automatically with this aim.

## 1.3 Forecasts quality control

The modeled forecasts quality control consists of comparing time-series forecasts with in situ observations (e.g., wave-buoys, tidal gauge, weather stations, etc.) through automatically-run algorithms. Also, gridded data forecasts may be compared automatically with observed data (e.g., CMEMS gridded data observations). This process calculates several statistic parameters (e.g., correlation coefficient, bias, RMSE, skill, etc.) to assess the quality of forecasts.





## 2 Data integration and fusion

It is not obvious which are the best strategies to merge datasets from different data sources, build the best available datasets, or fuse various data sources to produce aggregated data, indices, and products. A possible solution when we have different solutions with different resolutions for the same area is to make a fusion of these data and offer a unique integrated dataset. Another option is to provide all datasets separately with the option of an integrated solution. No matter which solution is adopted, the final objective of the data integration and fusion is to:

- improve data accuracy and robustness of the model's initial and boundary conditions
- to provide to the users' comprehensive data that results from the merge of different data sets using clear criteria.

For example, let's assume a user is interested in wave data for a specific location. Still, he realizes that for the period in which he has an interest, time series from different wave-buoys exist. He may be interested in creating a unique time series that results from merging in a compatible fashion the different time-series. This process may require complex actions regarding the levels of accuracy of the various measuring devices, the measuring time rate and units, etc.





## 3 Data management

This section summarizes the concept of metadata that will be adopted by the HiSea data platform, following the INSPIRE data specification template in its relevant parts, i.e., dataset-level, services metadata, and data quality. It also contains detailed technical documentation on the XML source-code level and therefore provides specific guidelines to create and maintain metadata in the XML format correctly.

### 3.1 Data vocabulary

The use of standard vocabularies in all meta-databases and data formats is an essential prerequisite for the consistency and interoperability with existing Earth Observing systems and networks. Standard vocabularies consist of lists of standardized terms of reference covering a broad spectrum of disciplines of relevance to the oceanographic and wider community. The use of standardized terms of reference solves the problem of ambiguities related to the data structure, organization, and format is solved. Therefore, standard algorithms for data processing may be applied. This process allows the interoperability of datasets in terms of their manipulation, distribution, and long-term reuse.

HiSea will adopt an Essential Variables list of terms (aggregated level) that was defined and published in June 2016 on the NERC/BODC Vocabulary Server<sup>1</sup>.

This new vocabulary maps to the standards recommended for HiSea parameter metadata: P01 (parameter), P07 (CF variable), P06 (units) from SeaDataNet controlled vocabularies managed by NERC/BODC and the internationally assured AphiaID from the World Register of Marine Species (WoRMS)<sup>2</sup>.

### 3.2 Metadata

Metadata refers to the description of datasets and services in a compliant form as defined by the Directive 2007/2/EC (INSPIRE) and Commission Regulation No 1205/2008.

We can consider Metadata as the **data about the data**. Metadata describes how, when, and by whom a set of data or services was collected or prepared, and how the data is formatted, or the service is available. Metadata is essential for understanding the stored information and has become increasingly important. Metadata consists of structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, and manage an information resource.

Metadata is also data about services. Metadata describes the content, quality, condition, and other characteristics of a data set or the capabilities of service. Creating metadata or data documentation for geospatial datasets is crucial to the data development process. Metadata is a valuable part of a dataset and can be used to:

- **Organize** data holdings (Do you know what you have?);

---

<sup>1</sup> [https://www.bodc.ac.uk/data/codes\\_and\\_formats/vocabulary\\_search/A05/](https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/A05/)

<sup>2</sup> <http://www.marinespecies.org/aphia.php?p=webservice>





- Provide **information about** data holdings (Can you describe to someone else what you have?);
- Provide information **to data users** (Can they figure out if your data are useful to them?);
- **Maintain the value** of your data (Can they figure out if your data sets will be valuable 20 years from now?).

In the geographical domain, we can have a description of spatial data (**spatial data** metadata), a service (**service** metadata), or a particular analysis process (**process** metadata). Most of the standardization work is for metadata applied to data sets; however, service and process metadata has become increasingly important. Metadata is used in discovery mechanisms to bring spatial information providers and users together. The following mechanisms are recognized:

- **Discovery**: which data source contains the relevant information for the user;
- **Exploration (or evaluation)**: do I find within the data sources the right information to suit my information needs?;
- **Exploitation (use and access)**: how can I obtain and use the data sources?

Each mechanism has its use of metadata. The selected standards should fulfill the needs to carry out services using these mechanisms. Metadata is required to provide information about an organization's data holdings. Data resources are a major asset, and the information of what datasets exist within different organizations, particularly in the public sector, is required to improve efficiencies and reduce data duplication. Data catalogs and data discovery services enable potential users to find, evaluate, and use that data, thereby increasing its value. This process is also becoming principal at the European level. Furthermore, metadata received from an external source may require further information supplied to metadata to allow easy process and interpretation. In this context, for all types of data, the following information is required (SeaDataNet, 2010):

- **Where** the data were collected: location (preferably as latitude and longitude) and depth/height;
- **When** the data were collected (date and time in UTC or specified local time zone);
- **How** the data were collected (e.g., sampling methods, instrument types, analytical techniques). How do we organize the data (e.g., in terms of station numbers, cast numbers);
- **Who** collected the data, including the name and institution of the data originator(s) and the principal investigator;
- **What** processes were applied to the data (e.g., details of processing and calibrations applied, algorithms used to compute derived parameters);
- **Watchpoints** for other users of the data (e.g., problems encountered and comments on data quality).

The ICES Working Group on Data and Information Management (WGDIM) has developed data type guidelines that itemize the elements required for thirteen different data types. These guidelines were created with the expertise of the oceanographic data centers of ICES Member Countries. The guidelines describe the elements of data and





metadata considered relevant to the ocean research community. These guidelines target most physical-chemical-biological data types collected on oceanographic research vessel cruises. Each guideline addresses the data and metadata requirements of a specific data type, covering three main areas:

- What the data collector should provide to the data center (e.g., collection information, processing, etc.);
  - How the data center handles data supplied (e.g., value-added, quality control, etc.);
  - What the data center can provide in terms of data, referral services, and expertise back to the data collector.
- Annex 1 this document includes a selection of these guidelines.





## 4 HiSea datasets

This section describes the structure and the content for the proposed HiSea metadata profile on the dataset-level. It includes general guidelines for the metadata from two points of view, the HiSea metadata and HiSea data quality issues. The metadata model adopted by HiSea is described in more detail in Annex I.

### 4.1 Dataset-level metadata

Metadata can be reported for each spatial object (spatial object-level metadata) or once for a complete dataset or dataset series (dataset-level metadata). For data quality elements at the spatial object level, the documentation shall refer to the appropriate definition in the Data Quality Info section of this document. This section only specifies the dataset-level metadata elements.

For dataset-level metadata elements, in particular on data quality and maintenance, using a specific scope allows the definition of metadata at sub-dataset level, e.g., separately for each spatial object type. The rules to encode metadata using ISO 19115/19139 are:

- The scope element (of type DQ\_Scope) of the DQ\_DataQuality subtype will encode the scope;
- For the level element of DQ\_Scope the following values should be used: series, dataset, featureType;
- If the level is featureType<sup>3</sup> then the levelDescription/MD\_ScopeDescription/features element (of type Set <GF\_FeatureType>) shall be used to list the feature type names.

