



UNITED
NATIONS

EP

UNEP/MED WG.467/Inf.10



UNITED NATIONS
ENVIRONMENT PROGRAMME
MEDITERRANEAN ACTION PLAN

8 August 2019
Original: English

7th Meeting of the Ecosystem Approach Coordination Group

Athens, Greece, 9 September 2019

Agenda Item 6: IMAP Pilot Info System and Related Quality Assurance Issues; Data Standards and Data Dictionaries; MAP Data Management Policy

IMAP Pilot Information System: User Needs Analysis

For environmental and economic reasons, this document is printed in a limited number. Delegates are kindly requested to bring their copies to meetings and not to request additional copies.

UNEP/MAP
Athens, 2019

Table of contents

I. Introduction	1
1. Premise	1
2. Background	1
3. Relevant MAP Supporting Documents	2
II. Requirements of the Pilot Information System (IMAP Pilot Info System)	3
1. User needs analysis - Non-functional Requirements.....	4
2. User needs analysis - Functional Requirements.....	5
III. Components of the Pilot Information System (IMAP Pilot Info System)	6
1. General characteristics	6
2. Software Components	8
2.1. Application software components - database	8
2.2. Application software components - elaboration	9
2.2.1. Central Module.....	10
2.2.2. Data Acquisition Module	11
2.2.3. Reporting Module	13
2.2.4. Data Representation Module	13
2.2.5. User Management Module	14
2.2.6. Login Module.....	15
3. Applicative Software Components - publication	15
3.1 Technology Infrastructure	15
3.2 Data processing and consultation environment - client side	15
3.3 Processing, consulting and publishing environment - server side	15
IV. Process of data processing	17
1. Information Flow.....	18
2. Data Processing Phases	18
3. User profiles	21

List of Abbreviations/Acronyms

COP	Conference of Parties
EcAp	Ecosystem Approach
ETL	Extract, Transform, Load
GES	Good Environmental Status
IMAP	Integrated Monitoring and Assessment Programme
LBS	Land-based sources
MPA	Marine Protected Areas
NAP	National Action Plans
NIS	Non-indigenous species
PIS	Pilot Info System
RBDMS	Relational database management system
UN Environment	United Nations Environment
UN Environment/ MAP	United Nations Environment– Mediterranean Action Plan
W3C	World Wide Web Consortium

I. Introduction

1. Premise

1. The establishment of an adequate information system to support IMAP implementation is a key priority for UN Environment/MAP and the EU funded EcAp-Med II project.
2. The focus of the work of INFO/RAC is to undertake a Mediterranean specific assessment (user need analysis) on how to further develop the UN Environment/MAP InfoMAP platform in line with other regional data-management platforms/reporting obligations, with inter-operability and capable data-management system.
3. INFO/RAC, in full compliance with the needs of MAP system, and for the implementation of Integrated Monitoring and Assessment Programme of the Mediterranean Sea and Coast and Related Assessment Criteria (IMAP) is going to provide a web software platform for data and information management of all datasets related to marine ecosystem monitoring and assessment under its mandate.
4. The platform should allow:
 - comply with the information requirements in the light of the obligations laid down in the Barcelona Convention;
 - evaluate, define and monitor decision-making strategies through achieving the goal of adequacy and usability of the data.
5. Within the framework of this project, the creation of the IMAP Pilot Info System, including the integrated communication infrastructure, is of particular importance. The system will collect and integrate data from different sources and data providers and provide information to different target user groups.
6. The following document provides information to characterize both the current operational situation for the proper framing of the functional aspects required for the IMAP Pilot Info System as a data and information management tool and to characterize the technological and functional aspects considered critical and essential for the executive project in order to be properly implemented.

2. Background

7. The 19th Meeting of Contracting Parties (COP 19), held in February 2016, adopted the Integrated Monitoring and Assessment Programme (IMAP) of the Mediterranean Sea and Coast and Related Assessment Criteria (Decision IG. 22/7), with a list of regionally agreed good environmental status descriptions, common indicators and targets, with principles and clear timeline for its implementation.
8. IMAP, through Decision IG.22/7 lays down the principles for an integrated monitoring, which will, for the first time, monitor biodiversity and non-indigenous species, pollution and marine litter, coast and hydrography in an integrated manner. As such, IMAP aims to facilitate the implementation of article 12 of the Barcelona Convention and several other monitoring related provisions under different Protocols with the main objective to assess GES. Its backbone are the 11 Ecological Objectives and their 27 common indicators as presented in Decision IG. 22/7.

3. Relevant MAP Supporting Documents

9. A number of assessment products have been prepared by the Secretariat since the adoption of the Ecosystem Approach Roadmap, with the contribution of Contracting Parties, all MAP Components, relevant partners and regional stakeholders, with the view to provide a clearer image of the state of environment in the Mediterranean, and define the main drivers, pressures and impacts as well as their anthropogenic sources.

