Project Title:	ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS
Project number:	641762
Project Acronym:	ECOPOTENTIAL
Proposal full title:	IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS
Type:	Research and innovation actions
Work program topics addressed:	SC5-16-2014: “Making Earth Observation and Monitoring Data usable for ecosystem modelling and services”

D5.2 Metadata for pre-existing datasets

Version:	v1
Main Authors:	Poursanidis Dimitris (FORTH), Johannes Peterseil (EAA), Christoph Wohner (EAA), Chrysoulakis Nektarios (FORTH), Florian Wetzel (MfN), Joaquim Alonso (ICETA), Pedro Castro (ICETA), Carl Beierkuhnlein (UBT), Asja Bernd (UBT), Alaitz Zabala (UAB), Joan Masó (CREAF), Cristina Domingo (CREAF), Ole Vetaas (UiB),Tessa Bargmann (UiB), Samuel Bosch (UNESCO).

This project has received funding from the *European Union’s Horizon 2020 research and innovation programme* under grant agreement No 641762

Table of Contents

1 SUMMARY 5

2 INTRODUCTION 7

2.1 METADATA FOR DATA DRIVEN SCIENCE 7
2.2 THE ECOPOTENTIAL COMMUNITY 8

3 STATE OF THE ART 9

3.1 RELEVANT METADATA SCHEMES 9
3.1.1 Dataset documentation 9
3.1.2 Documentation of data quality 14
3.1.3 Documentation of observation sites 15
3.1.4 Other relevant meta-information 18
3.2 EUROPEAN AND GLOBAL CATALOGUES INTEGRATING METADATA 18
3.2.1 GEOSS (Group on Earth Observation System of Systems) 18
3.2.2 LifeWatch 20
3.2.3 INSPIRE 22
3.2.4 Data One 23
3.2.5 LTER Europe & DEIMS-SDR 24
3.2.6 GBIF Integrated Toolkit 27
3.2.7 EuMon 28
3.2.8 SeaDataNet 29

4 IDENTIFICATION OF COMMUNITY METADATA NEEDS 32

4.1 QUESTIONNAIRE ON METADATA SCHEMES AND NEEDS 32
4.2 RESULTS ON METADATA SCHEMES AND NEEDS 33
4.3 METHOD APPLIED TO EVALUATE RELEVANT QUALITY ELEMENTS. 36

5 ECOPOTENTIAL COMMUNITY PROFILES 39

5.1 OBSERVATION FACILITY [SITE] 39
5.1.1 Definition 39
5.1.2 Recommended MD elements for observation facilities 39
5.1.3 Mapping of site information to INSPIRE EF Schema 52
5.1.4 Possible Harvesting Mechanisms of site information 53
5.2 DATA PRODUCT [DP] 54
5.2.1 Definition 54
5.2.2 Recommended MD elements for data products 55
5.3 DATASET [DS] 61
5.3.1 Definition 61
5.3.2 Recommended MD Elements for In-Situ Data 61
5.3.3 Recommended MD Elements for EO Data 70
5.3.4 Recommended MD elements for documentation of Data Quality 78

6 TECHNICAL IMPLEMENTATION 93

6.1 METADATA AUTHORING 93
6.2 DISCOVERY CATALOGUE FOR IN-SITU DATA 94

7 STATISTICS OF THE AVAILABLE DATASETS 96

8 DISCUSSION 98

8.1 METADATA AND ESSENTIAL BIODIVERSITY VARIABLES 98

8.2 DEIMS-SDR AND MARINE SITES 98

8.3 LINKING METADATA FOR SCIENCE AND RESEARCH 99

9 REFERENCES 101

9.1 LIST OF TABLES 102

9.2 LIST OF FIGURES 103

10 ANNEXES 104

10.1 ANNEX 1 LIST OF DOCUMENTED DATA PRODUCTS IN DEIMS-SDR 104

10.2 ANNEX 2 LIST OF DOCUMENTED DATASETS IN DEIMS-SDR 110

Terms and abbreviations

ABCD	Access to Biological Collection Data
CSW	Open Geospatial Consortium (OGC) standard on Catalogue Web Service (CSW)
DC	Dublin Core
DEIMS-SDR	Dynamic Ecological Information Management System Site and Dataset Registry
DP	Data product (as defined in DEIMS-SDR)
DQ	Data quality
DS	Dataset (as defined in DEIMS-SDR)
DwC	Darwin Core
EML	Ecological Metadata Language
EO	Earth Observation
GEO	Group on Earth Observation
GEOSS	GEO System of Systems (Data integration)
ISO	International Standardisation Organisation
KVP	Key/value pair
LTER	Long Term Ecological Research
MD	Metadata
PA	Protected area
RDA	Research Data Alliance
RDF	Resource Description Framework
SI	Site or observation facility (as defined in DEIMS-SDR)
TDWG	Biodiversity Information Standards, aka Taxonomics Databases Working Group

1 Summary

The ECOPOTENTIAL project covers a wide range of protected areas (hereafter PA's) in different climate and biogeographic zones in Europe, South Africa and Caribbean Sea. What is common among them are the in-situ data that each Protected Area collects. These data can be used for monitoring purposes (biodiversity, ecosystem, and climate), for use in Ecosystem and Ecological modelling and for Earth Observation data analysis in order to understand the ecosystem and species functions and distributions as well as the impacts of the human activities on the natural resources. All above mentioned tools and methods have a unique target, to secure them for the next generations.

In-situ data are collected using various tools and methods whilst are referred to a specific period (e.g. year, month, larger periods). Thus, the description of these, by metadata, should be a mandatory step in order to allow other end users for further use of the in-situ data. Metadata describe the data per se, whilst can include several additional information that can be crucial for a specific application but not for another type of analysis.

Scientists from the Protected Areas, due to an overwhelming of diverse duties and activities, often overlook the importance and necessity of the metadata avoiding the preparation of them. On top of that, due to the vast diversity of different standards and protocols, the choice and use of the "most suitable" is another difficult task for them. On the other hand, scientists such as modellers, earth observation analysts, or ecologists who will make use of the in-situ data require metadata in order to understand and evaluate them for the analysis performed. Nevertheless, for the resulting data products metadata need to be provided (e.g. including provenance). This also adds an additional task to the processing and analysis chain of the data analysts (e.g. modellers).

The present study aimed to evaluate community requirements on metadata provision and enabling the reuse of data from scientific workflows and monitoring programs. This should result in recommendations for the provision of metadata across different communities. We started with the evaluation of existing metadata schemes and protocols, as well as a survey on the knowledge and application of metadata within the ECOPOTENTIAL community. We also focused on the identification of relevant existing in-situ data for the Protected Areas that are participating in the project. The work included the assessment of the awareness of the scientists (participating in the project) with regard to metadata and data quality for in-situ data.

The quality of data—that is, the degree of congruency between recorded data and current conditions that the historical data represent—is a central issue for global monitoring and assessment. It influences the accuracy of our descriptions of historical and contemporary patterns, determining our ability to provide realistic models of the future impacts of environmental change. Consequently, controlling for biological data quality is becoming increasingly important as advances in information technology promote ever-faster gathering and access to biodiversity information.

We used **DEIMS SDR** (***D**ynamic **E**cological **I**nformation **M**anagement **S**ystem **S**ite and **D**ataset **R**egistry*) for the documentation of research facilities, data products and datasets. DEIMS-SDR is a web based platform where each PA can register environmental monitoring facilities, data products/activities and datasets. It allows exporting e.g. site information to a INSPIRE EMF (http://inspire.ec.europa.eu/Themes/120/2892) compliant metadata record. The aim of DEIMS-SDR is to be the globally most comprehensive catalogue of

environmental research facilities, featuring foremost but not exclusively information about all LTER sites on the globe and providing that information to science, politics and the public in general. In addition, we adapted the metadata models applied in DEIMS-SDR to the needs of the project. For the surveys, we used online tools such as Google Forms with fixed semi-closed questions to assess the awareness on metadata and data quality and circulated the survey among the ECOPOTENTIAL members.

The results from the surveys, as well as the experiences with DEIMS SDR show that proper data documentation is a key prerequisite in order to make in-situ (e.g. biodiversity) data discoverable, accessible, and reusable beyond the original purposes for which they were collected. The absence of metadata degrades the usefulness of the in-situ data and often makes them not useable in other contexts. Therefore, raising the awareness for scientists on the importance of metadata as additional information on the data they are using is one of the tasks of this work. Nevertheless, the existence of a wide range of standards, their complexity and their extent often hampers the provision of metadata. It is important to find a compromise between the completeness for the use and easiness for the provision. The report provides recommendations for the documentation of research facilities, data products and datasets based on existing standards. We also focused on metadata elements describing the data quality.

2 Introduction

2.1 Metadata for data driven science

The importance and challenges of habitability and land sustainability imply environmental observation and monitoring (GEOSS, COPERNICUS, Earth Systems Data Cube) as well as, environmental/ecological modelling (ME) relating of biodiversity management and ecosystem services assessment (BISE, GBIF, LIFEWATCH, IPBES). These purposes require the increase of (spatial) data capture, modelling, management, sharing and access (big, open and linked data) included in thematic applications, systems and (cyber) geographic information infrastructures development framework. The recent quantity and diversity of spatial data are associated with data life cycle changes and demands enhancement on data quality assessment, management processes, knowledge management, reinforce training opportunities/education, and individual/organizations capacity building needs, as well as improving communication and technical-political decision processes.

In the last decades, data quality assessment and management concerns influences data modelling, datasets/database management, but mainly metadata profiles definition and metadata catalogue fulfilment. In this period, concepts, benchmarks, standards and practices have evolved. This ranges from: i) the internal/intrinsic assessment of the data producer's/provider compliance with the previous and specified data model or production processes; ii) to external quality when evaluating the satisfaction/adequacy of the data to the request/purposes defined by each users for specific uses highlighting the usability or “fitness for use”. The scientific and technological community of (spatial) data quality assessment presents new conceptual approaches and methodological frameworks, methods and instruments/tools in the direct (total or partial) spatial data evaluation (ISO19157) namely, in the generation of indicators and representation of the spatial variability of quality elements. In recent years, efforts have been made in the meta-evaluation of external and (in)direct quality by the end-user(s), taking advantage of possibilities of documentation and communication of quality elements in metadata profiles. This requires simple and (semi)automatic metadata fulfilment, integration and metadata catalogue interoperability, in parallel, to the operationalization of the theoretical methods foreseen in the spatial data quality assessment and management standards (ISO19115, ISO19139, ISO19157 and ISO19158).

Metadata permits to describe, discover and communicate elements about (spatial) datasets and data quality elements in groups, communities and network contexts. Metadata promotes networks development and management namely at scientific, technical and political open collaborative frameworks. Quality is defined by ISO 8402 as the “totality of characteristics of a product that bear on its ability to satisfy stated or implied need”, and by ISO 9000 as the “degree to which a set of inherent characteristics fulfils requirements”. The purpose of describing the quality of geographic data is to facilitate the comparison and selecting of the dataset best suited to application needs or requirements as well as to reduce possible data degradation across time (Tessarolo et al., 2017). Complete descriptions of the quality of a dataset will encourage the sharing, interchange and use of appropriate datasets. Information on the quality of geographic data allows a data producer to evaluate how well a dataset meets the criteria set forth in its product specification and assists data users in evaluating a product’s ability to satisfy the requirements for their application.

Data producers are interested in the widest reuse of their data. This is only possible when (a) the datasets satisfy the requirements of the targeted users, and (b) the degree to which these requirements are fulfilled is documented. In the first, based on user’s requirements, specific targets on data quality are established that have to be achieved in the course of data production or transformation. The second aspect corresponds to

documenting the quality of the data that is delivered to the users. For each of these tasks a common way of expression is necessary that comprises an agreed terminology, evaluation, and reporting methods, both aspects lead to formalising data quality. The advantages of the metadata are related with the information that capture and while the production of them sometimes is time consumer for the researchers. Research challenges and technical advances was oriented to (semi)automatic fulfilment metadata as well as, publish and share metadata elements at real/opportune time and agile way. Data providers and data user/usage recognition permits initial fulfilment of metadata fields/elements and highlights the advantages of automatic edition as well as, the regular and opportune updating of relevant data quality elements.

2.2 The ECOPOTENTIAL community

52 partners coming from Universities and Research Institutes and Protected Areas Management Bodies/Management authorities from 24 sites across Europe, South Africa and Caribbean Sea constitute the ECOPOTENTIAL community. A detailed description for each partner can be found in the official website of the project (see http://www. ECOPOTENTIAL-project.eu/partners) bringing expertise on different fields ranging from Earth Observation Analysis, Ecological Modelling, Environmental Modelling, and In-situ Scientist (Ecologist/Biologist).