Mandatory or conditional metadata elements are specified in the next sub-section, while optional metadata elements are specified in subsequent sub-Section. The tables describing the metadata elements contain the following information:

- the first column provides a reference to a more detailed description;
- the second column specifies the name of the metadata element;
- the third column specifies the multiplicity;
- the fourth column specifies the condition under which the given item becomes mandatory (only for the first and second tables).

---

<sup>3</sup> The value *featureType* is used to denote spatial object types







## 4.2 Data format standards

### 4.2.1 Ocean Data View data model and netCDF Format

As part of the HiSea services, data sets will be accessible via download services. Delivery of data to users requires common data transport formats, which interact with other standards (Vocabularies, data quality control). The SeaDataNet project decided that Ocean Data View (ODV) and NetCDF format are mandatory.

The CF metadata conventions (<http://cf-pcmdi.llnl.gov/>) promotes the processing and sharing of files created with the [NetCDF API](#). The conventions define metadata that provides a definitive description of what the data in each variable represents and the spatial and temporal properties of the data. This option enables users of data from different sources to decide which quantities are comparable and facilitates building applications with powerful extraction, re-gridding, and display capabilities.

The standard is both mature and well-supported by formal governance for its further development. The documentation of this standard is available on a PDF manual accessible from a link from the CF metadata homepage (<http://cf-pcmdi.llnl.gov/>). Note that CF is a developing standard, and consequently, access via the website rather than through a direct URL to the document will ensure access to the latest version. The version used in this document was version 1.6 of the conventions dated 5 December 2011.

The approach taken with the development of the SeaDataNet profile based on CF 1.6 was to classify data based on feature types and produce a SeaDataNet specification for storage of each of the following:

- **Point time series**, such as current-meter or sea level data, have *row\_groups* made up of measurements from a given instrument at different times. The metadata date and time is the time when the first measurement occurred. The primary variable is time (UT) encoded either as:
  - A real number, representing the Chronological Julian Date, defined as the time elapsed in days from 00:00 on January 1<sup>st</sup>, 4713 BC. When choosing this option, the column must have the heading 'Chronological Julian Date [days].'
  - A string containing the UT date and time to sub-second precision corresponding to ISO8601 syntax (YYYY-MM-DDThh:mm:ss.sss) for example 2009-02-12T11:21:10.325. When choosing this option, the column must have the heading 'time\_ISO8601'. If the time is not known to sub-second precision, then use the ISO8601 form appropriate to the known accuracy. For example, a timestamp to the accuracy of one hour is 2009-02-12T11:00. A timestamp to a precision of a day is 2009-02-12.Rows within the *row\_group* are ordered by increasing time. Note that the z co-ordinate (e.g., instrument depth), essential for many types of time series data, needs to be stored as a data variable and could have the same value throughout the *row\_group*.
- **Profile data**, such as CTD or bottle data, have *row\_groups* made up of measurements at different depths. The metadata date and time use the time when the profile measurement started. The primary variable is the 'z co-





ordinate,' which for SeaDataNet is either depth in meters or pressure in decibars. Rows within the row\_group are ordered by increasing depth.

- **Trajectories**, such as underway data, have row\_groups made up of a single measurement, making the metadata time and positions the Spatio-temporal co-ordinate channels. The primary variable is the 'z co-ordinate,' which for SeaDataNet is standardized as depth in meters. Rows within the row\_group are ordered by increasing time;
- **TimeSeriesProfile** ( $x, y, z$  fixed;  $t$  variable), but some variables can have measures at different depths at the same time  $var=f(t, z)$ . The specification given is for storage of time series profiles such as moored ADCP.
- **TrajectoryProfile** ( $x, y, z, t$  all variable), but some variables can have measures at different depths at the same time  $var=f(t, z)$ . The specification given is for storage of trajectory profiles such as shipborne ADCP.

The specification was then developed through discussions on a collaborative e-mail list involving participants in SeaDataNet, MyOcean, USNODC, NCAR, and AODN. The working objective focussed on producing profiles with the following properties:

- CF 1.6 conformant;
- Have maximum interoperability with CF 1.6 implementations in use by MyOcean (OceanSITES conventions), USNODC (USNODC NetCDF templates) and two contributors to AODN (IMOS and METOC);
- Include storage for all labels, metadata, and standardized semantic mark-up included in the SeaDataNet ODV format files for the equivalent feature type.

A significant list discussion focussed on the version of NetCDF should be used for SeaDataNet. NetCDF 4 was selected as the best option wherever possible, but NetCDF 3, although strongly discouraged, should not be forbidden.

On ANNEX II some examples of the structure of these files are presented.





## 5 Data privacy policy

### 5.1 General principles

The basic principles regulated by the data protection that apply to Hisea are:

- ✓ HiSea only holds the necessary personal data to offer services provided by its platform.
- ✓ Data is only used for the purposes described in the Data Protection Register Form and the Informed Consent Form.
- ✓ Personal data is only stored for as long as necessary. Once data are no longer needed, it will be deleted from HiSea records by the HiSea platform Administrator (namely, the CLS Chief Technical Officer (CTO) / IT platform manager). More specifically, in case a specified period (one year) is passed without the entry of an end-user in the platform, CLS will alert him through a standardized electronic message on the destruction of personal data.
- ✓ Personal data storage will be secured to ensure that data are not accessible to unwanted third parties and is protected against disaster and risk.
- ✓ HiSea will regularly email website news and information updates only to those end-users and customers who have specifically subscribed to our email service. All subscription emails sent by the HiSea platform will contain clear information on how to unsubscribe from our email service.
- ✓ In any event, no personal data will be shared with any third party for direct marketing. HiSea will never sell, rent, or exchange mailing lists of personal data.
- ✓ All HiSea partners shall comply with the data protection and privacy laws applicable in their country of origin, including their national laws applicable to exporting data into the EU.
- ✓ HiSea partners from non-EU countries have provided signed declarations that they will meet all relevant H2020 ethical standards and regulations. Exporting personal data from the EU to non-EU countries must comply with the applicable EU rules on cross-border transfer of personal data.
- ✓ In accordance with the Privacy and Electronic Communications (EC Directive) Regulations 2003, HiSea will never send bulk unsolicited emails, (popularly known as Spams) to any email addresses.
- ✓ HiSea may send emails to existing end-users and customers or prospective end-users and customers having enquired or registered in HiSea platform, regarding products or services directly provided by HiSea platform.
- ✓ All emails sent by HiSea will be clearly marked as originating from this platform. All such emails will also include clear instructions on how to unsubscribe from HiSea email services. Such instructions will either include a link to a page to unsubscribe or a valid email address to which the user should reply, with “unsubscribe” as the email subject heading.





## 5.2 Use of Cookies

Cookies are small text files placed on your computer by websites that you visit. They are commonly used to make websites work, or work more efficiently, as well as providing information to the owner of the site.

HiSea platform may generate cookies to work more efficiently. These will enhance features such as platform search and optimized page loading.