10. More in particular:

a. The **Initial Integrated Assessment of the Mediterranean Sea and Coastal Areas** was prepared in 2011, based on available knowledge and information. This initial assessment was undertaken at sub-regional and regional levels. A summary for decision makers, providing the main findings and priorities was submitted and endorsed by the COP17 with its Decision IG. 20/4.

b. In addition, the **State of the Mediterranean Marine and Coastal Environment Report** (SoER MED 2012) was prepared and published in 2012, which provides valuable information on the main human induced pressures, the state and impacts on the Mediterranean Ecosystems, while its last part presents the regional regulatory framework and identifies gaps and key steps for the implementation of the Ecosystem Approach.

c. The third assessment report since the adoption of the Ecosystem Approach Roadmap is the **Quality Status Report 2017**. Based on the COP 19 Decision on the Integrated Monitoring and Assessment Programme of the Mediterranean Sea and Coast and Related Assessment Criteria (IMAP, Decision IG. 22/7) and the MAP Programme of Work 2016-2017 (Decision IG.22/20), UN Environment/MAP Secretariat with the input of all components and the experts of the various Correspondence Groups on Monitoring (CORMONs) have prepared the draft Quality Status Report (QSR 2017).

d. The QSR 2017 is fully based on the **IMAP Common Indicators**, using data provided by the Contracting Parties and MAP RACs or collected through research projects and made available by other sources of information.

f. In line with the above, **Indicator Guidance Fact Sheets** have been developed for each Common Indicator to ensure coherent monitoring, with specific targets defined and agreed in order to deliver the achievement of Good Environmental Status (GES) and as such, provide concrete guidance and references to Contracting Parties to support implementation of their revised national monitoring programmes towards the overall goal of implementing the Ecosystem Approach (EcAp) in the Mediterranean Sea and achieving GES.

g. The structure of a **Common Indicator Factsheet** can be summarized looking at the different organization levels of the developed factsheet templates. A common set of relevant policy and science based information is required on each (ie. Indicator Title, Rational, Policy Context and Targets, Indicator analysis methods and Methodology for monitoring (temporal and spatial scope), Contacts and Document Registration). In each, detailed definitions, methodologies, references, gaps, uncertainties, data analysis approaches, basis for aggregation (if applies) and outputs complete the guidance factsheets.

II Requirements of the IMAP Pilot Info System

11. According to information model paradigm represented by Data-Information-Knowledge-Wisdom (DIKW) pyramid, the first step to implement an information work flow is to identify the appropriate layer in the pyramid. For IMAP, the Data layer is the most fundamental and is represented by monitoring data sets to be used to populate and calculate indicators for ecosystem assessment required by the EcAp programme.

12. Collection of monitoring data can be properly set if Information Standards are defined in this process. Such activity is not to be confused with the process of the implementation of indicators which is still on-going, implies the definition of common specific and quantitative indexes and metrics to be used on collected monitoring data and actually belongs to the Information layer of the DIKW pyramid.

13. IMAP Info System as a node to be developed for the collection and sharing of monitoring data and the implementation of IMAP indicators, has started with a proposal of Information Standards.

14. Marine monitoring programs have a consolidated history that goes back to the 70s-80s and have been further developed in the framework of European Union Directives on coastal (WFD – 2000/60/CE) and marine waters (MSFD -2008/56/EC). In the last 10-15 years a certain number of Mediterranean countries have already implemented monitoring programs and collected data according to shared information standards which may contribute to many IMAP indicators.

15. IMAP indicators also covers issues that have not been formally standardized and some of these are rather new and challenging from the monitoring protocols point of view (for ex. marine litter). Furthermore, Countries usually implement monitoring programmes through regional environmental agencies, research institutes, national/local observatories that jointly participate to monitoring/research projects at the Mediterranean level where Information Standards for data collection have been proposed for some specific issues.

16. Taking into account such context, the first proposal of Information Standards for the collection of monitoring data is based on the **Indicator Guidance Fact Sheets** where the method including standards etc are defined.

17. An information work flow cannot be properly defined without the identification of users or groups of users profiles which, for the IMAP Pilot Info System, can be naturally categorized in the first step as the following:

- UN Environment/MAP;
- INFO/RAC;
- MAP Components;
- Contracting Parties;
- MAP Focal Points (MAP FPs);
- International Institutes;
- National Institutes

Once the fundamental elements of the information work flow have been generally identified in terms of monitoring data for IMAP indicators and users, user needs analysis can be carried on to define functional and non-functional requirements.

1. User needs analysis - Non-functional requirements

Support

18 The IMAP Pilot Info System will have to provide support to the users, so it must be able to share information and summary data that will also be useful to identify critical issues and possible rationalizations both in the internal context and in relationships with external entities such as data providers.

Resources

19 The resources needed to manage the system must be properly planned and allocated, while respecting service quality levels. Systems and IT environments will therefore have to be highly efficient and reliable. In fact, a fundamental and generally overlooked role is also played by IT management staff which are actually responsible for the maintenance of the system.

Scalability

20 User needs change and scale up very easily in a development context like IMAP implementation process. The platform should provide for the possibility of integrating application, adding modules and technological infrastructure even at later times, managing increased needs with minimum effort and maximum performance.

Open technology

21 For investment protection, the platform must use open, flexible and sound technology based on *standard and de facto* widespread standards (such as ISO, OGC, OPEN WEB etc.) nowadays available. This means having broad possibilities of integration to the outside IT environment, as long as the design of the internal modules and their relationships follows a unified vision.

Architecture

22 Software development team is a major part of IT staff and the possibility to work on different functions without prejudice plays a key role in architecture choices. In compliance with regulatory, operational, resource and budget constraints, the choice of a "**modular**" architecture will allow a more flexible planning of the software development process.

Organizational requirements

23 The IMAP Pilot Info System must play a central role in supporting the whole process. It will therefore have to avoid rigidity of the processes, and in fact guarantee high adaptability, easy activation of new procedures or reconfiguration of existing ones.

Technological requirements

24 The IMAP Pilot Info System meets the following technological requirements:

Modular programming

25 A modular information system has been developed in terms of modules and components around different modules that encompass a cohesively whole solution or system. Each module will

contain everything necessary to execute only one aspect of the desired functionality. A module interface expresses the elements that are provided and required by the module. The elements defined in the interface are detectable by other module.

26 This allows the system to expand through the composition of a personalized solution obtained by combining the only modules of interest.