ECOPOTENTIAL include a number of Protected Areas which cover several climate zones and biogeographic regions (Fig. 2-1) in Europe and beyond and have been classified into three distinct main ecosystem types: a) mountain ecosystems, b) transitional zones between aquatic (including marine) and terrestrial ecosystems, and c) semi-arid and arid ecosystems.

Fig. 2-1 The protected areas that participate in ECOPOTENTIAL.

The protected areas[1] have several statuses of protection; Fig. 2-1 show for each site the protection status as well as the biogeographic region the site belongs to. Most of them have more than one, with the common being the status of the National Park and the Natura 2000 site classification.

With regard to metadata and data provision, all different groups act mainly as data provider as well as data user. The importance of metadata in order to make datasets discoverable, accessible and understandable is therefore for all of them an issue.

Experts on Earth Observation Analysis are focusing on the analysis of Earth Observation (EO) data of different spatial/temporal resolutions and satellite platforms. The derived data (e.g. land cover, NDVI, LST, albedo, soil moisture, DSM) are used to identify ecological processes and changes (e.g. on land cover) over broad spatial and temporal scales. In addition, these, data are also used as input layers to models. The provision of sufficient metadata is an important task in order to allow the discoverability and re-usability of the data.

The In-situ scientists provide data for calibration and validation of both the production of the variables from the analysis of Earth Observation data as well as the training of the Ecological Modelling/Environmental Modelling. The in-situ data provide the information on the ecosystem status and behaviour often over a longer time span. I.e. design and instrumentation of the observation are important information, which needs to be recorded. The provision of metadata in order to ensure the reusability of the data is an important task, also because in-situ observation often cannot be repeated or redone once the data are lost because of poor documentation of the datasets.

Experts on Ecological Modelling and/or Environmental Modelling use data and products that come from the in-situ or EO data providers in order to model environmental or ecosystem processes (nutrients cycle, natural hazards, climate change, biodiversity distribution, etc.) or to forecast their behaviour in future. Metadata describing the datasets used in the models are important to understand the context, resolution and quality of the data and their fitness to use for the problem addressed.

3 State of the art

3.1 Relevant Metadata Schemes

As starting point for the formulation of a common metadata profile for the ECOPOTENTIAL community, relevant metadata standards for the different information objects were evaluated. This included (a) the documentation of datasets, (b) the documentation of data quality, (c) the documentation of research sites and research facilities, and (d) other relevant metadata standards.

3.1.1 Dataset documentation

Dublin Core metadata (DC)[2] is a small set of vocabulary terms that can be used to describe web resources (e.g. video, images, web pages) as well as physical resources (e.g. books, CDs) or objects like artworks. The DC metadata terms are managed by the Dublin Core Metadata Initiative (DCMI)[3]. The basic elements[4] defined

[1] see http://www.ecopotential-project.eu/2016-05-24-14-52-12/protected-areas

[2] see http://dublincore.org/

[3] see http://dublincore.org/documents/dcmi-terms/

[4] see http://dublincore.org/documents/dces/

in DC could be used to annotate any information resource by:

- *title* as the name given to the resource
- *identifier* as an unambiguous reference to the resource within a given context
- *description* of the resource
- *subject* as the topic of the resource (e.g. keywords)
- *language* as the language of the resource
- *type* as the nature or genre of the resource (defined by http://dublincore.org/documents/dcmi-type-vocabulary/)
- *creator* as the primarily responsible entity for making the resource
- *contributor* as any entity making contributions to the resource.
- *publisher* as the entity responsible for making the resource available (e.g. metadata creator)
- *date* as point or period of time associated with an event in the lifecycle of the resource (e.g. metadata date, publication date)
- *coverage* as the spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
- *format* as the file format, physical medium, or dimensions of the resource
- *relation* as the relation to a related resource
- *rights* as the information about rights held in and over the resource
- *source* as the related resource from which the described resource is derived

These metadata elements can be used to describe ecological data as well as mapping existing MD schema to a common metadata model.

Data Catalogue Vocabulary (DCAT)[5] is W3C standard. DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogues published on the Web. By using DCAT to describe datasets in data catalogues, publishers increase discoverability and enable applications easily to consume metadata from multiple catalogues (Fig. 3-1). It further enables decentralized publishing of catalogues and facilitates federated dataset search across sites. Aggregated DCAT metadata can serve as a manifest file to facilitate digital preservation.

The DCAT class Dataset [dcat: Dataset] is defined as collection of data, published or curated by a single agent, and available for access or download in one or more formats. For each dataset information on *title, description, language, identifier, contact point, distribution, frequency, keyword, landing page, publisher, release date, spatial coverage, temporal coverage, theme, update date* can be given. DCAT is not domain specific but could be applied to any ecological dataset integrating information from more detailed metadata schemata's as the EML, Ecological Metadata Language.

[5] see https://www.w3.org/TR/vocab-dcat/

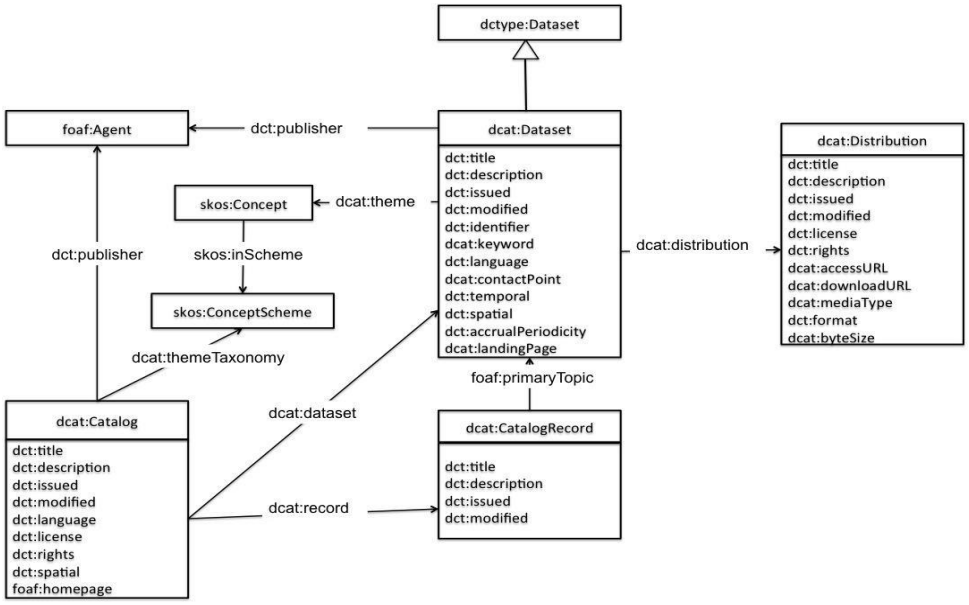

Fig. 3-1. Information model for DCAT (see https://www.w3.org/TR/vocab-dcat/)

INSPIRE and INSPIRE MD Specification[6] is a European Community Directive, which entered into force in May 2007. The INSPIRE directive defines the guidelines for the establishment of a spatial data infrastructure in Europe in order to support the Community environmental policies, and policies or activities why may have an impact on the environment. The data infrastructure is based on the infrastructures for spatial information established (and?) operated by the 27 member states of the European Union. The descriptive metadata are based on ISO19115/19139 as defined in the INSPIRE Metadata regulation (2008)[7]. For the implementation, technical guidelines and implementing rules have been specified. The INSPIRE MD Specification is describing metadata elements for datasets and dataset series as well as for data services; this can encompasses the e.g. following fields in Tab. 3-1.

Tab. 3-1. *INSPIRE Metadata elements (based on http:// INSPIRE.ec.europa.eu/documents/ INSPIRE-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1)*

Metadata element	*Dataset*	*Services*
Resource title	x	x
Temporal reference	x	x
Responsible organisation	x	x
Geographic bounding box	x	x
Resource language	x	
Topic category	x	
Spatial resolution	x	x
Resource abstract	x	x
Temporal extent	x	
Resource locator	x	x

[6] see http://inspire.ec.europa.eu/document-tags/metadata

[7] see http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32008R1205

Metadata element	*Dataset*	*Services*
Metadata language	x	x
Metadata point of contact	x	x
Metadata date	x	x
Resource type	x	x
Unique resource identifier	x	
Keyword	x	x
Conformity	x	x
Conditions for access and use	x	x
Limitations on public access	x	x
Coupled resource		x
Spatial data service type		x

LTER Europe adopted beside EML the INSPIRE MD specification as the main supported MD schemata for environmental data in the ecological domain.

ISO 19115/19139 is a generic metadata schema for describing geographic information and services by the means of metadata. It provides information about the identification, the extent, the quality, the spatial and temporal aspects, the content, the spatial reference etc. of digital geographic data and services. The standard consists of different parts namely *ISO19115-1:2014 Fundamentals, ISO 19115-2:2009 Extensions for imagery and gridded data* and *ISO19139:2007 Metadata XML schema implementation.* ISO19115/19139 is the underlying metadata schema adopted by the INSPIRE directive. As ISO19115 generically can be used to describe any geospatial dataset it can also be applied for any observational data having a spatial context. Nevertheless limitations in the documentation of methodological aspects may arise.

Ecological Metadata Language (EML) is a metadata specification for data resulting from the ecological domain (Mitchener et al. 1997). EML is implemented as a series of XML document types that can be used in a modular and extensible manner to document data. Each EML module is designed to describe one logical part of the total metadata that should be included with any ecological dataset[8]. EML was adopted by ILTER and GBIF as the main supported metadata standard.

Darwin Core (DwC)[9] is a Biodiversity Information Standards (TDWG) standard for sharing biodiversity data (Wieczorek et al. 2012). DwC is a set of terms, which can be seen as extension of the Dublin Core metadata standard for the biodiversity domain. DwC supports four types of data sources:

- *resource metadata* - where metadata on an available dataset are provided
- *checklist data* - where a list of species for a given area is published
- *occurrence data* - where occurrence observations are published

[8] see http://knb.ecoinformatics.org/

[9] see http://rs.tdwg.org/dwc/

- *Sampling event data* - where complex species observations (e.g. butterfly survey or vegetation survey) data are published.

The data type 'resource metadata'[10] describes a biodiversity information resource (of any other type) where currently no digital data are shared. For this DwC provides a pure set of metadata elements. *Tab. 3-2* shows the list of required and optional information which is based on the DC elements.

Tab. 3-2. Minimum set of DwC terms for resource metadata

DwC Element	***required***	***optional***
Title	x	
Description	x	
publishing organisation	x	
Type	x	
License	x	
contact(s)	x	
creator(s)	x	
metadata provider(s)	x	
sampling methodology		x
Citation		x

Its main use is the publication of specimen and observation records (e.g. adopted by GBIF and OBIS) but lately extends to include the needs for sharing more complex sample based (DwC Event) and ecological and environmental data bundled into DwC Archive (DwC-A). DwC includes metadata and data in one file within GBIF. DwC files are described by additional EML metadata.

Meta(data) standard for biological and geological collections – (ABCDEFG)[11] is a Biodiversity Information Standards (TDWG) standard providing more comprehensive information on observation data. It covers metadata on the dataset, the observation event, the identification of species information and the measurement. Mapping between ABCD and DwC exist in order to share information with GBIF. An extension of ABCD covers objects and collection units of geosciences (EFG). Metadata is an essential part of information also for specimens of museum collections, as it gives information on the collecting event (e.g. the time, region and the collector) and further information that is needed for a scientific analysis of data (identifier person, references to the digital representation of the specimen or observation). There are also metadata elements that describe the associated collection, holding institution and contact persons. ABCD consists of about 677 elements holding the data of the individual collection plus 72 elements with metadata about the entire dataset (Holotschek et al 2012). It is a standard that is XML-based and that was developed for structuring data elements, it is machine-readable but can also be read as a plain text. One important feature of ABCD is its hierarchical structure, this means that it clusters related items into semantic groups that relate to real-world objects or concepts (Holotschek et al 2012). The standard ABCD and its extension EFG are constantly

[10] see https://github.com/gbif/ipt/wiki/resourceMetadata

[11] see https://github.com/tdwg/abcd

improved, e.g. to further process XML-based standards so that they meet the requirements of the Semantic Web (Güntsch et al. 2016).

Generic Earth Observation Metadata Standard (GEOMS) metadata and data structure requirements developed to facilitate the use of geophysical datasets by improving their portability and accessibility, and by making their contents self-describing. This approach was originally selected to deal with atmospheric and oceanographic datasets, but has been recently expanded to support all measurements from Earth observation instruments. The definitions have been carefully chosen to allow applicability to other scientific endeavours. GEOMS metadata and data structure requirements may be applied to any project where data are to be exchanged.