HiSea may use Google Analytics to collect quantitative information on the platform's performance and end-users' interaction with the platform. HiSea will use this information to improve the service and experience offered by the platform. The use of Social Media buttons on some of the pages links to third party websites and services, like Facebook and Twitter also creates cookies. These services use cookies when clicking the button. Privacy policies will be available for all these services, and users should be able to read them to be informed on how their information is used, and how they can opt-out, should they wish to do so.





## 6 References

AtlantOS, 2016, *Data Management Handbook*

Carval Thierry, 2016, *Catalogue of data and platforms at Network GDAC level, including the example of Copernicus In Situ TAC*. IMN/IDM/ISI/FM/16-024. <http://doi.org/10.13155/45063>

Columbus Project, 2017, *Use and sharing of marine observations and data by industry*

Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata

Corrigendum to Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

Draft Commission Regulation (EU) No/... of [L] Implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial s and services

EN ISO 19113:2002 Geographic information – Quality principles

EN ISO 19114:2003/Corr. 1 :2005 Geographic information – Quality evaluation procedures

EN ISO 19115:2003/19115-1 :2014: Geographic information – Metadata

EN ISO 19115:2003/Cor. 1:2006

EN ISO 19108: Geographic information – Temporal schema

EN ISO 19108:2002/Cor. 1:2006

EN ISO 19119:2005/19119 :2016: Geographic information - Services

EN ISO 8601:2004 Data elements and interchange formats – Information exchange – Representation of dates and times

INSPIRE DS-D2.3, Definition of Annex Themes and Scope, v3.0,

INSPIRE DS-D2.5, Generic Conceptual Model, v3.4,

INSPIRE DS-D2.6, Methodology for the development of data specifications, v3.0,

INSPIRE DS-D2.7, Guidelines for the encoding of spatial data, v3.1,

INSPIRE DS-D2.8.I.7, INSPIRE Data Specifications on Transport Networks – Guidelines, v3.0

INSPIRE DS-D2.8.I.8, INSPIRE Data Specifications on Hydrography – Guidelines, v3.0

INSPIRE DS-D2.8.I.9, INSPIRE Data Specifications on Protected Sites – Guidelines, v3.0

INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119, v. 1.1





ISO/TS 19138:2006 Geographic information – Data quality measures

ISO/TS 191139:2007 Geographic information – Metadata – XML Schema Implementation

SeaDataNet, 2010, *Data Quality Control Procedures*

SeaDataNet, 2017, *Data File Formats: ODV, Medatlas, netCDF. Deliverable D8.5*





## ANNEX I: Inspire mandatory and optional metadata elements

The metadata describing a spatial dataset or a spatial dataset series shall comprise the metadata elements required by Commission Regulation No 1205/2008/EC (implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata) for spatial datasets and spatial dataset series (*cf. Table 1*) as well as the HiSea - specific mandatory metadata elements specified in Table 2 and optionally elements specified in **Error! Reference source not found.**

*Table 1: Metadata for spatial datasets and spatial dataset series specified in the INSPIRE Metadata Regulation [REGULATION 1205/2008/EC]*

| Metadata           |                               |              |   |
|--------------------|-------------------------------|--------------|---|
| Regulation Section | Metadata element              | Multiplicity | Condition   |
| 1.1                | Resource title                | 1            |   |
| 1.2                | Resource abstract             | 1            |   |
| 1.3                | Resource type                 | 1            |   |
| 1.4                | Resource locator              | 0..*         | Mandatory if a URL is available to obtain more information on the resource, and/or access related services. |
| 1.5                | Unique resource identifier    | 1..*         |   |
| 1.7                | Resource language             | 0..*         | Mandatory if the resource includes textual information.   |
| 2.1                | Topic category                | 1..*         |   |
| 3                  | Keyword                       | 1..*         |   |
| 4.1                | Geographic bounding box       | 1..*         |   |
| 5                  | Temporal reference            | 1..*         |   |
| 6.1                | Lineage                       | 1            |   |
| 6.2                | Spatial resolution            | 0..*         | Mandatory for datasets and dataset series if an equivalent scale or a resolution distance can be specified. |
| 7                  | Conformity                    | 1..*         |   |
| 8.1                | Conditions for access and use | 1..*         |   |
| 8.2                | Limitations on public access  | 1..*         |   |
| 9                  | Responsible organisation      | 1..*         |   |
| 10.1               | Metadata point of contact     | 1..*         |   |
| 10.2               | Metadata date                 | 1            |   |



|      |                   |   |  |
|------|-------------------|---|--|
| 10.3 | Metadata language | 1 |  |
|------|-------------------|---|--|

Table 2: Mandatory and conditional theme-specific metadata for the HiSea metadata profile

| HiSea metadata profile Section | Metadata element            | Multiplicity | Condition  |
|--------------------------------|-----------------------------|--------------|--|
| 1                              | Coordinate reference system | 1            |  |
| 2                              | Encoding                    | 1..*         |  |
| 3                              | Character Encoding          | 0..*         | Mandatory, if a non-XML-based encoding is used that does not support UTF-8 |
| 4                              | Source title                | 0..1         | Mandatory, if a source title can be specified                              |
| 5                              | Source date of mapping      | 0..*         | Mandatory, if source date of mapping can be specified                      |

### Coordinate Reference System

| Metadata element name          | Coordinate Reference System  |
|--------------------------------|--|
| Definition                     | Description of the horizontal coordinate reference system used in the dataset.   |
| ISO 19115 number and name      | 13. referenceSystemInfo  |
| ISO/TS 19139 path              | referenceSystemInfo  |
| INSPIRE obligation / condition | mandatory  |
| INSPIRE multiplicity           | 1  |
| Data type (and ISO 19115 no.)  | 189. MD CRS  |
| Domain                         | Either the referenceSystemIdentifier (RS_Identifier) or the projection (RS_Identifier), ellipsoid (RS_Identifier) and datum (RS_Identifier) properties shall be provided.  |
| Implementing instructions      | –  |
| Example                        | referenceSystemIdentifier:<br>code: ETRS_89<br>codeSpace: INSPIRE RS registry  |
| Example XML encoding           | <pre>&lt;gmd:MD_Metadata .. .. &lt;gmd:referenceSystemInfo&gt; &lt;gmd:MD_ReferenceSystem&gt; &lt;gmd:referenceSystemIdentifier&gt; &lt;gmd:RS_Identifier&gt; .. &lt;gmd:code&gt; &lt;gco:CharacterString&gt;3035&lt;/gco:CharacterString&gt; &lt;/gmd:code&gt; &lt;gmd:codeSpace&gt; &lt;gco:CharacterString&gt;EPSG&lt;/gco:CharacterStr</pre> |





|         |  |
|---------|--|
|         | <pre> ing&gt; &lt;/gmd:codeSpace&gt; .. &lt;/gmd:RS_Identifier&gt; &lt;/gmd:referenceSystemIdentifier&gt; &lt;/gmd:MD_ReferenceSystem&gt; &lt;/gmd:referenceSystemInfo&gt; .. &lt;/gmd:MD_Metadata&gt; </pre>                                |
| Comment | To ensure unified approach of the codes across the whole Europe, we recommend an approach that uses standardize EPSG codes for the expression of the reference system (like 4326 stands for the coordinate system WGS84 or 3035 for ETRS89). |