User interface

27 Must be conceived to ensure usability through product and process. The most common criteria of software ergonomics, such as: consistency, efficiency, design, memory and context help, will be used.

Application logs

28 The application log mechanism must be fully configurable both in terms of size and tracking level.

Database

29 Must be of a relational type that can ensure transactionality, multi-user, concurrency and, where required by the data schema, compliance with the rules of referential integrity; it must be periodically updated, scalable, assuring portability between different platforms and equipped with a sound back-up system to manage data recovery in case of failures of the system.

Internet Technologies

30 Open web technologies must be used as described in the [Open Web Platform](#) by the World Wide Web Consortium (W3C).

Application Integration

31 The solution must have the ability to share information and services with external systems, both as a data exchange opportunity and as a means of providing external functionality (RACs, Contracting Parties, Research Organizations, Institutions, etc.).

2. User needs analysis - Functional requirements

32 Below there are the functional requirements that will characterize the information system based on user needs analysis not related to IT development and maintenance of the system:

ETL (Extract, Transform, Load)

33 The feeding and updating of the analytical database is in itself the most delicate and characteristic part of the whole discipline. The ETL process is developed through the implementation of automated procedures for extracting data from homogeneous or heterogeneous sources, performing intermediate processing for storing in the proper format or structure for the purposes of querying and analysis, and finally loading the flow into the target database.

34 The key feature required for the ETL module is that the user must be able to design and implement the necessary procedures in intuitive visual environments (which do not require the user to know any programming language).

Manual Data Entry

35 To assist user in manual insertion of information specific web form must be developed, in order to reduce the probability of errors, (if possible and if the number of alternatives is reasonable) the selection mode from list values rather than mere typing must be chosen. However, the formal correctness, consistency and plausibility of the entered data shall be checked. As a matter of fact, monitoring data are usually massively uploaded using excel, csv or db access files and web form are actually used only to modify or provide additional information on already stored data sets.

Reporting

36 Reporting refers to data products and documents that are to be produced from database queries in order to comply with both internal and external information requirements.

37 All information presented to the user must be printable or exported in the most common formats to ensure integration and data exchange with other application environments. Therefore, the main feature of the reporting tool is the availability of a visual editor that allows you to create reports easily and intuitively and the ability to export documents in various formats (.pdf, .html, .doc, .rtf, .txt, .xls, .csv).

38 The reporting system allow Contracting Parties to produce their own reports for policy makers.

Dashboards

39 Creating dashboards can be implemented through interactive pages through which you can manage a series of operations such as:

- displaying charts or synthetic charts;
- management of the compilation phase of the reporting modules;
- interchange of data.

Data Mining

40 In addition to the classic charts, dashboards, and reports, the platform will need to use tools to perform advanced statistics. The availability of huge amounts of data allows the application of statistical techniques falling within the Data Mining.

Web portal

41 The platform is able to be run on a web portal that allows its data feeding, and the publication of the produced information outputs.

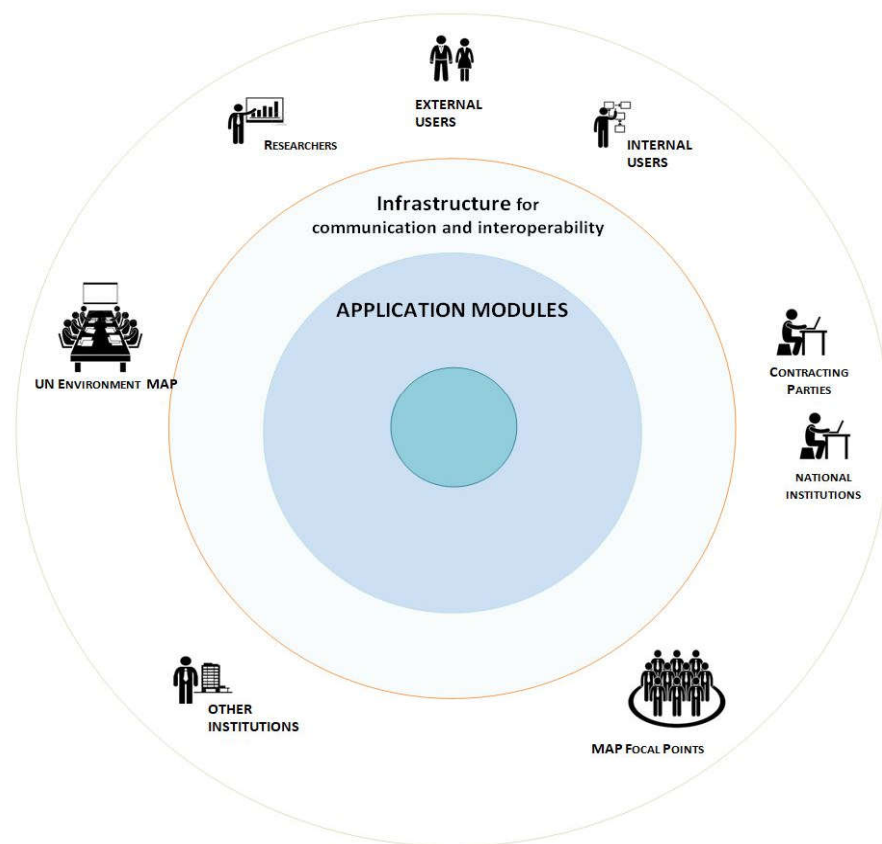
III Components of IMAP Pilot Info System

1.General characteristics

42 The IMAP Pilot Info System operates in a highly integrated and consistent way, guaranteeing coherence for data sharing with agreed policy.