3.1.2 Documentation of data quality

(Spatial) data quality is a central/critical issue in project and network management namely, in promoting clear and effective communication between data providers (and data users (researchers, technicians, political and strategic decision-makers). The development of spatially explicit models implies to choose and integrate adequate data and present results/ effective proposals supported on data quality assessment and management principles and procedures. Data quality and data quality evaluation has received attention in geographical information science community (ISO/TC 211 Geographic information/Geomatics) with standards and guidance's implementation. The last decades (in particular the last five) years big, open and linked data trends highlight the dynamics and need of formalizing spatial data quality orientations and standards. Recent ISO standards data quality defines concepts and methods of spatial data quality assessment related to quality insurance of data supply.

ISO 19157:2013 "Data quality", replaces ISO/TS 19138:2006 "Data quality measures", ISO 19114:2003 "Quality evaluation procedures", and ISO 19113:2002 "Quality principles", and establishes the principles for describing the quality of geographic data. It i) defines components for describing data quality; ii) specifies components and content structure of a register for data quality measures; iii) describes general procedures for evaluating the quality geographic data; iv) provides guidelines on how to describe, evaluate and report data quality. ISO 19157:2013 also defines a set of data quality measures for use in evaluating and reporting data quality. It is applicable to data producers providing quality information to describe and assess how well a data set conforms to its product specification and to data users attempting to determine whether or not specific geographic data are of sufficient quality for their particular application.

ISO/TS 19157-2:2016 "Data quality - Part 2: XML schema implementation", defines data quality encoding in XML. It is an XML schema implementation derived from ISO 19157:2013 and the data quality related concepts from ISO 19115-2:2009 "Metadata - Part2: Extensions for imagery and gridded data". ISO/TS 19157-2:2016 utilizes encoding rules from ISO 19118:2011 "Encoding" and ISO/TS 19139:2007 "Metadata - XML schema implementation", and the implementation approach from ISO/TS 19115-3:2016 "Metadata - Part3: XML schema implementation for fundamental concepts" to define an XML schema implementation of ISO 19157:2013, and the data quality related concepts from ISO 19115-2:2009.

ISO/TS 19158:2012 "Quality assurance of data supply" provides a framework for quality assurance specific to geographic information. It is based upon the quality principles and quality evaluation procedures of geographic information identified in ISO 19157 and the general quality management principles defined in ISO 9000. The framework defined in ISO/TS 19158:2012 enables a customer to satisfy itself that its suppliers, both internal and external, are capable of delivering geographic information to the required quality. Fundamental to the framework is the assurance of the supplier's ability to understand and meet the quality requirements. Through the quality assurance framework both the customer and the supplier are able to consider the quality required at the earliest opportunity in the production/update process. Principles and responsibilities of the relationship between the customer and the supplier that facilitate the framework are provided. The responsibility for the quality assessment procedure is shared between the customer and the supplier. ISO/TS 19158:2012 is applicable to customers and suppliers of all geographic information where the quality of the product may be impacted upon by the supplier's processes in any of the following scenarios: i) there is an agreement or legislation for the supply of data acquisition services, ii) data acquisition services are being tendered for, and iii) one or more suppliers exist in the supply chain.

3.1.3 Documentation of observation sites

LTER Site Metadata Model (SMM)[12] provides a set of metadata elements describing observation and experimental facilities for the environmental domain. Starting from the needs from the global ILTER network a set of elements was defined providing documentation about the organisation (e.g. organisation (e.g. contact, information, and network), the location, the characteristics (e.g. climate, habitats) or available equipment to name just a few. A 'Site' is defined as basic generic in-situ observation or experimentation facility (of any size and organisation level), where the collection of e.g. bio-physical, environmental, biological or socio-ecological data is conducted (Fig. 3-2).

The basic metadata elements are in line with the data specification of the INSPIRE data theme *Environmental Monitoring Facilities (EF)*. A mapping between the standards has been done within the ECOPOTENTIAL project implementing the publication of site meta-information as extended EF record. Details can be found in Chapter 5.1.3.

The Site Metadata Model encompasses the following metadata fields:

- *Name and general description*
- *Contact details*
- *Metadata providers*
- *Geographic location*
- *Ecosystem and environmental characteristic*
- *Network affiliation*
- *Site classification*
- *Status and history*
- *Focus, design and scale of site*
- *Protection status and resource management*
- *Infrastructure and operation*
- *Data sharing policy*

[12] see https://data.lter-europe.net/deims/documentation/site

- *Data management*

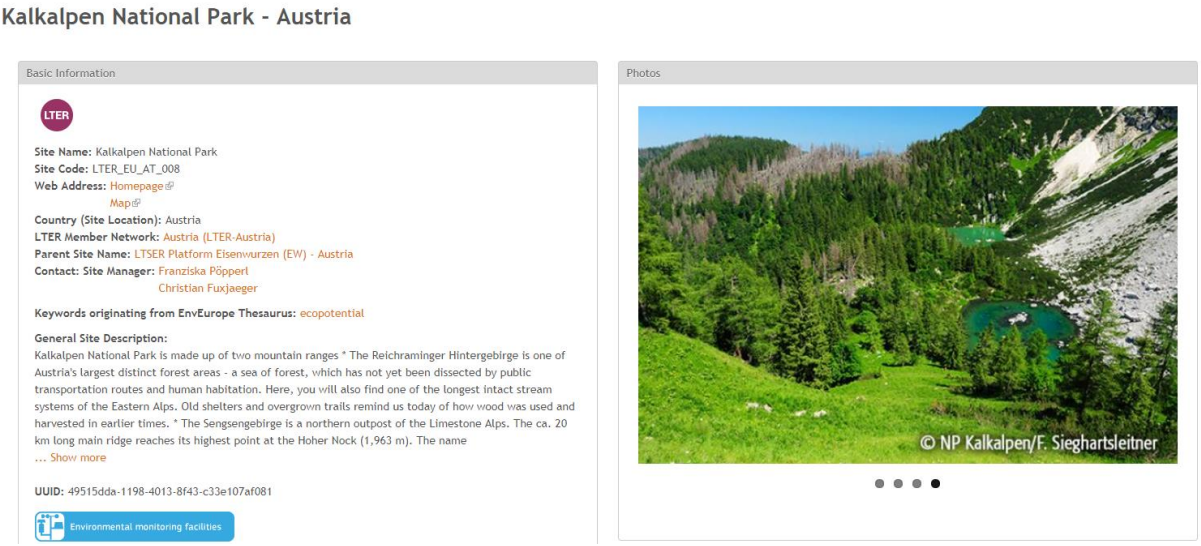

Fig. 3-2. Documentation of the protected area NP Kalkalpen (AT) using SMM in DEIMS-SDR (landing page)

With the DEIMS-SDR Site and Dataset Catalogue, an editor is provided for the documentation of observation and experimentation facilities. Currently extensions to include a persistent global unique identifier (PID) are under preparation.

INSPIRE Environmental Monitoring Facility [EF] Application schema[13] is the INSPIRE implementation guideline to encode spatial information on observation facilities and the resulting observation values in a harmonised manner. The legal definition of INSPIRE EF is "*Location and operation of environmental monitoring facilities includes observation and measurement of emissions, of the state of environmental media and of other ecosystem parameters (biodiversity, ecological conditions of vegetation, etc.) by or on behalf of public authorities [Directive 2007/2/EC]. A generic model [for monitoring facilities], which can be used across various domains and leave the necessary freedom to thematic domains to bring in specific needs while keeping a shared structure. So the data specification provides a common structure but not a thematic harmonisation across domains*"[14].

Environmental monitoring facilities are linked to information describing aggregations/collections of monitoring facilities and their thematic or organisational grouping and background. Additionally, environmental monitoring facilities link to observations and measurements taken. This is covered by using the Observations and Measurements standard tailored to the common structure in the INSPIRE data specifications. The EF model is a common framework to describe environmental monitoring facilities in a common way across domains.

In general, the aim of INSPIRE EF is to provide a standardised format to answer the following questions:

[13] see http://inspire.ec.europa.eu/Themes/120/2892

[14] see http://inspire.ec.europa.eu/id/document/tg/ef

- What is being observed?
- Where is it being observed?
- When is it being observed?
- Who is carrying out the observation?

The specification itself covers all kinds of environmental monitoring using fixed stations, moving equipment or remote sensing, however in the context of DEIMS-SDR it is currently limited to documenting fixed stations. INSPIRE EF has a multi-layered structure. The model provides a recursive hierarchical link between Environment Monitoring Facilities, thus allowing mapping the hierarchy of a platform, its sites, stations and sensors.

Sensor Markup Language (SensorML) is an Open Geospatial Consortium (OGC) standard that provides standard models and an XML encoding for describing sensors and measurement processes. It includes information for the sensor discovery, the sensor location, the processing of sensor observations, as well as mechanisms for the programming and the alert of sensors. SensorML is an intrinsic part of the *Sensor Observation Service (SOS[15])* which provides a standardized interface for managing and retrieving metadata and observations from heterogeneous sensor systems. It is part of the Sensor Web Enablement (SWE[16]) initiative and is an Open Geospatial Consortium (OGC) approved standard in version 2.0 (and in the previous version 1.0.0). The SOS is based on the exchange of standard messages (requests and responses) between a "service" and a "consumer", through the usage of the HTTP protocol. The requests can be sent in two different ways: through an HTTP POST method or through an HTTP GET method (a KVP). In both cases, the response of these queries is an XML file compliant with the specifications. According to OGC, a SOS service must implement the three mandatory requests for the SOS Core profile, while other operations of the Transactional profile, the Enhanced profile and the Result Handling Extension profile (introduced in version 2.0) are optional. The three mandatory operations for the Core profile are the following:

- *GetCapabilities*, that gives a self-description of the service.
- *DescribeSensor*, giving information about the sensor itself, encoded in a Sensor Model Language (SensorML) instance document.
- *GetObservation*, giving the pure sensor data encoded in Observation & Measurements (O&M) standard.

Thanks to these three operations, a Consumer can retrieve data from a specific Sensor in a SOS service.

In addition, some other terms and definitions need to be explained because all of them are the key objects for the SOS standard:

- *Procedure*: indicates the provider of observations (generally the sensor) and is represented as SensorML standard data model.
- *Observations*: represents the values measured at given time instants and represented according to the O&M

[15] http://www.opengeospatial.org/standards/sos

[16] http://www.opengeospatial.org/ogc/markets-technologies/swe

standard data model.

- *Observed Properties*: represents the observed phenomena (e.g. phenomenon, air-temperature) and is represented with a URI according to O&M standard.
- *Feature of interest*: is the feature that relates to the observations.
- *Offering*: represent a collection of sensors used to conveniently group them up.

The SOS is implemented to publish and share observation data in a harmonised manner. Common parameter names can be selected e.g. from a vocabulary or a list of terms.

3.1.4 Other relevant meta-information

To ensure the reusability of the data in scientific workflows information on the sources and the applied processing steps is important. This information should be captured from the workflows and provided as provenance for the dataset. To document the provenance of data different provenance models (and tools) are provided by the scientific community. The Open Provenance Model (OPM)[17], PROV[18], or Provenir (Sahoo 2011) to name just a few. While the concept of provenance is gaining importance in data driven science, these aspects were currently not addressed for the in-situ data, but would be important for the extension of metadata for model based results from the analysis WP.

3.2 European and global catalogues integrating metadata

Metadata are often provided by local or regional data holders and shared catalogues. By this a user can discover a wide range of information in either domain specific or general data catalogues. In the following chapter a number of relevant metadata catalogues are listed, where the ECOPOTENTIAL community can contribute data sources to.

3.2.1 GEOSS (Group on Earth Observation System of Systems)

Global Earth Observation System of Systems (GEOSS)[19] seeks to connect the producers of environmental data and decision-support tools with the end users of these products. It has the aim of enhancing the relevance of Earth observations to global issues. The result is focused on a global public infrastructure that generates comprehensive, near real-time environmental data, information and analyses for a wide range of users. One of the first achievements of the Group on EO was the acceptance of a set of high level Data Sharing Principles as a foundation for GEOSS. The 10-Year Implementation Plan says that "*the societal benefits of Earth observations cannot be achieved without data sharing*" and sets out the GEOSS Data Sharing Principles:

- There will be full and open exchange of data, metadata and products shared within
- GEOSS, recognizing relevant international instruments and national policies and legislation;
- All shared data, metadata and products will be made available with minimum time delay and at

[17] see http://eprints.soton.ac.uk/271449/1.opm.pdf

[18] see https://www.w3.org/TR/prov-overview/

[19] See https://www.earthobservations.org/geoss.php

minimum cost;

- All shared data, metadata and products being free of charge or no more than cost of reproduction will be encouraged for research and education.

GEO recognizes that the societal benefits arising from Earth observations can only be fully achieved through the sharing of data, information, knowledge, products and services. GEO has therefore promoted fundamental principles for data sharing, expanding the trend towards open data worldwide. Thus, as it embarks on its second decade, GEO aims to implement the following GEOSS Data Sharing Principles for 2015:

- Data, metadata and products will be shared as Open Data by default, by making them available as part of the GEOSS Data Collection of Open Resources for Everyone (Data-CORE) without charge or restrictions on reuse, subject to the conditions of registration and attribution when the data are reused;
- Where international instruments, national policies or legislation preclude the sharing of data as Open Data, data should be made available with minimal restrictions on use and at no more than the cost of reproduction and distribution; and
- All shared data, products and metadata will be made available with minimum time delay.