### Encoding

| Metadata element name          | Encoding  |
|--------------------------------|---|
| Definition                     | Description of the computer language construct that specifies the representation of data objects in a record, file, message, storage device or transmission channel.  |
| ISO 19115 number and name      | 271. distributionFormat   |
| ISO/TS 19139 path              | distributionInfo/MD_Distribution/distributionFormat   |
| INSPIRE obligation / condition | mandatory   |
| INSPIRE multiplicity           | 1..*  |
| Data type (and ISO 19115 no.)  | 284. MD_Format  |
| Domain                         | The following property values shall be used for default and alternative encodings specific and developed in the HISEA project:<br>Default Encoding<br>name: HISEA<br>version: version 3.0; GML, version 3.2.1   |
| Implementing instructions      | –   |
| Example                        | name: HISEA<br>version: version 3.0, GML, version 3.2.1   |
| Example XML encoding           | <pre> &lt;gmd:MD_Metadata .. .. &lt;gmd:distributionInfo&gt; &lt;gmd:MD_Distribution&gt; &lt;gmd:distributionFormat&gt; &lt;gmd:MD_Format&gt; &lt;gmd:name&gt; &lt;gco:CharacterString&gt;<b>GML</b>&lt;/gco:CharacterString&gt; &lt;/gmd:name&gt; &lt;gmd:version&gt; &lt;gco:CharacterString&gt;<b>3.2.1</b>&lt;/gco:CharacterString&gt; &lt;/gmd:version&gt; &lt;/gmd:MD_Format&gt; </pre> |



|         |  |
|---------|--|
|         | <pre> .. &lt;/gmd:distributionFormat&gt; &lt;/gmd:MD_Distribution&gt; &lt;/gmd:distributionInfo&gt; .. &lt;/gmd:MD_Metadata&gt; </pre> |
| Comment | -  |

### Character Encoding

| Metadata element name          | Character Encoding   |
|--------------------------------|--|
| Definition                     | Full name of the character coding standard used for the dataset  |
| ISO 19115 number and name      | 4. characterSet  |
| ISO/TS 19139 path              | identificationInfo/*/characterSet  |
| INSPIRE obligation / condition | Mandatory, if a non-XML-based encoding is used that does not support UTF-8   |
| INSPIRE multiplicity           | 0..*   |
| Data type (and ISO 19115 no.)  | 40. MD_CharacterSetCode  |
| Domain                         | Codelist (See B.5.10 of ISO 19115)   |
| Implementing instructions      | -  |
| Example                        | -  |
| Example XML encoding           | <pre> &lt;gmd:MD_Metadata .. .. &lt;gmd:identificationInfo&gt; .. &lt;gmd:characterSet&gt; &lt;gmd:MD_CharacterSetCode codeList="<a href="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodetests.xml#MD_ClassificationCode">http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodetests.xml#MD_ClassificationCode</a>" codeListValue="utf8"&gt;utf8&lt;/gmd:MD_CharacterSetCode&gt; &lt;/gmd:characterSet&gt; .. &lt;/gmd:identificationInfo&gt; .. &lt;/gmd:MD_Metadata&gt; </pre> |
| Comment                        | -  |

### Source title

| Metadata element name | Source title |
|-----------------------|--------------|
|-----------------------|--------------|



|                                |   |
|--------------------------------|---|
| Definition                     | Full name by which the input data source of the described dataset is known  |
| ISO 19115 number and name      | 360. title  |
| ISO/TS 19139 path              | dataQualityInfo/*/lineage/*/source/*/sourceCitation/*/title   |
| INSPIRE obligation / condition | 0..1  |
| INSPIRE multiplicity           | Mandatory, if a source title can be specified.  |
| Data type (and ISO 19115 no.)  | 96. sourceCitation  |
| Domain                         | Free text   |
| Implementing instructions      | Full name of the source dataset without an expression of the scale – a scale (in a form of a denominator) should be expressed in section 4.2.4.   |
| Example                        | Map of pedogenetic associations, HISEA Dataset  |
| Example XML encoding           | <pre> &lt;gmd:MD_Metadata .. .. &lt;gmd:dataQualityInfo&gt; .. &lt;gmd:lineage&gt; &lt;gmd:LI_Lineage&gt; &lt;gmd:source&gt; &lt;gmd:LI_Source&gt; .. &lt;gmd:sourceCitation&gt; &lt;gmd:CI_Citation&gt; &lt;gmd:title&gt; &lt;gco:CharacterString&gt;<b>HISEA maritime dataset</b>&lt;/gco:CharacterString&gt; &lt;/gmd:title&gt; .. &lt;/gmd:CI_Citation&gt; &lt;/gmd:sourceCitation&gt; &lt;/gmd:LI_Source&gt; &lt;/gmd:source&gt; .. &lt;/gmd:LI_Lineage&gt; &lt;/gmd:lineage&gt; &lt;/gmd:dataQualityInfo&gt; .. &lt;/gmd:MD_Metadata&gt; </pre> |
| Comment                        | –   |

#### Source date of mapping





| Metadata element name          | Source date of mapping   |
|--------------------------------|--|
| Definition                     | Reference date for the input data source of the described dataset.   |
| ISO 19115 number and name      | 362. date  |
| ISO/TS 19139 path              | dataQualityInfo/*/lineage/*/source/*/sourceCitation/*/date/*/date  |
| INSPIRE obligation / condition | Mandatory, if a source date of mapping can be specified  |
| INSPIRE multiplicity           | 0..*   |
| Data type (and ISO 19115 no.)  | 393. CI_Date   |
| Domain                         | Described in ISO 19108 and ISO 8601  |
| Implementing instructions      | This metadata should be filled with these elements:<br>- date (e.g. 2010, 2010-04, 2010-04-09)<br>- dateType (i.e. creation, revision or publication)  |
| Example                        | –  |
| Example XML encoding           | <pre> &lt;gmd:MD_Metadata .. .. &lt;gmd:dataQualityInfo&gt; .. &lt;gmd:lineage&gt; &lt;gmd:LI_Lineage&gt; &lt;gmd:source&gt; &lt;gmd:LI_Source&gt; .. &lt;gmd:sourceCitation&gt; &lt;gmd:CI_Citation&gt; .. &lt;gmd:date&gt; &lt;gmd:CI_Date&gt; &lt;gmd:date&gt; &lt;gco&gt;Date&gt;2010-04-09&lt;/gco&gt;Date&gt; &lt;/gmd:date&gt; &lt;gmd:dateType&gt; &lt;gmd:CI_DateTypeCode codeList="http://www.isotc211.org/2005/resources/codeList.xml#CI_ DateTypeCode" codeListValue="creation" /&gt; &lt;/gmd:dateType&gt; &lt;/gmd:CI_Date&gt; &lt;/gmd:date&gt; .. &lt;/gmd:CI_Citation&gt; &lt;/gmd:sourceCitation&gt; &lt;/gmd:LI_Source&gt; &lt;/gmd:source&gt; &lt;/gmd:LI_Lineage&gt; </pre> |





|         |  |
|---------|--|
|         | <pre>&lt;/gmd:lineage&gt; &lt;/gmd:dataQualityInfo&gt; .. &lt;/gmd:MD_Metadata&gt;</pre> |
| Comment | -  |