43 Therefore, a central core is represented by a central RDBMS – Relational Data Base Management System, by a set of autonomous application modules (software components) and by an infrastructure for communication and interoperability with other systems.

44 The figure below is the conceptual model of the proposed IMAP Pilot Info System. In the external circle users are represented at the same level. Of course, the system allows the hierarchical distribution of users (external circle) according to Contracting Parties needs.



45 The modular approach allows to considerably reduce critical interferences between system components, as they can be developed independently and then plugged together.

46 The modular structure has to assure interoperability between functional application components for integration purposes, while remaining autonomous and independent both operational and functional.

47 The proposed solution envisages a IMAP Pilot Info System not only modular but also scalable, that is to say that it can be functionally upgraded with the addition or development of new application modules, and capable of managing the increments of traffic and users expected and desirable, simply by increasing the processing and / or transmission capacity.

2. Software Components

2.1. Application software components – database

48 The technology platform on which the application solutions will be hosted will have to provide as the main base component a central RDBMS – Relational Data Base Management System, which will have the following features:

- it must be based on relational systems that will allow all types of data and information to be managed;

- must use query, reporting and reporting languages, based on standard SQL;
- it must ensure transactionality, multi-user, concurrency and wherever required by the data schema, respect for referential integrity.

49 A portion of the database (catalog or dictionary) contains a centralized description of the data, which can be used by all modules.

50 In particular, the catalog should include all the metadata on:

- stations / areas;
- parameters;
- habitats;
- assessment areas;
- species list;
- unit of measure;
- laboratories;
- users;
- etc.

51 The logical structure of the database will be designed in the form of ER diagrams with standard SQL script creation scripts. This documentation will define the typology and characteristics of each field in the tables, the existing constraints on the fields and among the fields, and ultimately providing precise information about the data use mode. The documentation must also be consistently aligned with the actual database structure.

2.2.Application software components - elaboration

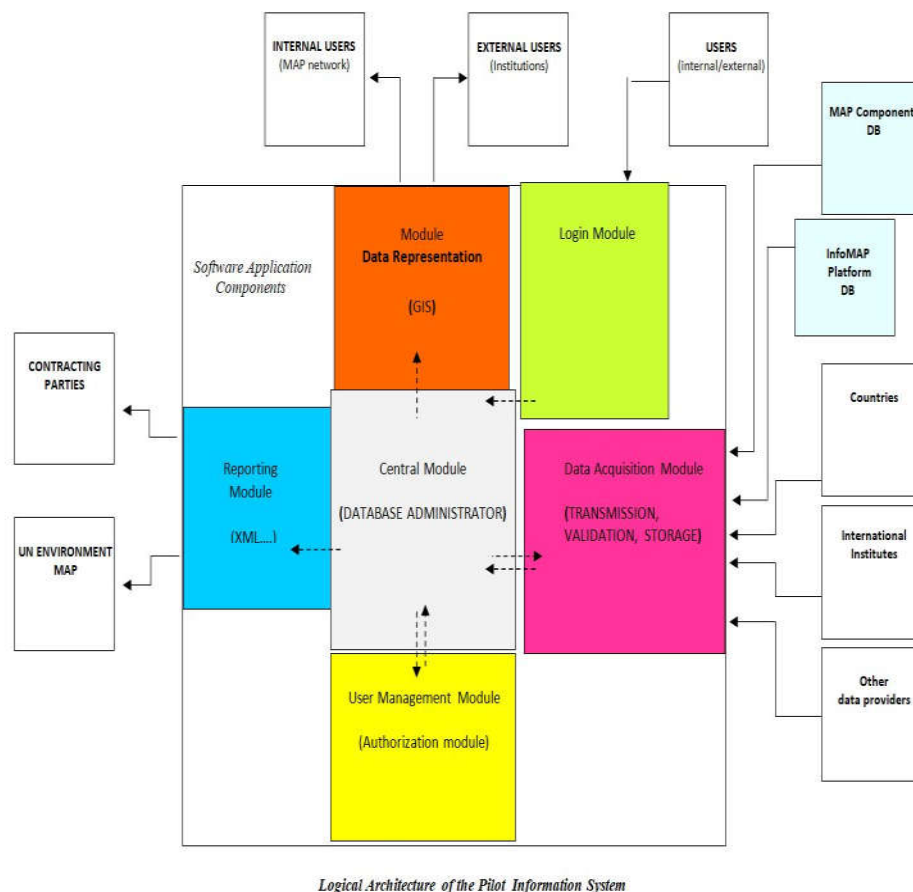
52 The logical architecture of the system, as mentioned above, is flexible and modular and lends itself to being combined in a broad spectrum of architectural solutions: from monolithic, hosted on a single server, to more complex situations at logical and physical level.

53 Six main modules are defined:

- 1. CENTRAL MODULE**
- 2. DATA ACQUISITION MODULE**
- 3. REPORTING MODULE**
- 4. DATA REPRESENTATION MODULE**
- 5. USER MANAGEMENT MODULE**
- 6. LOGIN MODULE**

54 The **CENTRAL MODULE** is responsible for database management and is the module through which all other application components communicate.

55 The logical architecture is represented in the following diagram:



2.2.1. Central Module

56 The Central Module has the task of storing, organizing, centralizing and making available all the information.

57 It will therefore have to be able to handle:

- all types of data and information that are involved in the monitoring activity;
- the congruence and consistency of the information entered and / or modified.

58 However, the form must allow not only to manage the current data, but also to historicise the data and information useful to reporting, analysis, etc.

59 It must also provide backup procedures (eg incremental, integral, etc.) and restore procedures to minimize the risk of information loss.

60 All static data will be managed through a web based interface.

61 The module allows through the catalogue management to share information in an up-to-date format and stay aligned after each change.