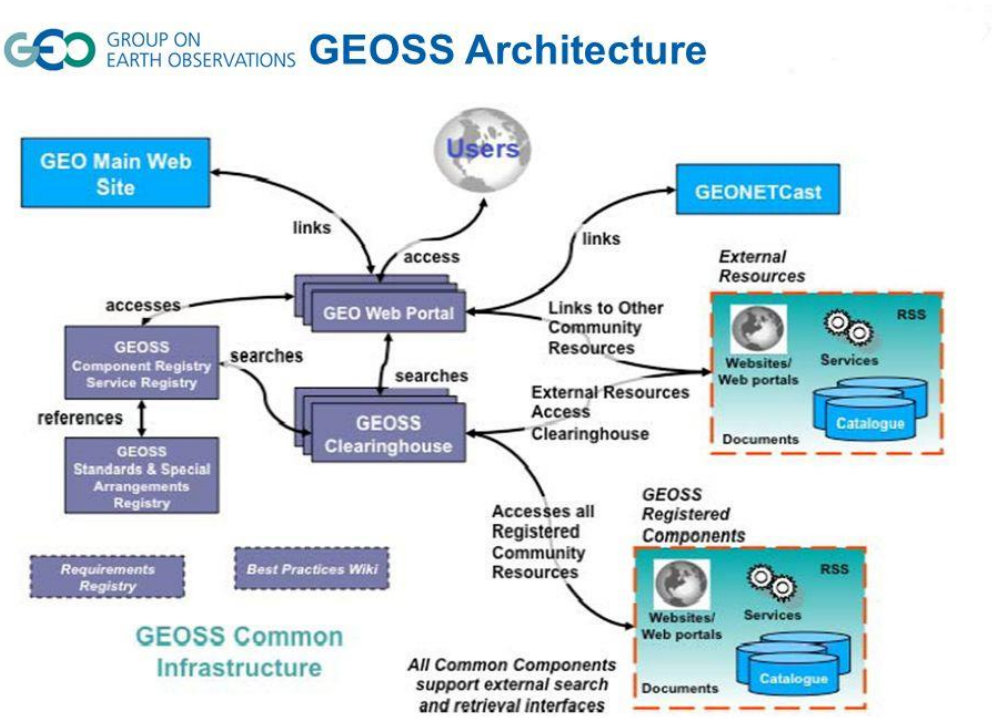

Fig. 3-3. GEOSS Architecture

Discovery broker (GI-cat, Fig. 3-4): a component, which is able to connect disparate (distributed and heterogeneous) metadata sources, exposing them through a set of standard catalogue, interfaces. By means of metadata harmonization and protocol adaptation, it is able to search metadata from different sources and transform query results to a uniform and consistent metadata model. GI-cat mediates among the connected metadata sources interfaces, and harmonizes their metadata mapping them to an internal schema based on ISO 19115 (GI-cat metadata model). Each query request sent through the external interfaces is performed

against all the connected sources based on the internal schema. GI-cat supports both distributed queries (for external sources exposing a catalogue service) and harvesting. Harvesting can be adopted for enhance query performances for catalogues, or to enable search also on inventory services providing metadata without catalogue functionalities. The choice between distributed query and harvesting can be made per data sources. In case of harvesting also, the repetition time can be defined per data source.

Fig. 3-4. GEOSS Discovery and Access Broker20

3.2.2 LifeWatch

The mission of LifeWatch is to advance biodiversity research and to provide major contributions to addressing the big environmental challenges, including knowledge-based solutions to environmental managers for its preservation. This mission is achieved by providing access through a single infrastructure to a multitude of sets of data, services and tools enabling the construction and operation of Virtual Research Environments (VREs) linked to LifeWatch, and where specific issues related with biodiversity research and preservation are addressed. LifeWatch was included in the Roadmap of the European Strategy Forum on Research Infrastructures (ESFRI[21]), the body that identifies the new research infrastructures (RIs) of pan-European interest with the goal of promoting the long-term competitiveness of European Research and Innovation. Its statutes and governance were modified following the comments provided by the European Commission in January 2014 resulting from the step-1 submission of LifeWatch ERIC application back in July 2013.

The LifeWatch e-infrastructure meets a number of key requirements:

- 'Fit for Purpose': flexible, secure, adaptable, robust, resilient, scalable, and maintainable.
- Integration of "external resources", provided by institutions and networks concerned with ICT

[20] see e.g. http://www.eurogeoss-broker.eu/

[21] see http://www.esfri.eu

technologies and biodiversity research.

- Offering an attractive set of capabilities to users and other stakeholders.
- User-friendly at different levels of knowledge in both science and policy domains.
- Non-proprietary – based on open standards (in application of EU Openness Directives).
- Based on existing technological solutions wherever appropriate, and Adaptive to the heterogeneous IT landscape of Europe wide research IT.
- In selected areas, parallel research into cutting-edge technologies to ensure adoption of new approaches and to contribute to ERA bioinformatics development.
- Staged approach to construction and deployment. Long-term outlook on all desired functionality within a realistic, controlled and manageable construction process.

Four major layers (Fig. 3-5) compose LifeWatch e-Infrastructure.

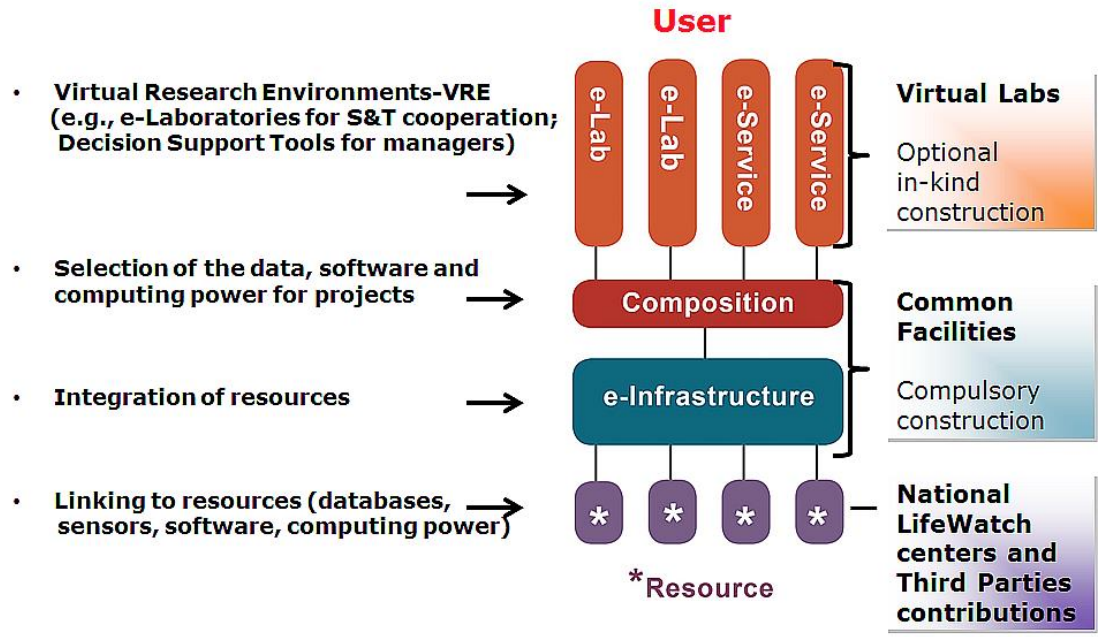

Fig. 3-5. Architecture of LifeWatch e-Infrastructure

The ***Resource layer*** contains the specific resources, such as data repositories and collections (i.e., LTER, GBIF, CETAF), computational capacity and sensor networks, and High Performance Computing (HPC) resources, which contribute to the LifeWatch system. It is supported by contributing facilities and in turn integrated in an e-Infrastructure layer which also serves shared workflows. The Resource and e-Infrastructure layers will incorporate tools from existing networks and e-Infrastructures such as, LTER, GBIF, CETAF, among others. This will provide a basis for interoperability between LifeWatch and other existing and future systems.

The ***e-Infrastructure layer*** enables to share the specific resources as generic services in a distributed environment spread across multiple administrative domains. Some of the capabilities of this layer will be provided by underlying Europe-wide e-Infrastructures (for example, EGI.eu and its supporting IBERGRID through a proper LifeWatch EGI.eu Competence Centre) and these will also play a prominent role in delivering the Composition layer of LifeWatch, under the coordination and supervision of the above-mentioned ICT e-Infrastructure Technical Office of the common facilities. Similarly, in order to ensure

commonality of data management modalities for differing data sets / data providers, the ICT e-Infrastructure will have to take into account the data management guidelines that are likely to emerge over the coming years.

The ***Composition layer*** supports the selection and combination of services for task completion. It offers resources for new workflow development in a semantic metadata frame.

The ***User laye***r enables the different research communities to create their own Virtual

Research Environments-VRE (e.g., e-Labs, decision-making tools, etc.); users may share their data and analytical and modelling tools with others while controlling access to them. At the Composition and User layers, LifeWatch expect to adapt and extend mechanisms from existing networks and e-Infrastructures, as well as from relevant e-Science projects.

3.2.3 INSPIRE

The INSPIRE directive came into force on 15 May 2007 and was implemented in various stages, with full implementation required by 2021 based on the roadmap[22]. The INSPIRE directive aims to create a European Union (EU) spatial data infrastructure. This should enable the sharing of environmental spatial information among public sector organizations and better facilitate public access to spatial information across Europe. A European Spatial Data Infrastructure assists in policy-making across boundaries. Therefore, the spatial information considered under the directive is extensive and includes a great variety of topical and technical themes.

INSPIRE is based on a number of common principles:

- Data should be collected only once and kept where it can be maintained most effectively.
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications.
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes.
- Geographic information needed for good governance at all levels should be readily and transparently available.
- Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

INSPIRE is based on the infrastructures for spatial information established and operated by the 27 Member States of the European Union. The Directive addresses 34 spatial data themes needed for environmental applications, with key components specified through technical implementing rules. This makes INSPIRE a unique example of a legislative “regional” approach.

[22] http://inspire.ec.europa.eu/index.cfm/pageid/44

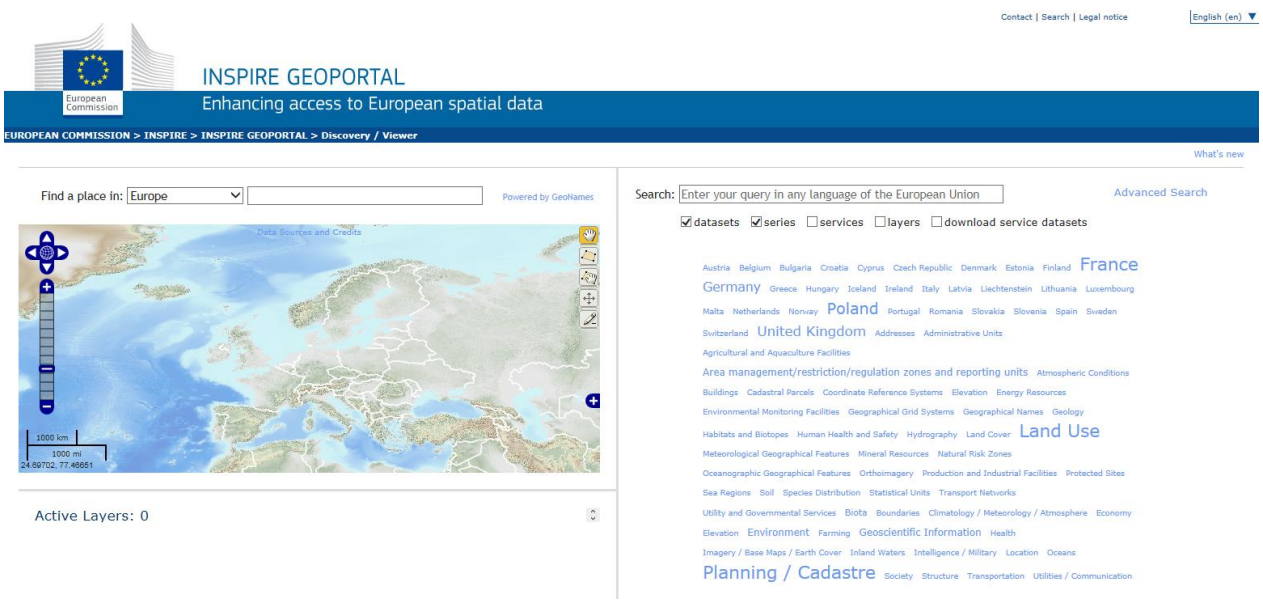

Fig. 3-6 INSPIRE Data Discovery Interface (see http://inspire-geoportal.ec.europa.eu/discovery/)

3.2.4 Data One

DataONE (Data Observation Network for Earth) is a network of different member repositories, the so-called member nodes. The aim of DataONE is to support enhanced search and discovery of earth and environmental data and, beyond that, universal access to the data itself (cf. https://www.dataone.org/). The global network (Fig. 3-7) builds on existing data centres, consists of coordinating nodes and member nodes and data can be searched by a common search interface. Most nodes use MetaCat, which is an open source metadata catalogue and data repository for scientific data. Some datasets can also be accessed via LTER, Dryad or other repositories. Dryad for example is a repository specifically for data that is used for peer reviewed scientific papers.