## ANNEX II: Examples of data files structure

### Example of ISO 19139 XML encoding

The XML source code written below represents an example of a full HiSea metadata profile – i.e. all metadata elements according to the Commission Regulation No 1205/2008/EC (implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata) for spatial datasets and spatial dataset series as well as all HiSea metadata elements described in the metadata profile documentation written above.

```
<?xml version="1.0" encoding="utf-8" ?>
<gmd:MD_Metadata xsi:schemaLocation="http://www.isotc211.org/2005/gmd
http://www.isotc211.org/2005/gmd/metadataEntity.xsd" xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gmx="http://www.isotc211.org/2005/gmx"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:gml="http://www.opengis.net/gml">
<gmd:fileIdentifier>
<gco:CharacterString>0de9b6b7-641c-441b-bf37-f03db0960cb8</gco:CharacterString>
</gmd:fileIdentifier>
<gmd:language>
<gmd:LanguageCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmxCodelists.xml#LanguageCode" codeListValue="eng">eng</gmd:LanguageCode>
</gmd:language>
<gmd:hierarchyLevel>
<gmd:MD_ScopeCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_ScopeCode" codeListValue="dataset">dataset</gmd:MD_ScopeCode>
</gmd:hierarchyLevel>
<gmd:contact>
<gmd:CI_ResponsibleParty>
<gmd:organisationName>
<gco:CharacterString>HISEA Project</gco:CharacterString>
</gmd:organisationName>
<gmd:contactInfo>
<gmd:CI_Contact>
<gmd:address>
<gmd:CI_Address>
<gmd:electronicMailAddress>
<gco:CharacterString>HISEA@HISEA.eu</gco:CharacterString> </gmd:electronicMailAddress>
</gmd:CI_Address>
</gmd:address>
<gmd:onlineResource>
<gmd:CI_OnlineResource>
<gmd:linkage>
<gmd:URL>http://HISEAplatform.eu</gmd:URL>
</gmd:linkage>
</gmd:CI_OnlineResource>
</gmd:onlineResource>
</gmd:CI_Contact>
</gmd:contactInfo>
<gmd:role>
<gmd:CI_RoleCode
```





```
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_RoleCode"
codeListValue="pointOfContact">pointOfContact</gmd:CI_RoleCode>
</gmd:role>
</gmd:CI_ResponsibleParty>
</gmd:contact>
<gmd:dateStamp>
<gco:Date>2010-04-09</gco:Date>
</gmd:dateStamp>
<gmd:spatialRepresentationInfo>
<gmd:MD_VectorSpatialRepresentation>
<gmd:topologyLevel>
<gmd:MD_TopologyLevelCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmx/Codelists.xml#MD_TopologyLevelCode"
codeListValue="geometryOnly">geometryOnly</gmd:MD_TopologyLevelCode>
</gmd:topologyLevel>
<gmd:geometricObjects>
<gmd:MD_GeometricObjects>
<gmd:geometricObjectType>
<gmd:MD_GeometricObjectTypeCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmx/Codelists.x
ml#MD_GeometricObjectTypeCode" codeListValue="surface">surface</gmd:MD_GeometricObjectTypeCode>
</gmd:geometricObjectType>
</gmd:MD_GeometricObjects>
</gmd:geometricObjects>
</gmd:MD_VectorSpatialRepresentation>
</gmd:spatialRepresentationInfo>
<gmd:referenceSystemInfo>
<gmd:MD_ReferenceSystem>
<gmd:referenceSystemIdentifier>
<gmd:RS_Identifier>
<gmd:code>
<gco:CharacterString>WGS 1984</gco:CharacterString> </gmd:code>
<gmd:codeSpace>
<gco:CharacterString>EPSG</gco:CharacterString>
</gmd:codeSpace>
</gmd:RS_Identifier>
</gmd:referenceSystemIdentifier>
</gmd:MD_ReferenceSystem>
</gmd:referenceSystemInfo>
<gmd:identificationInfo>
<gmd:MD_DataIdentification>
<gmd:citation>
<gmd:CI_Citation>
<gmd:title>
<gco:CharacterString>Sample HISEA Dataset metadata</gco:CharacterString>
</gmd:title>
<gmd:date>
<gmd:CI_Date>
<gmd:date>
<gco:Date>2018-01-01</gco:Date>
</gmd:date>
</gmd:dateType>
<gmd:CI_DateTypeCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_DateTypeCode"
codeListValue="publication">publication</gmd:CI_DateTypeCode>
</gmd:dateType>
```





```
</gmd:CI_Date>
</gmd:date>
<gmd:identifier>
<gmd:RS_Identifier>
<gmd:code>
  <gco:CharacterString>HISEA_Xpto</gco:CharacterString> </gmd:code>
<gmd:codeSpace>
  <gco:CharacterString>MU</gco:CharacterString>
  </gmd:codeSpace>
  </gmd:RS_Identifier>
  </gmd:identifier>
  </gmd:CI_Citation>
  </gmd:citation>
<gmd:abstract>
  <gco:CharacterString>This dataset represents aggregated HISEA data sources.</gco:CharacterString>
  </gmd:abstract>
<gmd:pointOfContact>
<gmd:CI_ResponsibleParty>
<gmd:organisationName>
  <gco:CharacterString>HISEA Project</gco:CharacterString>
  </gmd:organisationName>
<gmd:contactInfo>
<gmd:CI_Contact>
<gmd:address>
<gmd:CI_Address>
<gmd:electronicMailAddress>
  <gco:CharacterString>HISEA@HISEA.eu</gco:CharacterString> </gmd:electronicMailAddress>
  </gmd:CI_Address>
  </gmd:address>
  </gmd:CI_Contact>
  </gmd:contactInfo>
<gmd:role>
<gmd:CI_RoleCode
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
  Schemas/resources/Codelist/ML_gmxCodlists.xml#CI_RoleCode" codeListValue="processor">processor</gmd:CI_RoleCode>
  </gmd:role>
  </gmd:CI_ResponsibleParty>
  </gmd:pointOfContact>
<gmd:descriptiveKeywords>
<gmd:MD_Keywords>
<gmd:keyword>
  <gco:CharacterString>HISEA</gco:CharacterString> </gmd:keyword>
<gmd:thesaurusName>
<gmd:CI_Citation>
<gmd:title>
  <gco:CharacterString>GEMET - INSPIRE themes, version 1.0</gco:CharacterString> </gmd:title>
<gmd:date>
<gmd:CI_Date>
<gmd:date>
  <gco:Date>2018-01-01</gco:Date>
  </gmd:date>
<gmd:dateType>
<gmd:CI_DateTypeCode
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
  Schemas/resources/Codelist/ML_gmxCodlists.xml#CI_DateTypeCode"
  codeListValue="publication">publication</gmd:CI_DateTypeCode>
  </gmd:dateType>
  </gmd:CI_Date>
  </gmd:date>
```







```
</gmd:CI_Citation>
</gmd:thesaurusName>
</gmd:MD_Keywords>
</gmd:descriptiveKeywords>
<gmd:resourceConstraints>
<gmd:MD_Constraints>
<gmd:useLimitation>
  <gco:CharacterString>no conditions apply</gco:CharacterString> </gmd:useLimitation>
  </gmd:MD_Constraints>
  </gmd:resourceConstraints>
<gmd:resourceConstraints>
<gmd:MD_LegalConstraints>
<gmd:accessConstraints>
<gmd:MD_RestrictionCode
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
  Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_RestrictionCode"
  codeListValue="otherRestrictions">otherRestrictions</gmd:MD_RestrictionCode>
  </gmd:accessConstraints>
<gmd:otherConstraints>

  <gco:CharacterString>intellectual property rights</gco:CharacterString> </gmd:otherConstraints>
  </gmd:MD_LegalConstraints>