2.2.2. Data Acquisition Module

62 The definition of data acquisition refers to the set of procedures by which the data sent from the periphery to the centre is subjected to a set of controls defined by appropriate rules before it can be inserted into the consolidated database.

63 The system of rules to be built will have to be based on the consistency between the boost towards the central government and the instances of autonomy of the suburbs, without forgetting that the central government will still remain the functions of addressing, coordinating and controlling.

64 It was intended to identify the procedures:

- data transmission;
- validation of data (formal check, consistency check ...);
- data storage.

Data Transmission

65 The monitoring data collection function must be based on the application of precise rules as described in the **Indicator Guidance Fact Sheets** for each indicator. A prerequisite will be the assurance that data production through monitoring activities is managed in a uniform and standardized way through a systematic approach.

66 In order to transmit the information from the periphery to the platform in a congruent and consistent manner, the following actions must be implemented:

- design of information flows:
 - analysis of the information that should be made available at the national level by peripheral systems (RACs, Contracting Parties, National institutes, etc.) and identification of information should be returned from central to peripheral level;
 - defining the channels through which this information should be transmitted (transmission mode).
- creating information and coding standards in line with community standards.

67 The proposal of data standards for data exchange, in line with IMAP Common Indicators and data user needs, is elaborated by INFO/RAC, according to the **Indicator Guidance Fact Sheets**. These information standards are made up of a set of spreadsheets containing detailed data tables and data dictionaries supported by compilation guidelines to ensure that the data can be easily verified, analyzed and clearly interpreted. The data standards will be uploaded on IMAP Pilot Info System.

Validation of data

68 The reliability of an entire set of data collected in the various monitoring surveys is a cornerstone of the decision-making process that accompanies the whole process.

69 The checks to be carried out are both formal and coherent:

Formal check

70 Validation procedures will be implemented to check the formats (text, date/time, numerical, boolean), minimum and maximum size, specific constraints on numerical intervals and expressions, and compliance where relevant with default enumeration lists, and so on.

Consistency check

71 The need to validate the data is obvious: the data collected are affected by multiple sources of error (sampling, analytic and "intrinsic") whose single evaluation is not absolutely determinable. The validation procedure has the task of applying the established criteria (eg statistical methods) to the data set that allow to calculate concentrations and / or measurements corresponding to a predetermined probability level. The Data Standard and Data Dictionary are provided as excel formats and tis format can also be used to upload data into the system. As quality control are implemented through development of XML schema, there is also the possibility to share data by using XML format. In this case, the procedure will apply to the file (XML) formal checks and data consistency and will operate in two phases:

- Phase 1: Provides formal checks on an XSD schema. Failure to comply with these checks causes the entire XML file to be lost;
- Phase 2: (only accessible if Phase 1 checks are exceeded) involves applying data consistency checks. Failure to comply with these controls implies the scrap of the single record and not the entire XML file.

72 Therefore, the system should include the implementation of application modules for the quality control of analytical and cartographic data, with particular reference to: control of the data format (text, numeric, date, logical, etc.), compliance with lists of predefined values, consistency with acceptable range of values, geographic correctness of both punctual and polygonal information. Control procedures should be implemented in such a way as to allow the user to validate appropriate parameters without any modification of the software application module. The list of control procedures must be indexed and managed using a specific application form that contains each code, description, target, and output procedure. The control procedure target can be a table on the relational database or a .xls, .xlsx, .mdb, .csv file in a default schema.

73 The output of control procedures must be returned in .csv, .xls, and xlsx format and inserted into dedicated tables in the relational database. The "unit control procedure" is a single control to be performed on a specific analytic field of a table: the date format, the numeric format, and the text format, etc.

74 If there are multiple fields of the same format on the same table or on different tables, the number of unit controls is equal to the number of fields to be controlled: i.e. if there are two data format fields in a table and a data format field in another table, the number of unit controls is 3.

75 Unit controls must be implemented by software procedures executable on a given field of a particular table regardless of the number of records in the table. Topological controls on cartography are not included in the unit control procedures and will be performed using GIS desktop software tools.

Data storage

76 The procedure is to consolidate validated data through storage operations in order to keep the information in time and to find them easily. The amount of information to be collected is such that the

data for being properly managed must be stored in a precise way so that search and consultation can be managed as efficiently as possible.

2.2.3. Reporting Module

77 The module will handle the reporting process by providing reporting creation and management capabilities, as well as with predefined reports that can:

- respect the information needs of the Barcelona Convention protocols and requirements therein specified;
- meet administrative requirements;
- allow more complex and articulate analysis activities for decision making purposes.

78 The module will systematically and articulate the extraction and correlation of data subject to reporting.

79 The reporting process is described in the diagram below.



2.2.4. Data Representation Module

80 Organized, validated and archived data are usable and can be consulted according to the data policy that will be agreed with the Contracting Parties. This module manages the process of interpretation, analysis and representation of data in order to produce "knowledge".

81 Through the Web-Gis tool, the module will use tools to represent information designed to make the data not only visible but also comprehensible.

82 Within the scope of the IMAP Pilot Info System for the purpose of processing data into information, it is necessary to implement a software tool falling within the general typology of the Information Design Tool, which offers a platform capable of "defining, planning and giving shape to a content considering the context in which it is presented, with the aim of meeting the information request of the recipients "(International Institute for Information Design).

83 The following minimum functionalities required for this type of software tool (Information Design Tool type) based on the main features, data aggregation capabilities, and data analysis properties:

Main features

- intuitive and interactive visual exploration;
- geographic display of data distribution related to monitored phenomena;
- identification of areas of interest to conduct in-depth analysis;

- display of georeferenced elements (common, sites, stations, points of interest ...) with clustering and punctuation.