DataONE was developed for biological and environmental research with many different stakeholders, primarily researchers from academia, government, private industry and non-profit organizations (Michener et al. 2012). One of the key elements is also to provide metadata for the datasets in order to obtain a long-term data preservation and to make these data discoverable and accessible. Data providers can choose their relevant metadata standards. DataONE also gives recommendations for standards to promote best practices (https://www.dataone.org/best-practices/metadata). These recommendations cover for example guidelines on choosing appropriate standards and explain the importance of basic metadata needs (e.g. what fields/information are required for a sound data reporting).

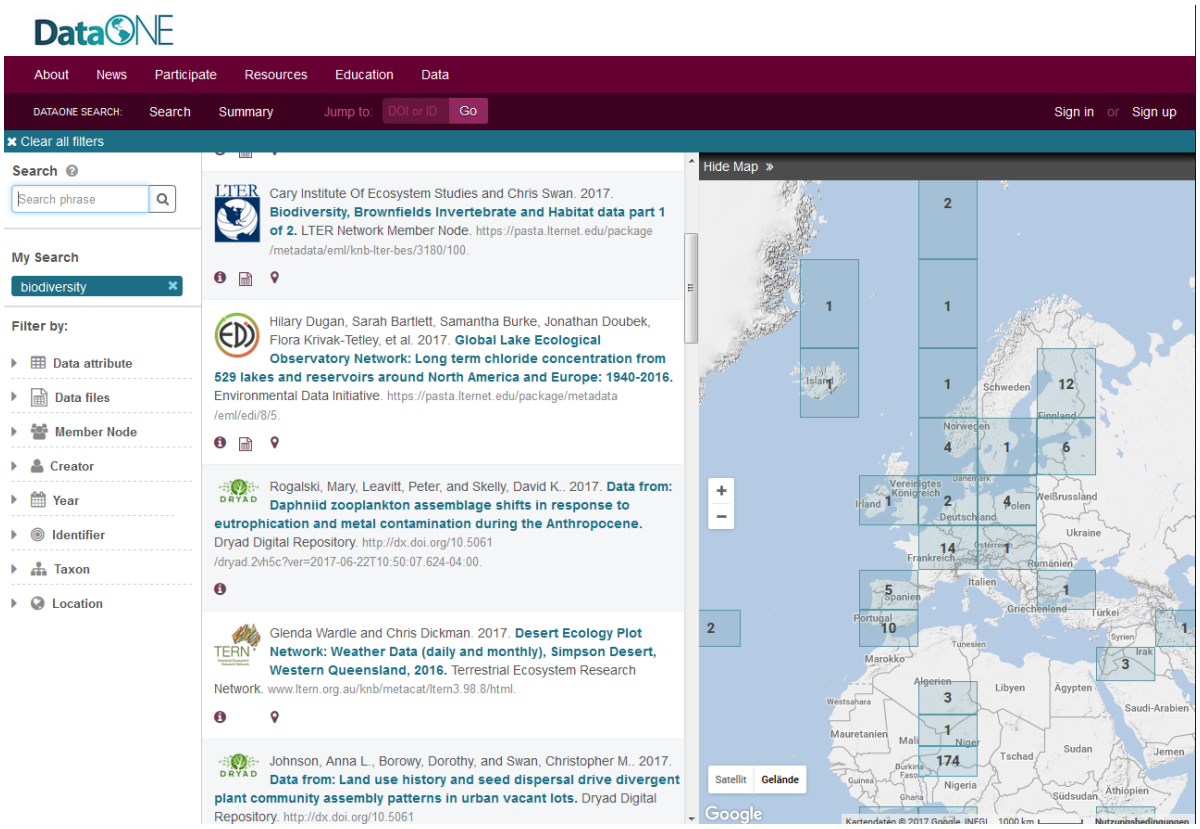

Fig. 3-7. Data One Search Interface for obtaining information on Earth and Environmental Data

3.2.5 LTER Europe & DEIMS-SDR

DEIMS-SDR (Dynamic Ecological Information Management System Site and Dataset Registry)[23] is the site catalogue of LTER Europe and also acts as a data node for dataset publication. US LTER, the University of New Mexico, the University of Puerto Rico, the University of Wisconsin, and Palantir.net using the content management system Drupal 7 developed the basic version of DEIMS. DEIMS-SDR core allows storing, editing, and sharing data and information about biological and ecological research and exposing that information in standardised dataset formats, such as ISO19139 (GMI and GMD), BDP (Biological Data Profile) and EML 2.1.1 (Ecological Metadata Language). Both LTER Europe and ILTER later on extended the core part in order to be capable of storing information about the sites where research is carried out. This information is exposed as INSPIRE EF. DEIMS-SDR was established in 2013 running on Drupal 6. In early 2017 a new version of DEIMS based on Drupal 7 was released. As of June 12th DEIMS-SDR hosts 1006 published site records. Site records include both LTER and non-LTER sites (46 non-LTER records).

[23] see https://data.lter-europe.net/deims/

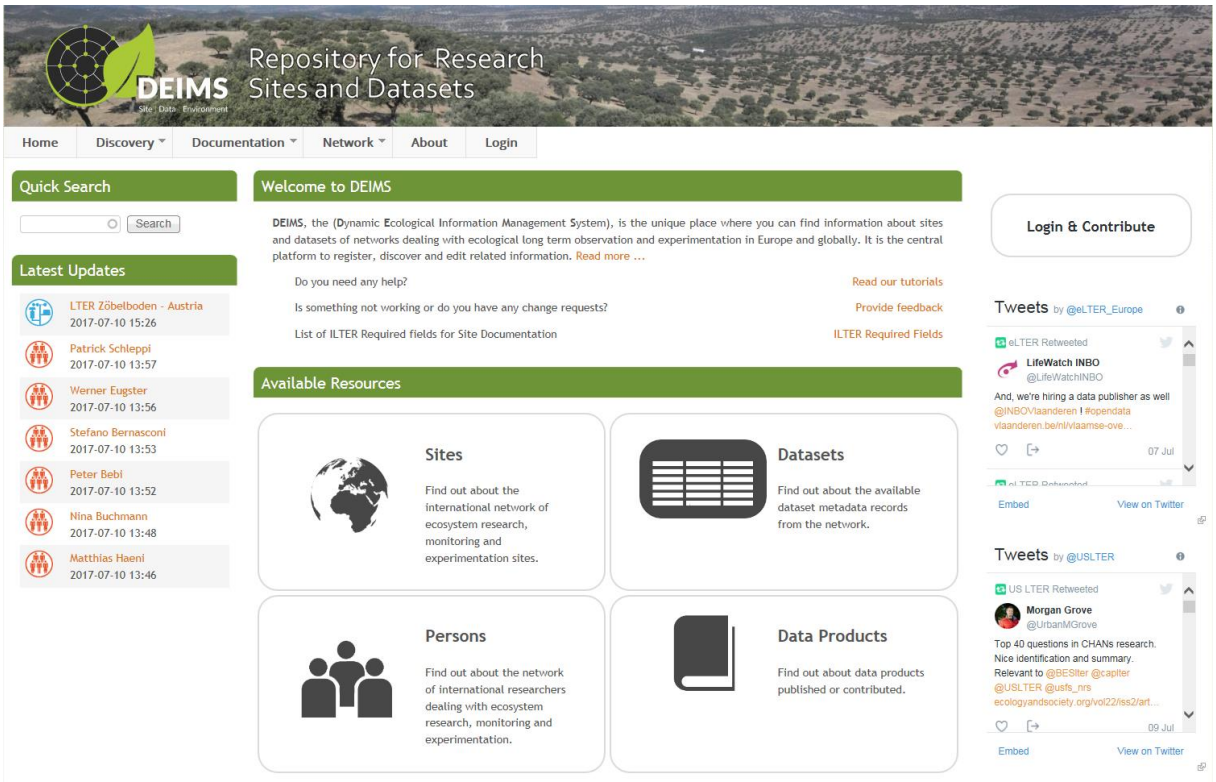

Fig. 3-8. DEIMS-SDR Site and Dataset Catalogue

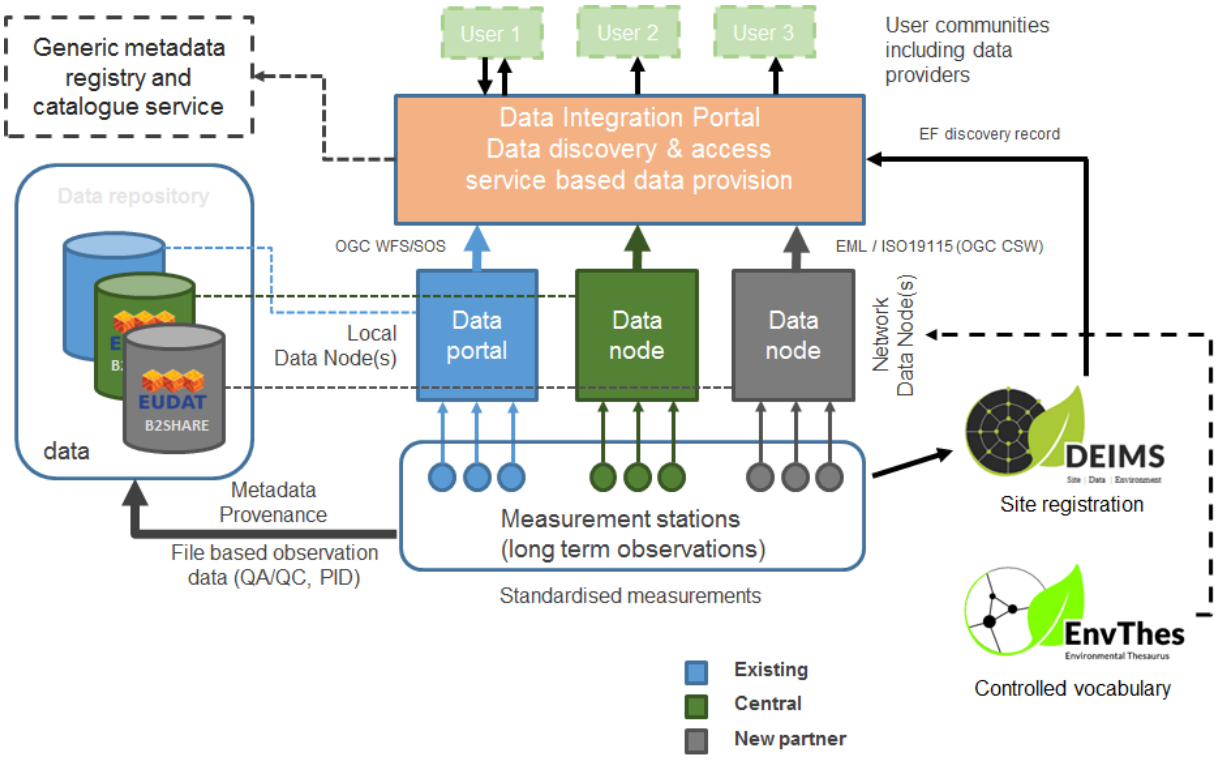

Fig. 3-9. Schematic architecture for the integration of data from the LTER network (reference eLTER H2020 project)

In addition to DEIMS, which primarily acts as a metadata editor, there is also the eLTER Data Integration Portal (DIP), which is currently being developed within the eLTER project. The aim of the DIP is to act as a central discovery portal of all information that is available in LTER Europe. All information available on DEIMS-SDR is already harvested by the DIP.