  </gmd:resourceConstraints>
<gmd:spatialResolution>
<gmd:MD_Resolution>
<gmd:equivalentScale>
<gmd:MD_RepresentativeFraction>
<gmd:denominator>
  <gco:Integer>50000</gco:Integer>
  </gmd:denominator>
  </gmd:MD_RepresentativeFraction>
  </gmd:equivalentScale>
  </gmd:MD_Resolution>
  </gmd:spatialResolution>
<gmd:spatialResolution>
<gmd:MD_Resolution>
  <gmd:distance gco:nilReason="missing" />
  </gmd:MD_Resolution>
  </gmd:spatialResolution>
<gmd:language>
<gmd:LanguageCode
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
  Schemas/resources/Codelist/ML_gmxCodelists.xml#LanguageCode" codeListValue="pt">pt</gmd:LanguageCode>
  </gmd:language>
<gmd:characterSet>
<gmd:MD_CharacterSetCode
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmx
  Codelists.xml#MD_ClassificationCode" codeListValue="utf8">utf8</gmd:MD_CharacterSetCode>
  </gmd:characterSet>
<gmd:topicCategory>
  <gmd:MD_TopicCategoryCode>geoscientificInformation</gmd:MD_TopicCategoryCode> </gmd:topicCategory>
<gmd:extent>
<gmd:EX_Extent>
<gmd:geographicElement>
<gmd:EX_GeographicBoundingBox>
<gmd:westBoundLongitude>
  <gco:Decimal>12.09</gco:Decimal>
  </gmd:westBoundLongitude>
<gmd:eastBoundLongitude>
```





```
<gco:Decimal>18.85</gco:Decimal>
  </gmd:eastBoundLongitude>
<gmd:southBoundLatitude>
<gco:Decimal>48.58</gco:Decimal>
  </gmd:southBoundLatitude>
<gmd:northBoundLatitude>
  <gco:Decimal>51.05</gco:Decimal>
  </gmd:northBoundLatitude>
  </gmd:EX_GeographicBoundingBox>
</gmd:geographicElement>
<gmd:temporalElement>
<gmd:EX_TemporalExtent>
<gmd:extent>
<gml:TimePeriod gml:id="ID6w561g20-4vwp-8l1i-6d5l-hhu5sv1h2fw1" xsi:type="gml:TimePeriodType">
  <gml:beginPosition />
  <gml:endPosition />
  </gml:TimePeriod>
  </gmd:extent>
</gmd:EX_TemporalExtent>
</gmd:temporalElement>
</gmd:EX_Extent>
</gmd:extent>
</gmd:MD_DataIdentification>
</gmd:identificationInfo>
<gmd:distributionInfo>
<gmd:MD_Distribution>
<gmd:distributionFormat>
<gmd:MD_Format>
<gmd:name>
  <gco:CharacterString>3.2.1</gco:CharacterString> </gmd:name>
<gmd:version>
  <gco:CharacterString>GML</gco:CharacterString>
  </gmd:version>
  </gmd:MD_Format>
  </gmd:distributionFormat>
<gmd:transferOptions>
<gmd:MD_DigitalTransferOptions>
<gmd:onLine>
<gmd:CI_OnlineResource>
<gmd:linkage>
  <gmd:URL>http://HISEAplatform.eu/</gmd:URL> </gmd:linkage>
  </gmd:CI_OnlineResource>
</gmd:onLine>
</gmd:MD_DigitalTransferOptions>
</gmd:transferOptions>
</gmd:MD_Distribution>
</gmd:distributionInfo>
<gmd:dataQualityInfo>
<gmd:DQ_DataQuality>
<gmd:scope>
<gmd:DQ_Scope>
<gmd:level>
<gmd:MD_ScopeCode codeListValue="dataset"
  codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
  Schemas/resources/Codelist/ML_gmxCodetlists.xml#MD_ScopeCode">dataset</
  gmd:MD_ScopeCode>
  </gmd:level>
  </gmd:DQ_Scope>
</gmd:scope>
```





```
<gmd:report>
<gmd:DQ_DomainConsistency xsi:type="gmd:DQ_DomainConsistency_Type"> <gmd:result>
<gmd:DQ_ConformanceResult xsi:type="gmd:DQ_ConformanceResult_Type">
<gmd:specification>
<gmd:CI_Citation>
<gmd:title>
<gco:CharacterString>HISEA dataset</gco:CharacterString>
</gmd:title>
<gmd:date>
<gmd:CI_Date>
<gmd:date>
  <gco:Date>2018-01-01</gco:Date>
  </gmd:date>
  <gmd:dateType>
  <gmd:CI_DateTypeCode
codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_
Schemas/resources/Codelist/ML_gmxCodellists.xml#CI_DateTypeCode"
codeListValue="revision">revision</gmd:CI_DateTypeCode> </gmd:dateType>
  </gmd:CI_Date>
  </gmd:date>
  </gmd:CI_Citation>
  </gmd:specification>
<gmd:explanation>
  <gco:CharacterString>See the Commission Regulation No 1205/2008/EC.</gco:CharacterString>
  </gmd:explanation>
<gmd:pass>
  <gco:Boolean>true</gco:Boolean>
  </gmd:pass>
  </gmd:DQ_ConformanceResult>
  </gmd:result>
  </gmd:DQ_DomainConsistency>
  </gmd:report>
<gmd:report>
<gmd:DQ_CompletenessOmission>
<gmd:result>
<gmd:DQ_QuantitativeResult>
<gmd:valueType>
  <gco:RecordType>percentage</gco:RecordType>
</gmd:valueType>
<gmd:valueUnit>
<gml:BaseUnit gml:id="perc">
  <gml:identifier codeSpace="" />
  <gml:unitsSystem />
  </gml:BaseUnit>
  </gmd:valueUnit>
<gmd:value>
  <gco:Record>98,4%</gco:Record>
  </gmd:value>
  </gmd:DQ_QuantitativeResult>
  </gmd:result>
  </gmd:DQ_CompletenessOmission>
  </gmd:report>
<gmd:report>
<gmd:DQ_AbsoluteExternalPositionalAccuracy>
<gmd:result>
<gmd:DQ_QuantitativeResult>
<gmd:valueType>
  <gco:RecordType>measure</gco:RecordType> </gmd:valueType>
<gmd:valueUnit>
```





```
<gml:BaseUnit gml:id="meters">
  <gml:identifier codeSpace="" />
  <gml:unitsSystem />
  </gml:BaseUnit>
</gmd:valueUnit>
<gmd:value>
  <gco:Record>1,5m</gco:Record>
  </gmd:value>
  </gmd:DQ_QuantitativeResult>
  </gmd:result>
  </gmd:DQ_AbsoluteExternalPositionalAccuracy>
  </gmd:report>
</gmd:report>
<gmd:DQ_ThematicClassificationCorrectness>
</gmd:result>
<gmd:DQ_QuantitativeResult>
<gmd:valueType>
  <gco:RecordType>percentage</gco:RecordType> </gmd:valueType>
</gmd:valueUnit>
<gml:BaseUnit gml:id="years">
  <gml:identifier codeSpace="" />
  <gml:unitsSystem />
  </gml:BaseUnit>
  </gmd:valueUnit>
<gmd:value>
  <gco:Record>2,4%</gco:Record>
  </gmd:value>
  </gmd:DQ_QuantitativeResult>
  </gmd:result>
  </gmd:DQ_ThematicClassificationCorrectness>
  </gmd:report>
</gmd:lineage>
<gmd:LI_Lineage>
<gmd:statement>
  <gco:CharacterString>This dataset has been created while using several underlying datasets..</gco:CharacterString>
  </gmd:statement>
</gmd:source>
<gmd:LI_Source>
<gmd:scaleDenominator>
<gmd:MD_RepresentativeFraction>
<gmd:denominator>
  <gco:Integer>50000</gco:Integer>
  </gmd:denominator>
  </gmd:MD_RepresentativeFraction>
  </gmd:scaleDenominator>
</gmd:sourceCitation>
<gmd:CI_Citation>
<gmd:title>
  <gco:CharacterString>HISEA Data</gco:CharacterString>
  </gmd:title>
</gmd:date>
<gmd:CI_Date>
<gmd:date>
  <gco:Date>2018-01-01</gco:Date>
  </gmd:date>
</gmd:dateType>
<gmd:CI_DateTypeCode
  codeList="http://www.isotc211.org/2005/resources/codeList.xml#CI_DateType
  Code" codeListValue="creation" />
```





```
</gmd:dateType>  
</gmd:CI_Date>  
</gmd:date>  
</gmd:CI_Citation>  
</gmd:sourceCitation>  
</gmd:LI_Source>  
</gmd:source>  
</gmd:LI_Lineage>  
</gmd:lineage>  
</gmd:DQ_DataQuality>  
</gmd:dataQualityInfo>  
</gmd:MD_Metadata>
```