Data aggregation and metadata

- correlation between information from different sources to obtain aggregate information;
- "data set" information merge according to common parameters or based on the distance between the georeferenced elements (common, sites, stations, points of interest);
- correlations between information from different sources for aggregate information;
- export of aggregated data to standard interoperable formats (excel, csv);
- description of metadata with ISO profile in accordance with the guidelines by using a web editor for their compilation;
- MS service implemented according to ISO standards (NETWORK SERVICES);

Data analysis

- charts and dashboards;
- navigable and exportable reports in Excel, etc;
- data segmentation (eg smart data browser);
- Web Forms as a collection of user interface components allows a user to enter data that is sent to a server data entry);
- analyzes and views can easily be "cloned" in other areas of interest for immediate comparisons.
- search the metadata catalogue integrated into the GIS viewer so that it can be accessed on the same web page:
 - the display of the cartographic layers in a default folder list;
 - the search and selection of map layers on the metadata catalogue;
 - adding a mapping layer selected by the dot to the list of map layers in view.

2.2.5. User Management Module

84 Data stored in the IMAP Pilot Info System must be protected by readings and / or modifications by unauthorized users. This module has the task of ensuring adequate security criteria and meeting current security standards when accessing each information.

85 Users, user groups, authorization profiles, and access privileges in terms of applications, application and data functions have to be properly configured into the system so to assure compliance with agreed data policy. For each configured user, the operating role to which it can be associated must be defined.

86 The module must allow:

- password assignment as well as management in terms of complexity, length and validity over time;
- detection of access anomalies, such as failed attempts, simultaneous access by the same user, etc
- instant control of all connected users;
- the session closes automatically after the expiration of a confusing period of inactivity.

87 Finally, the module will have to provide features that can trace the activities of each user continuously. All traced operations can be viewed by users with special permissions (i.e super-user) in order to identify and analyze operational critical situations.

2.2.6. Login Module

88 IMAP Pilot Info System will provide both a public and a restricted access area. Access to restricted area will be granted with a web user interface, probably offered directly on the institutional site. Access should also implement a single-sign-on feature to ensure coherence and security of access.

89 The Login module will need to integrate applications from all other modules by providing the same presentation layout and dynamically integrating all controls regarding user privileges, features that can be activated for each application, and the data that each user can manipulate.

90 The number of applications available must also depend on the level of security for which the user is credited and in accordance with what is configured in the user management module.

3. Applicative Software Components - publication

3.1 Technology Infrastructure

91 The SW architecture will have to take into account the needs of collecting, processing and publishing the web of sea-to-sea monitoring data. As a result, two operating environments are identified:

- Data processing and consultation environment - client side
- Processing, consulting and publishing environment - server side

3.2 Data processing and consultation environment - client side

92 The web client for the data processing must operate within a web browser window but which functions are a set of separate application.

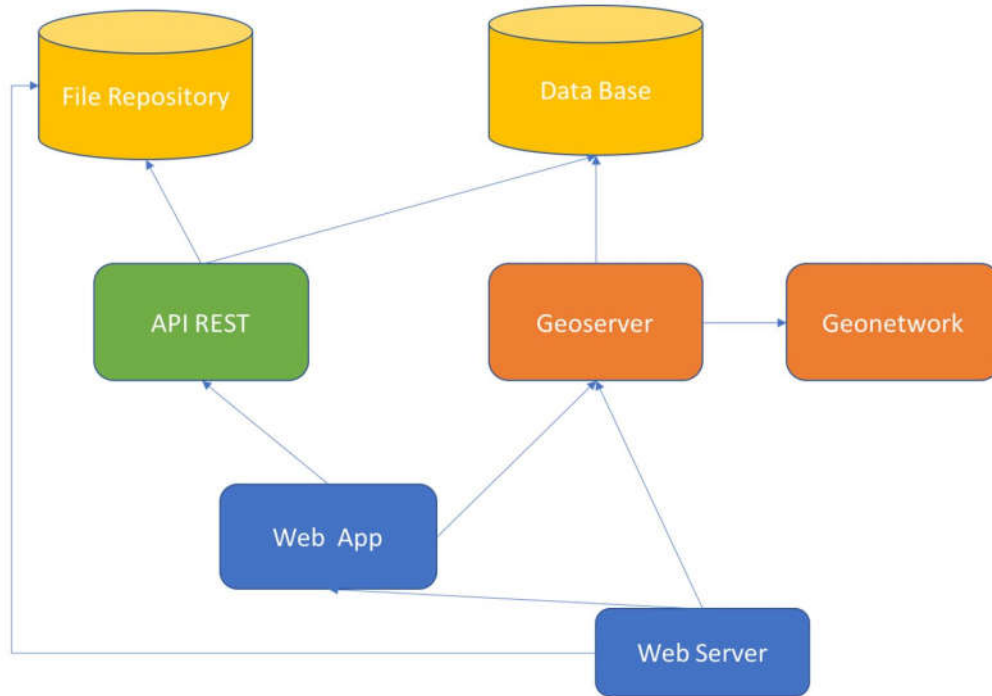
3.3 Processing, consulting and publishing environment - server side

93 This environment must be endowed with open-source software which can provide backup and restore functions in line with enterprise applications.