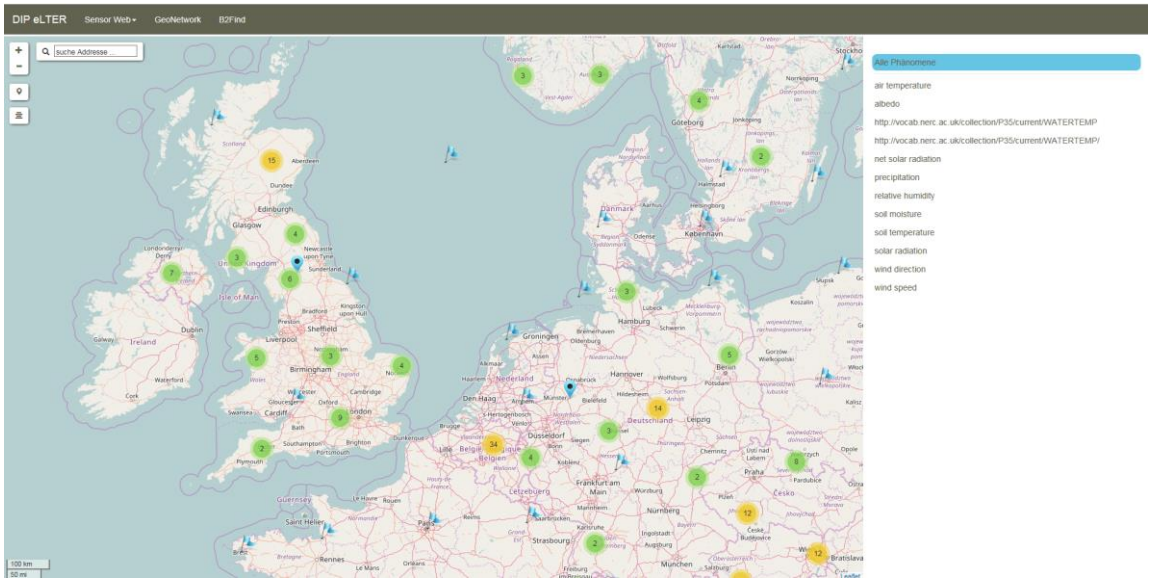

Fig. 3-10. eLTER Data Integration Portal (prototype)

In addition tools to visualise and download time series data (see Fig. 3-11) are implemented using Sensor Observation Services (OGC SOS) as provisioning services.

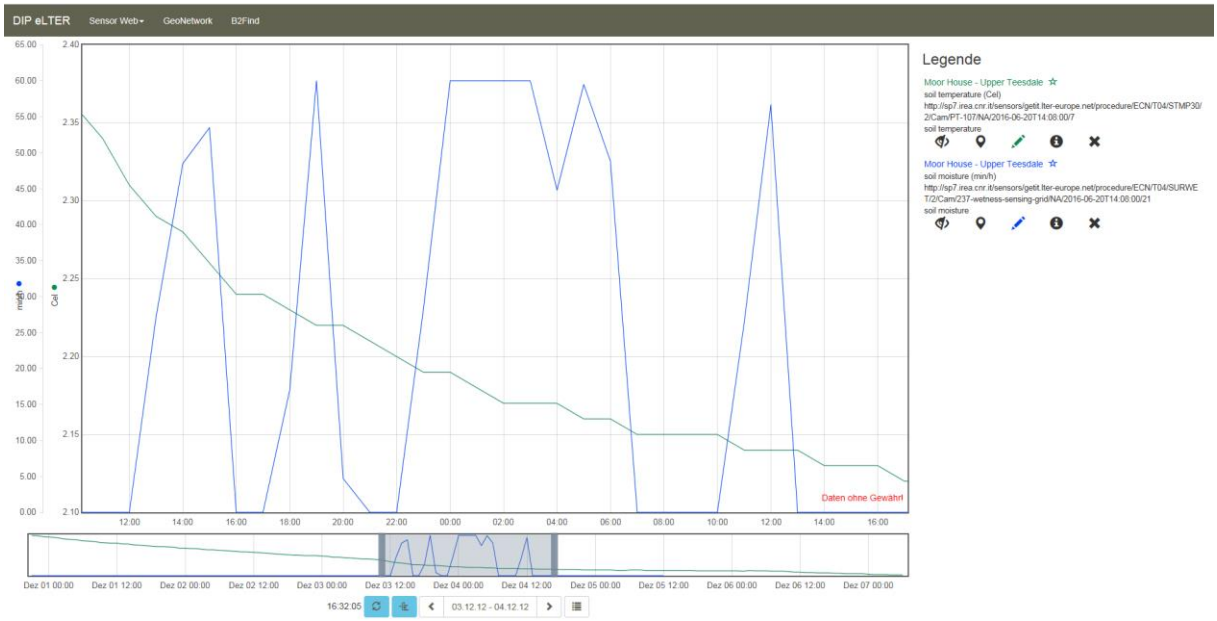

Fig. 3-11. eLTER Data Integration Portal (prototype) - time series viewer for sensor data

3.2.6 GBIF Integrated Toolkit

A way to upload biodiversity data in order to enable open access to metadata and data is provided by the Global Biodiversity Information Facility (GBIF). GBIF is an international open data infrastructure that is supported by several governments; it provides a research infrastructure for biodiversity data and can be accessed via a common portal (http://www.gbif.org/). GBIF fosters open access to data - the data can be uploaded to GBIF in several ways, via institutions that use middleware such as DiGIR, BioCASE or TAPIR and that provide data via GBIF nodes or GBIF tools such as the GBIF Integrated Publishing Toolkit (IPT). The GBIF IPT is also used in the Ocean Biogeographic Information System (OBIS), which focuses on marine data.

GBIF uses the standards Darwin Core (DwC) and Ecological Metadata Language (EML) and the IPT support three core types of data: checklists, occurrences, and events, plus data set level metadata (Smirnova et al. 2016a). A new development supported by the project EU BON is to also capture sample-based data, i.e. data that was collected by monitoring programmes according to a standardized data collection that revisits and resamples locations. These programmes are also characterized that they are comparable to surveys at other sites and that they collect, besides occurrence, also other factors such as abundance, standard methodologies and protocols (Braak and Robertson 2016). The OBIS-ENV extension to DwC extends this even further by providing a standard for exchanging environmental data collected together with occurrence data.

The tool GBIF IPT offers an easy way to submit data, for example species occurrence records and taxonomic checklists thus to publish local databases by using Darwin Core Archives (Robertson et al. 2014). It needs less technological expertise than for example DiGIR, BioCASE or TAPIR which make it especially useful for small and medium-sized projects with a less intensive IT-support. In addition to the biodiversity data itself, also metadata is provided (see Fig. 3-12 for an example of a sample-based dataset). The GBIF IPT is a free open source software that can be also used for ECOPOTENTIAL sites in order to upload biodiversity data and to provide open access.

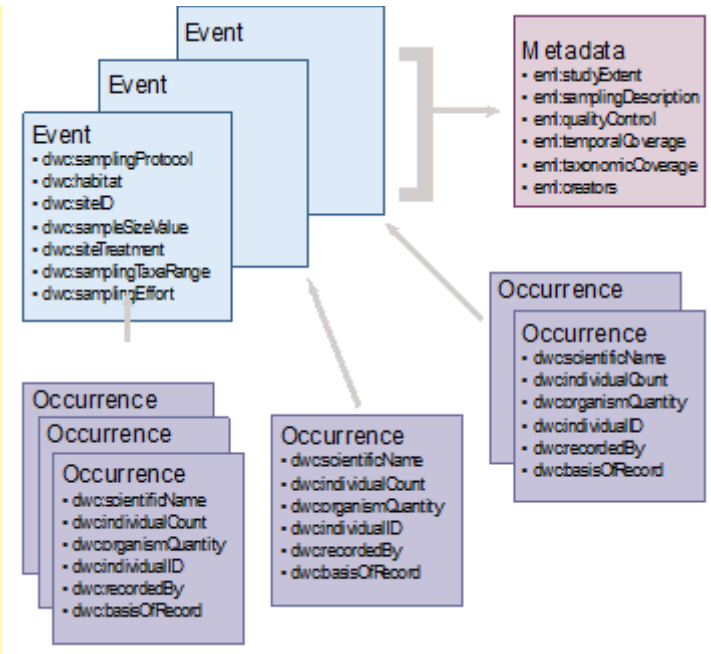

Fig. 3-12. Overview over the Darwin Core sample-based dataset (source: Braak and Robertson 2016).

There are many other options to submit data to GBIF, for example to publish data via a data curation platform (PlutoF, https://plutof.ut.ee) or by publishing data papers, for example via the Pensoft Arpha Publishing Platform (https://arpha.pensoft.net/) and its associated journals (cf. Smirnova et al. 2016b).

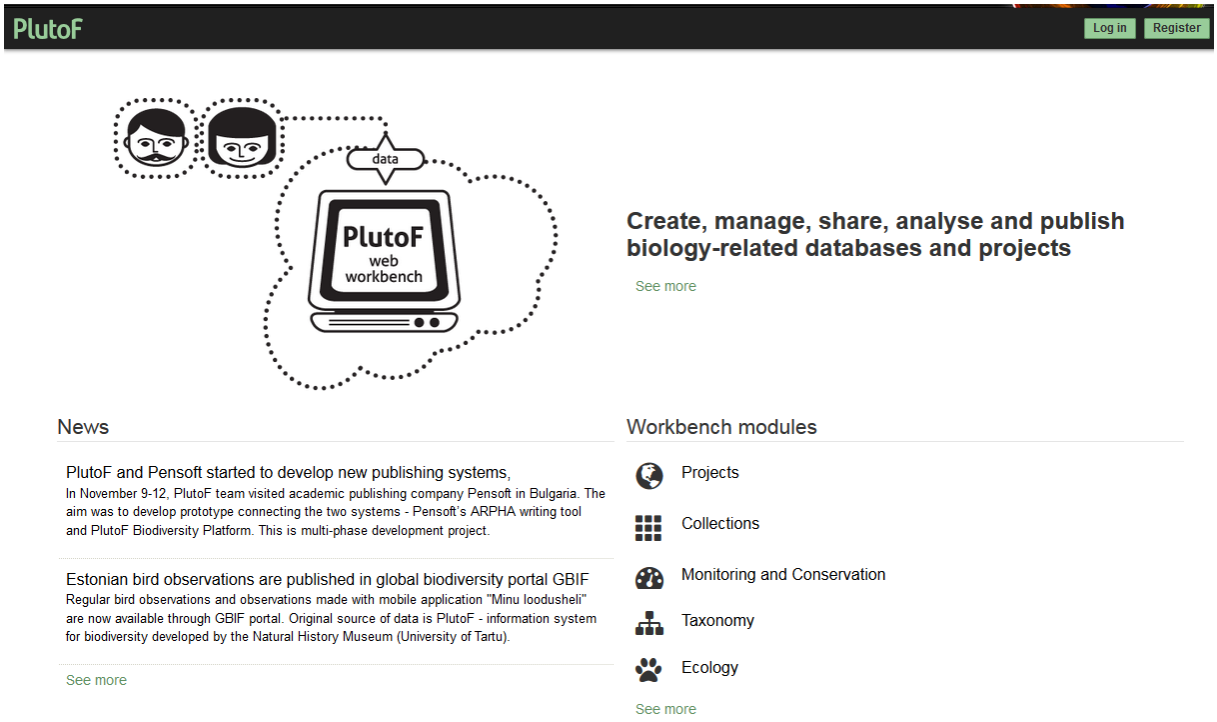

Fig. 3-13. Screenshot of the PlutoF welcome page

Pluto F (Fig. 3-13) is a cloud database for biological data that offers the opportunity to upload and curate data. All uploaded data is automatically linked (occurrence records, information on the collector, sound files, pictures etc.) and can be searched, analysed and published as a dataset (Köljalg 2016). There is also a feature implemented that allows the publication to GBIF. Targeted users are monitoring programmes, biological researchers, and citizen scientists. Date can be curated and uploaded via a Web-Interface or by using mobile phone applications that were specifically developed for PlutoF, such as I saw A Butterfly for sporadic butterfly observations or My Naturesound for animal sound recordings.

Another possibility to publish data to GBIF is to use ARPHA BioDiv, the Toolbox for Scholarly Publishing and Dissemination of Biodiversity Data that allows to publish data as data papers as well as submitting the data in turn to GBIF (Penev 2016).

3.2.7 EuMon

EuMon (EU-wide monitoring methods and systems of surveillance for species and habitats of Community interest) is a platform that offers metadata information on monitoring schemes and monitoring organizations across the countries of the European Union (http://eumon.ckff.si/monitoring/). The database lists currently (06/2017) 413 programmes: 470 monitoring schemes on species and 177 on habitats. The most monitored species groups are birds (175), followed by mammals (111), invertebrates (90) and plants (37, see also

http://eumon.ckff.si/biomat/1.1.2.php). Furthermore, over 4000 contact details of monitoring organizations were gathered. More than 4000 species are listed in the monitoring programmes. Most schemes have a local scope (55.4%) and most frequent targets are forests (28%), marine habitats (16%) and grasslands (14%, Lengyel et al. 2008).