## Example of CMEMS data file

CMEMS uses netCDF format for the data files that it is handling. An example of the CMEMS netCDF files header is presented below:

```
NetCDF-3 Classic global-analysis-forecast-phy-001-024_1506677952654.nc {
dimensions:
    time = 2 ;
    depth = 1 ;
    latitude = 361 ;
    longitude = 61 ;
variables:
    // Preference 'PRESERVE_FVD': false,
    // dimensions consistent with ncBrowse, not with native MATLAB netcdf package.
    single time(time), shape = [2]
        time:long_name = "Time (hours since 1950-01-01)" ;
        time:standard_name = "time" ;
        time:calendar = "gregorian" ;
        time:units = "hours since 1950-01-01 00:00:00" ;
        time:axis = "T" ;
        time:_CoordinateAxisType = "Time" ;
    int16 thetao(time,depth,latitude,longitude), shape = [2 1 361 61]
        thetao:_CoordinateAxes = "time depth latitude longitude " ;
        thetao:long_name = "Temperature" ;
        thetao:standard_name = "sea_water_potential_temperature" ;
        thetao:units = "degrees_C" ;
        thetao:unit_long = "Degrees Celsius" ;
        thetao:_FillValue = -32767 ;
        thetao:add_offset = 21 ;
        thetao:scale_factor = 0.000732444 ;
        thetao:cell_methods = "area: mean" ;
    single longitude(longitude), shape = [61]
        longitude:step = 0.083328 f;
        longitude:units = "degrees_east" ;
        longitude:unit_long = "Degrees East" ;
        longitude:long_name = "Longitude" ;
        longitude:standard_name = "longitude" ;
        longitude:axis = "X" ;
        longitude:_CoordinateAxisType = "Lon" ;
    single latitude(latitude), shape = [361]
        latitude:step = 0.083336 f;
```





```
latitude:units = "degrees_north" ;
latitude:unit_long = "Degrees North" ;
latitude:long_name = "Latitude" ;
latitude:standard_name = "latitude" ;
latitude:axis = "Y" ;
latitude:_CoordinateAxisType = "Lat" ;
single depth(depth), shape = [1]
depth:units = "m" ;
depth:positive = "down" ;
depth:unit_long = "Meters" ;
depth:long_name = "Depth" ;
depth:standard_name = "depth" ;
depth:axis = "Z" ;
depth:_CoordinateAxisType = "Height" ;
depth:_CoordinateZisPositive = "down" ;

//global attributes:
:title = "daily mean fields from Global Ocean Physics Analysis and Forecast updated Daily" ;
:institution = "MERCATOR OCEAN" ;
:references = "http://www.mercator-ocean.fr" ;
:source = "MERCATOR PSY4QV3R1" ;
Conventions = "CF-1.0" ;
history = "Data extracted from dataset http://opendap-glo.mercator-ocean.fr:8080/thredds/dodsC/global-analysis-forecast-phy-001-024" ;
:time_min = 594036 ;
:time_max = 594060 ;
:julian_day_unit = "hours since 1950-01-01 00:00:00" ;
:z_min = 0.494025 ;
:z_max = 0.494025 ;
:latitude_min = 20 ;
:latitude_max = 50 ;
:longitude_min = 0 ;
:longitude_max = 5 ;
```





## ANNEX III: Data Privacy Procedures

### Data protection principles

1. The data privacy aims to ensure compliance with the Act. The Act sets out eight principles with which any party handling personal data must comply. All personal data:
  2. Must be processed fairly and lawfully, meaning that at least one of the following conditions must be met:
    - The data subject has given his or her consent to the processing;
    - The processing is necessary for the performance of a contract to which the data subject is a party, or for the taking of steps at the request of the data subject with a view to entering into a contract;
    - The processing is necessary for compliance with any legal obligation to which the data controller is subject, other than an obligation imposed by contract;
    - The processing is necessary in order to protect the vital interests of the data subject;
    - The processing is necessary for the administration of justice, for the exercise of any functions of either House of Parliament, for the exercise of any functions conferred on any person by or under any enactment, for the exercise of any functions of the Crown, a Minister of the Crown or a government department, or for the exercise of any other functions of a public nature exercised in the public interest by any person;
    - The processing is necessary for the purposes of legitimate interests pursued by the data controller or by the third party or parties to whom the data is disclosed, except where the processing is unwarranted in any particular case by reason of prejudice to the rights and freedoms or legitimate interests of the data subject.
  3. Where the personal data is sensitive personal data (defined below in Part 4 of this Policy), at least one of the following conditions must be met:
    - The data subject has given his or her explicit consent to the processing of the personal data;
    - The processing is necessary for the purposes of exercising or performing any right or obligation which is conferred or imposed by law on the data controller in connection with employment;
    - The processing is necessary in order to protect the vital interests of the data subject or another person in a case where consent cannot be given by or on behalf of the data subject, or the data controller cannot reasonably be expected to obtain the consent of the data subject, or in order to protect the vital interests of another person, in a case where consent by or on behalf of the data subject has been unreasonably withheld;
    - The processing is carried out in the course of the legitimate activities of anybody or association which is not established or conducted for profit, and exists for political, philosophical, religious or trade-union purposes, is carried out with appropriate safeguards for the rights and freedoms of data subjects, relates only to individuals who either are members of the body or association or have regular contact with it in connection