94 Application software must be such as to guarantee the functionality required in this with particular regard to the functions of:

- RDBMS (Relational database management system)
- Application server
- Web GIS server

95 The solution provided is composed of the following software modules linked according to the following architectural diagram:



96 In particular, every component is realized by the following technology solutions:

- **Web Server: Apache 2.4.7.** It is responsible for management of each HTTP/HTTPS requests from the web portal and connects to Web App, Geoserver or File Repository according to specific functionalities
- **Web App:** developed in nodeJS and AngularJS technologies with Model-View-Controller architecture. It is responsible for login, user management, information standard management, files uploads and connects to Geoserver for WEB GIS functionalities. It also connects to Data Base by API REST (Application programming interface) modules for the queries on user and monitoring data;
- **Geoserver: GeoServer 2.9.2.** It allows to publish and manage GIS layers, web GIS services, WMS and WFS services. It connects to Data Base for the publishing of monitoring data on geodatabase and to Geonetwork for GIS layers catalogue
- **Geonetwork: GeoNetwork 2.10.4.** It manages GIS layers and services catalogues also in compliance with ISO 19215
- **API REST:** nodeJS and Json. It maps File repository and Data Base information to HTTP requests and allows Web App communications. It is also responsible for the implementation of quality data control on monitoring files and compliance with information standards;
- **Data Base:** PostgreSQL 9.3 and PostGIS 2.1.2. It is a relational data base management systems with GIS extension for the collection of user informations, monitoring data also with geographical features, information standards and monitoring files. It is equipped with scheduled back-up procedures;

- **File repository:** json. It allows storage of monitoring files, information standards and any other files relevant for the systems, assigning a unique identifier for each files and mapping the physical file on the file system to the logical file requested by Web App and stored on Data Base.

97 The system is realized with open source software distributed with Open GML user license and it is scalable and portable on Linux and Windows operating systems. It offers web services and pages accessible by widespread browser like Chrome, Firefox, IE, etc...

IV. Process of data processing

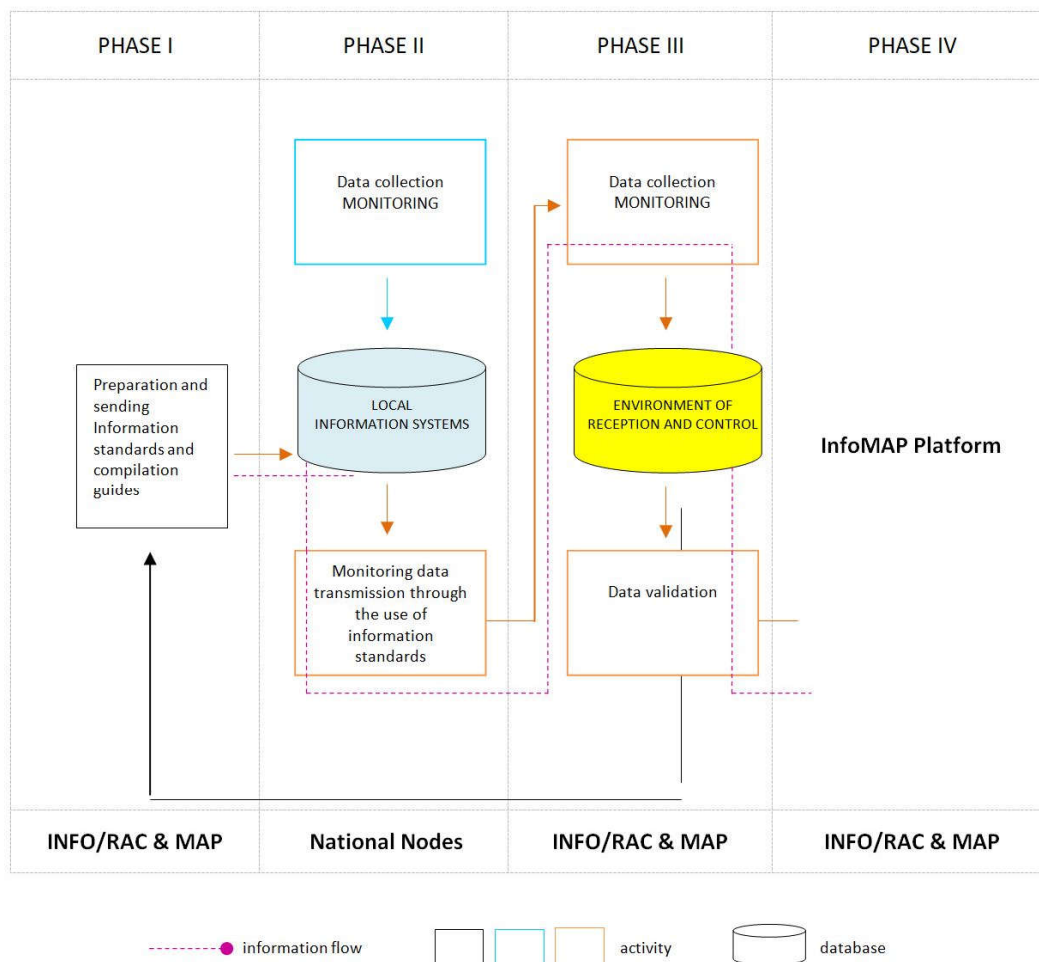
98 The **IMAP Pilot Info System** will need to be able to handle several heterogeneous data flows, such as:

- data deriving from monitoring programmes implemented pursuant to IMAP;
- relevant data used for IMAP assessment, including socio economic, definition of GES and determination of environmental targets;
- data belonging to other existing databases of MAP Components (included or not yet in the InfoMAP Platform).

1. Information flow

Scenario

114. Information flow is represented in the underlying scenario according to four distinct processing phases. The stages are in progressive sequence. For each phase, the tasks to be performed are distinguished.