Fig. 3-14. The EuMon database on European habitat and species monitoring (available under http://eumon.ckff.si/monitoring/)

The platform is also used to derive common and standardized protocols. The meta-database was developed in a EU-funded project that took place between 2004-2008 and metadata on further schemes were collected in turn, for example during the FP7 EU BON project period (2012-2017). The database contains metadata information on the coverage and characteristics of biodiversity monitoring in a comprehensive database (DaEuMon) and on participating volunteers. Further information is available on the geographic scope, the target country, the species included and on the coordinator of the programme. A specific search interface allows to search for certain species, species groups and habitats/habitat groups as well as on the geographical scope (International, EU, national, regional or local).

3.2.8 SeaDataNet

SeaDataNet is a standardized system for managing large and diverse data sets collected by the oceanographic fleets and the automatic observation systems.

All SeaDataNet metadata services (CSR, CDI, EDMED, EDMERP and EDIOS) make use of XML formats and exchange schema's (XSD). These are based upon the ISO 19115 content model. The CSR and CDI services have been upgraded to ISO 19139 in order to make these INSPIRE compliant and thus make the data

exchange more efficient by using automatic harvesting.

The Cruise Summary Reports (CSR) are the usual means in SeaDataNet for reporting on cruises or field experiments at sea. The Common Data Index (CDI) metadata format gives a highly detailed insight into the availability and geographical spreading of marine data sets. EDMED is a comprehensive reference to the marine data sets and collections held within European research laboratories, so as to provide marine scientists, engineers and policy makers with a simple mechanism for their identification. It covers a wide range of disciplines including marine meteorology; physical, chemical and biological oceanography; sedimentology; marine biology and fisheries; environmental quality; coastal and estuarine studies; marine geology and geophysics; etc. Data sets are described in EDMED irrespective of their format (e.g. digital databases or files, analogue records, paper charts, hard-copy tabulations, photographs and videos, geological samples, biological specimens etc). The European Directory of Marine Environmental Research Projects (EDMERP) covers marine research projects. The European Directory of Marine Environmental Research Projects (EDMERP) covers marine research projects. The European Directory of the Ocean Observing Systems (EDIOS) is an information system for marine observing and monitoring programmes, stations and platforms (including moored buoys, coastal installations, seabed stations, drifting buoys, repeated sections and sampling stations, airborne repeated tracks, etc) where there are routine, repeated, and consistent long-term observations of the marine environmental conditions, and where the data are made available for use in real-time, or near real-time. This directory includes discovery information on location, measured parameters, data availability, responsible institutes and links to data-holding agencies.

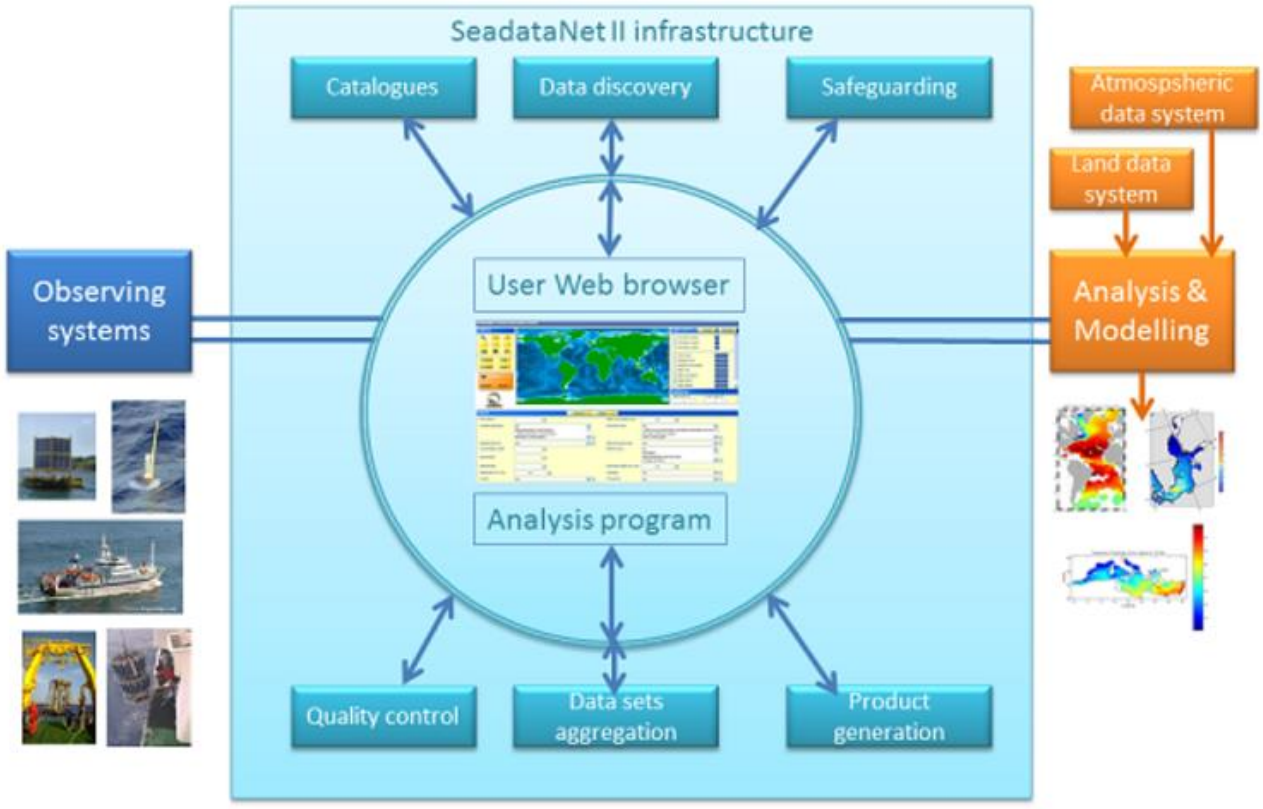

Fig. 3-15. The SeaDataNet infrastructure diagram (available at https://www.seadatanet.org)

4 Identification of community metadata needs

In order to assess the needs, use, and knowledge on metadata standards and metadata elements we developed online surveys and distributed them among the ECOPOTENTIAL community. In addition, the existing community profiles implemented in DEIMS-SDR for datasets and research sites were evaluated and combined with the results of the metadata surveys.

4.1 Questionnaire on metadata schemes and needs

Whilst the discovery and retrieval of the in-situ data for a certain geographical area is a challenging task, more difficult is the association of such data with their relevant metadata. Within the Task 5.3 we tried to identify the awareness of the ECOPOTENTIAL participants on the metadata use, the schemes that are used and the needs that exist and are open for improvements and implementation. We select the use of an online semi closed questionnaire, implemented using Google forms as platform with online access to the participants.

We developed six small and simple questions, which could easily be understood by the broad and diverse community in ECOPOTENTIAL (see 2.2). This encompassed the following questions:

In which WP you are involved? The purpose of this question is to identify the group that the participant is involved and how the in-situ data can be used by them.

Do you need metadata for your work? The purpose of this question is to understand the usefulness of the metadata during the use of the in-situ data. Moreover, to understand the awareness the user has and the significance of the metadata from the user's perspective.

If YES, which metadata are needed? The purpose of this question is to identify the critical metadata that are important to accompany the data. Six fixed answers and one open have been selected based on the experience of the in-situ data WP5 gained during the T5.1 implementation. The possible answers are:

- Time span for the collection of data,
- Methods applied of data creation and collection,
- Information on who collected the data,
- Information on the instrumentation used in the data collection,
- Observation design and frequency of the data collection,
- Data policy and terms of data use,
- Information on the quality for the data, and
- Any other information, which could be specified as free text.

Does the absence of metadata affect your work? The purpose of this question is to identify the impact of the partially lack of metadata on the work of Earth Observation Data analysis and in the Ecological/Environmental modelling.

Are in-situ data with no metadata useful for your work? The purpose of this question is to identify the impact of the lack of metadata from the in-situ data on the work of Earth Observation Data analysis and in the Ecological/Environmental modelling.

Which metadata protocol is easier to use? The purpose of this question is to identify the available protocols

for metadata creation that are easy to be prepared by using online tools, shared along with the data and understood by the end user of any level of awareness with metadata. The possible answers are:

- ISO19115/19139,
- EML (Ecological Metadata Language),
- DarwinCore,
- DublinCore,
- Any other domain specific metadata standard,
- Metadata according the INSPIRE metadata specification, and
- Other relevant metadata standards, provided as free text.

The questionnaire was open for contribution by all participants of ECOPOTENTIAL for a period of 3 months after the announcement in the ECOPOTENTIAL Basecamp, the communication platform of the project.

4.2 Results on metadata schemes and needs

We collected 22 responses from persons that are involved in almost all work packages; special focus has been for the participation of scientists from the work packages 4, 5 and 6 as well as 9 (see Fig. 4-1). The answers give as an insight on the awareness on the importance of metadata that the community has as well as the importance of them during the work. Below, we state the figures depicting the results of the online questionnaires.

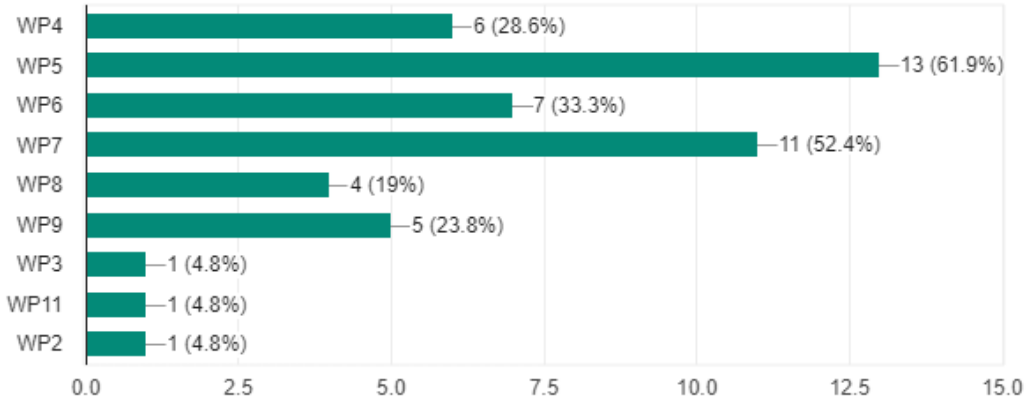

Fig. 4-1. Question 1 - Results on the involvement in the different WP (multiple answers possible); total = 22 answers.

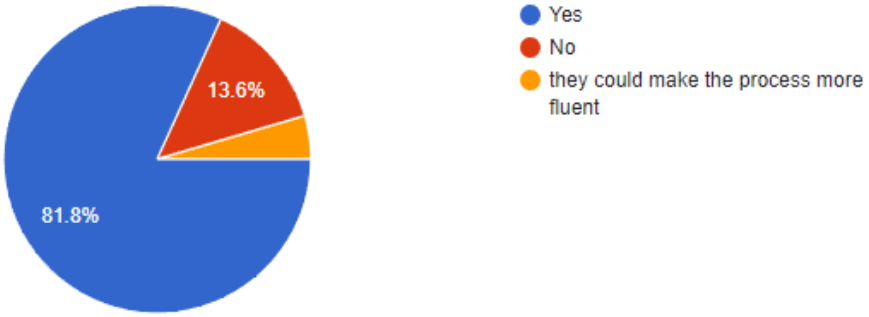

Fig. 4-2. Question 2 - Results on the importance of metadata; total = 22 answers.

The results of the questionnaire showed a high awareness on the importance of metadata for the scientific work (see Fig. 4-2). 86,4% (this are 19 answers) of the replies stated a high importance of the metadata for understanding the data and making the work with the data more fluent.

Providing the information on the time span of the data collection, the methods applied for the data collection (e.g. methods, observation design, or instrumentation) as well as information on how the data can be used were ranked as the most important metadata elements (see Fig. 4-3). The results comply with the basic elements of metadata describing the "When, How, Who and Where" of the data collections. This information can be seen as important for the discovery and reusability of the data-by-data analysts.

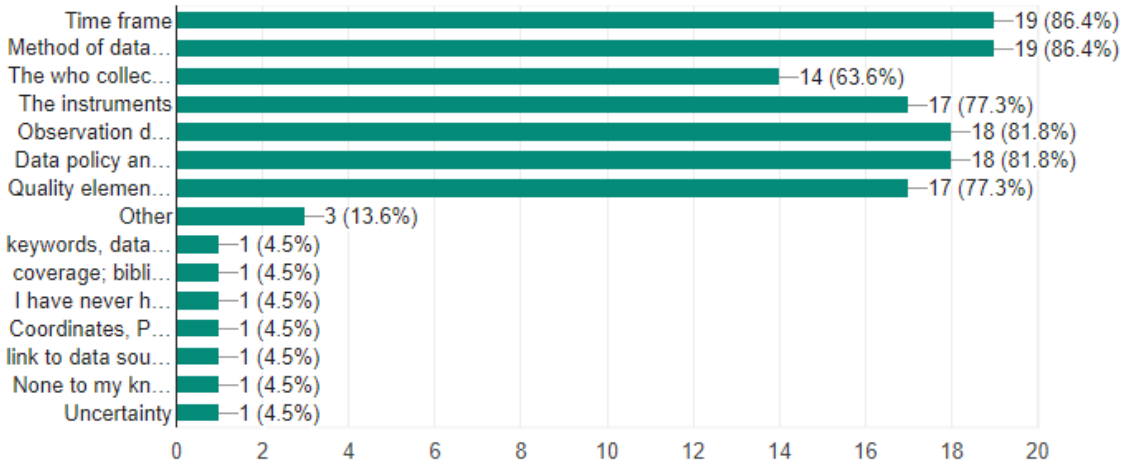

Fig. 4-3. Question 3 - Results on relevant metadata elements; total = 22 answers.