- with its purposes, and does not involve disclosure of the personal data to a third party without the consent of the data subject;
- The information contained in the personal data has been made public as a result of steps deliberately taken by the data subject;
  - The processing is necessary for the purpose of, or in connection with, any legal proceedings (including prospective legal proceedings), the processing is necessary for the purpose of obtaining legal advice, or is otherwise necessary for the purposes of establishing, exercising or defending legal rights;
  - The processing is necessary for the administration of justice, for the exercise of any functions conferred on any person by or under an enactment, or for the exercise of any functions of the Crown, a minister of the Crown or a government department;
  - The processing is either the disclosure of sensitive personal data by a person as a member of an anti-fraud organisation or otherwise in accordance with any arrangements made by such an organisation, or any other processing by that person or another person of sensitive personal data so disclosed, and is necessary for the purposes of preventing fraud or a particular kind of fraud;
  - The processing is necessary for medical purposes and is undertaken by a health professional, or a person who in the circumstances owes a duty of confidentiality which is equivalent to that which would arise if that person were a health professional;
  - The processing is of sensitive personal data consisting of information as to racial or ethnic origin, the processing is necessary for the purpose of identifying or keeping under review the existence or absence of equality of opportunity or treatment between persons of different racial or ethnic origins, with a view to enabling such equality to be promoted or maintained, and is carried out with appropriate safeguards for the rights and freedoms of data subjects.]
4. Must be obtained only for specified and lawful purposes and shall not be processed in any manner which is incompatible with those purposes;
  5. Must be adequate, relevant and not excessive with respect to the purposes for which it is processed;
  6. Must be accurate and, where appropriate, kept up to date;
  7. Must be kept for no longer than the necessary time in light of the purpose(s) for which it is processed;
  8. Must be processed in accordance with the rights of data subjects under the Act (for which, see Part 3 of this Policy);
  9. Must be protected against unauthorised or unlawful processing, accidental loss, destruction or damage through appropriate technical and organisational measures; and
  10. Must not be transferred to a country or territory outside of the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.





## Rights of Data Subjects

Under the Act, data subjects have the following rights:

- The right to access a copy of their personal data held by HiSea by means of a Subject Access Request (for which, see Part 8 of this Policy);
- The right to object to any processing of his or her personal data that is likely to cause (or that is causing) damage or distress;
- The right to prevent processing for direct marketing purposes;
- The right to object to decisions being taken by automated means (where such decisions will have a significant effect on the data subject) and to be informed when any such decision is taken (in which case the data subject has the right to require the data controller (by written notice) to reconsider the decision);
- The right to have inaccurate personal data rectified, blocked, erased or destroyed in certain circumstances;
- The right to claim compensation for damage caused by the Company's breach of the Act.

## Personal Data

Personal data are defined by the Act as data which relates to a living individual who can be identified from that data or from that data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

The Act also defines "sensitive personal data" as personal data relating to the racial or ethnic origin of the data subject; their political opinions; their religious (or similar) beliefs; trade union membership; their physical or mental health condition; their sexual life; the commission or alleged commission by them of any offence; or any proceedings for any offence committed or alleged to have been committed by them, the disposal of such proceedings or the sentence of any court in such proceedings.

HiSea only holds personal data that are directly relevant to its dealings with a given data subject. That data will be collected, held, and processed in accordance with the data protection principles and with this Policy. The following data may be collected, held and processed by HiSea:

- Name, Address, Phone Numbers and Email Addresses;
- Area of work, main requirements of the users activity that might be fulfilled by the offered services;
- Any other information provided to us by users.

## Processing Personal Data





Any and all personal data collected by HiSea are collected in order to ensure that HiSea can provide the best possible service to its customers.

Certain data collected by HiSea, such as IP addresses, certain information gathered by cookies, pseudonyms and other non-identifying information will nonetheless be collected, held and processed to the same standards as personal data. In particular, HiSea shall ensure that:

- All personal data are collected and processed fairly and lawfully;
- Data subjects are always made fully aware of the reasons for the collection of personal data and are given details of the purpose(s) for which the data will be used;
- Personal data are only collected to the extent that is necessary to fulfil the purpose(s) for which it is required;
- All personal data are accurate at the time of collection and kept accurate and up to date while they are being held and/or processed;
- No personal data are held for any longer than necessary in light of the purpose(s) for which they are required;
- A suitable online privacy policy is implemented, maintained and followed;
- Whenever cookies or similar technologies are used online, they shall be used strictly in accordance with the requirements of the Privacy and Electronic Communications Regulations, providing full details of cookie use and guidance on privacy;
- Individuals are provided with a simple, accessible method of amending any data submitted by them online;
- All personal data are held in a safe and secure manner taking all appropriate technical and organisational measures to protect the data.

## Data Protection Procedures

HiSea shall ensure that all of its employees comply with the following when working with personal data:

- All emails containing personal data must be sent securely;
- If personal data are being viewed on a computer screen and the computer in question is to be left unattended for any period of time, the user must lock the computer and screen before leaving it;
- No personal data should be transferred to any device personally belonging to an employee and personal data may only be transferred to devices belonging to agents, contractors, or other parties working on behalf of HiSea where the party in question has agreed to comply fully with the letter and spirit of this Policy and of the Act (which may include demonstrating to the Company that all suitable technical and organisational measures have been taken);
- All personal data stored electronically should be backed up regularly with backups stored [onsite] AND/OR [offsite].





- All electronic copies of personal data should be stored securely using passwords and data encryption;
- All passwords used to protect personal data should be changed regularly and should not use words or phrases that can be easily guessed or otherwise compromised. All passwords must contain a combination of uppercase and lowercase letters, numbers, and symbols [ All software used by the Company is designed to require such passwords];
- Under no circumstances should any passwords be written down or shared between any employees irrespective of seniority or department. If a password is forgotten, it must be reset using the applicable method. IT staff do not have access to passwords;
- All personal data shall be regularly reviewed for accuracy and completeness. Where HiSea has regular contact with data subjects, any personal data held about those data subjects should be confirmed at least annually. If any personal data are found to be out of date or otherwise inaccurate, they should be updated and/or corrected immediately where possible. If any personal data are no longer required, they should be securely deleted and disposed of;
- When personal data held by HiSea are used for marketing purposes, it shall be the responsibility of the Managing Director to ensure that no data subjects have added their details to any marketing preference databases including, but not limited to, the Telephone Preference Service, the Mail Preference Service, the Email Preference Service, and the Fax Preference Service. Such details should be checked at least annually.