2. Data Processing Phases

99 The following steps are described in detail in the above paragraph and related to the processing of data from the monitoring:

PHASE I. Preparation of the Information Standards and their guidelines

100 The information standards and the technical guidelines for their compilation shall be made available to the data providers. The Information Standard is made by a Data Standards (DSs) which contain the list of information to be provided (i.e. name of monitoring station and geographic coordinates, monitored parameter, monitoring methodology adopted) and a Data Dictionary (DD) which includes all the list of admissible values to be used in the DS. For each single information to be provided, the Information Standard should indicate:

- field name (i.e. PARAMETER);
- field format (one selected from: text, numeric, yes / no, date);
- any list of eligible values (i.e LIST OF CHEMICALS MONITORED or LIST OF SPECIES);
- description of the field;
- permissible range of values (i.e from 0 to 100 per percentage).

101 Typically, the information to be provided is aggregated for tables that have relationships between them. A table containing the information for the monitoring stations/areas, one for the values of the monitored parameters. In the guide to compiling the information standard, information on the relationships between the different tables must be provided. There may be static information defined once for all (the list of substances to be monitored or the list of monitoring station codes). Such lists, generally called master records, will have to be included in the technical guidelines. The format for the information standard can be CSV, XML or EXCEL.

PHASE II. Transmission and collection of data deriving from monitoring programs through Information Standards

102 Data providers, after obtaining the Information Standards made available by the information system in accordance with paragraph 1), compile the data resulting from the monitoring itself. The properly compiled information standard form represents the format by which the monitoring data is collected by the operators and transmitted to the Information System. Transmission modes will be all implemented and could be chosen by the users according its level of computerization.

103 Generally, the following modes are identified:

- (A) the monitoring data is inserted into a CSV, XML or EXCEL file and sent to the system by manual uploading on a dedicated web page;
- (B) the monitoring data is inserted into a CSV, XML or EXCEL file and made available on a folder of a server with regularity to be agreed upon. The information system retrieves periodically the available files on the server folder ;
- (C) monitoring data is inserted into a CSV file and transmitted in real time to the information system by automated procedures.

104 Each of the modes described above has advantages and disadvantages:

- Case A) The operator has greater control over the data in the compilation and transmission phases, but must manually carry out the transmission.
- Case B) It is not necessary to proceed with manual transmission but it is sufficient to deposit the files into a shared folder. This, however, involves the presence of a network server that is always accessible and managed by the operator.

Case C) Suitable for telemetry stations, the monitoring operator has little control over the transmission but is responsible for the physical maintenance of the monitoring station.

PHASE III. Control, validation and aggregation of data

105 Once you have collected data at the information system using one of the methods described in step 2 above, you must check and validate the information. In general, it is possible to identify the following types of controls:

Formal checks:

- data compliance checks on the single field:
 - field format (text, numeric, yes / no, date);
 - value entered according to the list of possible values;
 - entered value within the permissible range;
- data compliance checks on single field column:
 - single value entered in relation to the values entered across the table column (eg univocity of the station code).
 - data compliance checks against values entered in different tables (relational controls) i.e. the congruity of the station code in the station table with respect to the station code entered in the parameter values table.

106 Consistency checks:

- parameter-specific range checks (i.e. 0-14 for PH);
- checks on the correct geo-localization of monitoring stations;
- checks on the consistency of the list of values entered with respect to the list of eligible values (eg, species monitored with respect to the list of species);
- checks on QA / QC in general.

107 Typically, formal controls derive directly from the definitions in the information standard and are easily translatable in automated software procedures. Some consistency checks can be made automated by ad hoc software procedures, while consistency checks on proper geo-localization presupposes the use of GIS software by an operator and are performed manually. These GIS controls will be performed using GIS desktop software tools. The type of controls and the resulting execution procedures presuppose the existence of a development team devoted to this activity, which also carries out the appropriate form of return on the audit results to the monitoring actors so that they can correct any found errors.

108 In the work flow it is foreseeable that the data transmission and control / validation phase is repeated iteratively until an acceptable level of data quality is obtained.

109 The list of quality control procedures for Information Standards is defined by INFO/RAC with the contribution of other MAP components and agreed with Contracting Parties.

PHASE IV. Publication and sharing of data

110 The checked and validated, properly aggregated and selected data set must be made publicly available according to the existing Data Policy. The consultation methods are predictable either by means of an analytical search interface (inserting values to be searched in some predefined fields, i.e the name of the monitoring station and / or monitored parameter), or by using a GIS interface with the option of selecting one or multiple station points or areas simultaneously. The system will have to allow the download of selected data through the search as well as its possible sharing via web services to selected community platforms (i.e EMODNET, WISE-Marine, etc.)

3. User Profiles

111 The data processing phases allow you to foresee the following user profile set (each user is characterized by login credentials: username and password):

Administrator

112 It can perform any function available on the system and is responsible for managing users (create / delete users with username and password) for all profiles, including ADMINISTRATOR profile.

Operator

113 It can perform the following functions:

- access and download of information standards and guidelines;
- uploading data in format compliant with information standards;
- access and download of your data uploaded to the system;
- access and download of audit data outcomes;

Operator Controls

114 It can perform the following functions:

- access and download of information standards and guidelines;
- uploading data in format compliant with information standards;
- access and download to all data uploaded to the system;
- implementation of control procedures;
- loading checks on the system;
- access to and download of the results of all data controls;
- access to consolidated data according to the data policy agreed with the Contracting Parties .

Viewer

115 It can perform the following functions on the basis of agreed data policy with the Contracting Parties:

- access and download of information standards and technical guidelines;
- access and download to all data uploaded to the system;
- access and download of audit results for all data.

Public Subject

- 116 It can perform the following functions:
- access to consolidated data made accessible to everyone.