Additional metadata elements mentioned (under other) are keywords describing the content of the data, the coordinates as well as uncertainties and thematic coverage of the dataset. The results are in line with the experiences from the discussion with the LTER community (see Kliment & Oggioni 2011) leading to the first version of the DEIMS-SDR Dataset Metadata Model.

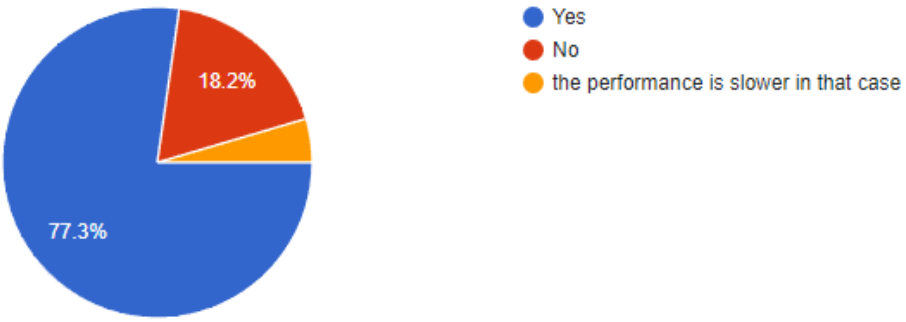

Fig. 4-4. Question 4 - Results on the effects of missing metadata; total = 22 answers.

Question 3 (see Fig. 4-4) can be seen somehow as complementary to the importance of metadata. We addressed whether a user assumes that the absence of sufficient metadata hampers the daily work with the data. About 82 % (= 18 replies) of the persons answering the survey stated that the absence of metadata has

significant impact on the workflows applied. Still 18,2% of the answers stated that there is no effect. This can be seen due to the fact, when the distance between data provider and data users, either in geographic or personal terms, is short and a direct communication and exchange of the data is done in a '*person to person*' manner, where metadata are exchanged directly and not in standardised forms. Nevertheless, the great majority of the users see the importance of metadata that also can be because relevant data are increasingly shared via data portals and personal communication on the data itself might be restricted.

Question 5 specifically addressed the importance of metadata for in-situ data (see Fig. 4-5). Interestingly in contrast to the previous question, only one third of the repliers stated that in-situ datasets are of no use. About two third stated with 'yes' or 'yes, but'. This can be interpreted, that metadata on in-situ datasets are not only provided by standardised metadata schemata but also via other channels, e.g. publications or personal communications. But, also the importance of metadata for the discovery of the data in a network of information was listed as aspect for the provision of metadata.

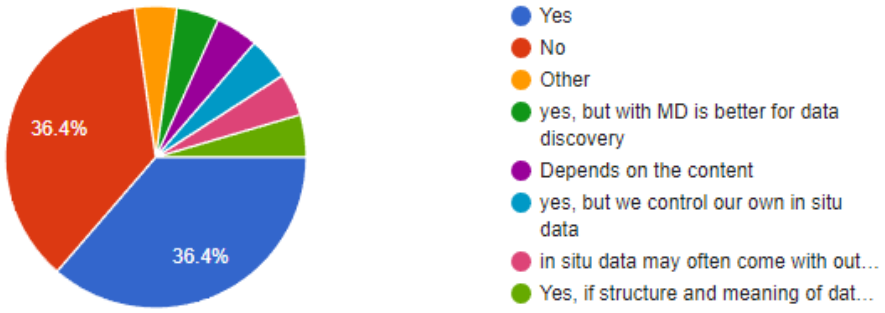

Fig. 4-5. Question 5 - Results on the importance of metadata on in-situ dataset; total = 22 answers.

With regard to important and implemented metadata standards in the Earth Observation and Ecological community in ECOPOTENTIAL ISO19115/19139, INSPIRE Metadata Specification and EML were listed as the most important metadata standards (see Fig. 4-6).

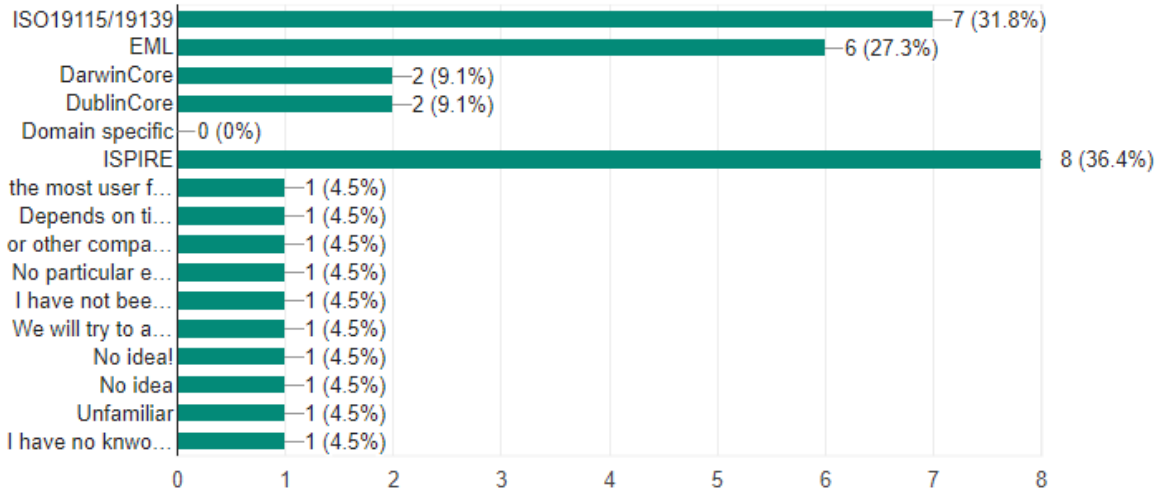

Fig. 4-6. Question 6 - Results on relevant metadata schemas (multiple answers possible) total = 22 answers.

In addition, Darwin Core (DwC) for biodiversity datasets (mainly species occurrence records) and Dublin Core (DC) for any other information was listed by two experts.

The results of the questionnaire on the importance of metadata and the relevant metadata elements was used to evaluate the current version of the community dataset metadata profile in DEIMS-SDR (see Kliment & Oggioni 2011). This led to the confirmation of the existing elements and if needed to an extension of the community profile for ECOPOTENTIAL (see chapter 5.3). In addition, the need for the development of the community profile for the data products, as intermediary metadata record describing the observation program, could be formulated (see chapter 5.2).

One of the results of the questionnaire was the importance of data quality elements as part of the metadata documenting the datasets (see Fig. 4-3). This led to the development of a second questionnaire on defining elements for the documentation of data quality in geo-spatial as well as in in-situ data.

4.3 Method applied to evaluate relevant quality elements.

The importance and challenges of habitability and land sustainability imply environmental observation and monitoring (GEOSS, COPERNICUS, Earth Systems Data Cube) as well as, environmental/ecological modelling (ME) relating to biodiversity management and ecosystem services assessment (BISE, GBIF, LIFEWATCH, IPBES). These purposes require the increase of (spatial) data capture, modelling, management, sharing and access (Big, Open and Linked Data) included in thematic applications, systems and (cyber) geographic information infrastructures development framework. The recent quantity and diversity of spatial data are associated with data life cycle changes and demands enhancement on data quality assessment and management processes, knowledge management, reinforce training opportunities/education and individual/organizations capacity-building needs, as well as improving communication and technical-political decision processes. In the last decades, data quality assessment and management concerns influenced data modelling, datasets/database management, but mainly metadata profiles definition and metadata catalogue fulfilment (European Commission INSPIRE Geoportal). In this period, concepts, benchmarks, standards and practices have evolved, from: (i) the internal/intrinsic assessment of the data producer's/provider compliance with the previous and specified data model or production processes; to (ii) external quality when evaluating the satisfaction/adequacy of the data to the request/purposes defined by each user for specific uses highlighting the usability or "*fitness for use*". The scientific and technological community of (spatial) data quality assessment presents new conceptual approaches and methodological frameworks, methods and instruments/tools in the direct (total or partial) spatial data evaluation (ISO19157: 2013 - Data quality - DQ) namely, in the generation of indicators and representation of the spatial variability of quality elements. In recent years, efforts have been made in the meta-evaluation of external and (in) direct quality by the end-user(s), taking advantage of possibilities of documentation and communication of quality elements in metadata profiles.

This requires simple and (semi)automatic metadata fulfilment, integration and metadata catalogue interoperability, in parallel, to the operationalization of the theoretical methods foreseen in the spatial data quality assessment and management standards (ISO19157 And ISO19158: 2012 - Quality assurance of data supply). In the last ten years, several national and European projects have been able to accumulate metadata profiles and catalogues, methods, and new tools for evaluating spatial data quality. The collaborative networks management challenges reveals relevant scientific needs and associated opportunities in the metadata (semi)automatic generation and improvement, expediting tools for requesting and evaluating the

external quality of spatial data (GEOSS), as well as advances in the processes of control, assurance and quality management. The internal and external data quality assessment and management guarantees an adjustment to the reality, objectivity and capacity of understanding, communication and decision between processes and actors involved. In the last decades there have been conceptual and instrumental advances in structure, format and modelling (ISO 19101) and specification of spatial data products (ISO 19131: 2007 - Geographic information - Data product specifications) as well as definition of quality evaluation principles (ISO 19113 ISO 19113:2002 - Geographic information -- Quality principles, ISO 19114: 2003 Geographic information - Quality evaluation procedures, ISO / TS 19138: 2006 - Geographic information - Data quality measures). Increased data and exponential number of users changed the data life cycle and the normal division between data producers/providers (internal quality perspective of compliance with a technical standard or data product specification) and data users (users quality perspective for a given application, specific utility for the satisfaction of a user data).

The need to standardize languages, define approaches and establish concepts and procedures is established by ISO 19157 in its relation to the control and quality assurance of data supply (ISO 19158; Geographic information - Quality assurance of data supply). The development of metadata profiles and metadata catalogues according global (ISO 19115 and ISO 19139), international (INSPIRE) or thematic (DEIMS) standards improve (in) direct quality assessment processes and promotes data quality practices and routines. The external/user perspective/fitness for use quality motivates the development and implementation of user oriented quality evaluation routines for a rapid/agile and adequate assessment of the internal and external quality of pre-existing data based on (spatial) metadata (Task 5.3) quality evaluation, following the principles of ISO 19157 and ISO 19158. The identification, analysis and specification of requirements for the user-oriented quality evaluation routines, supported on metadata implies the: (1) analysis of several candidate metadata profiles (INSPIRE; DEIMS; EML; ISO19157); (2) evaluation of the current DEIMS-SDR MD model in terms of adequacy for data quality assessment; and (3) proposal of new fields for the selected metadata profile, related to data quality elements (in cooperation with Task 5.3). The activity it´s focused on the user's input data quality requirements according to the application context (spatial modelling) and to the users' expertise (or "expected quality"). Therefore, was implemented an online questionnaire (also in cooperation with Task 5.5) on "knowledge and routines of data quality assessment and management" devised to inquire internal ECOPOTENTIAL researchers/data users.

This questionnaire was prepared and implemented on Google Forms (collaborative and oriented online questionnaire with closed and open questions) and it's targeted at both data providers and data users (including EO data production; field data production; data use in correlative or process-based modelling). More specifically the questionnaire intends inquiry about:

(1) Knowledge of data quality, including quality theory, concepts, elements (KNOWLEDGE);
(2) Current practices (e.g. practical experience) related to (spatial) data quality assessment and management routines (PRACTICAL EXPERIENCE); and
(3) Awareness, interests and willingness to implement data quality routines (UTILITY) of ECOP data providers and data users.

These activities benefits from the experience and tools from previous projects (e.g. BIO_SOS, EU BON), using open source software, in which fitness-for-use evaluation processes/tools were developed and tested with a strong input from data users. This will support the quality-driven selection of relevant data identified as well as the identification of data quality gaps and the planning of targeted data collection (see WP5.1). The results (*52 responses from 18 countries) indicates researchers with a multidisciplinary technical scientific background mainly with* Biology/Ecology and Remote Sensing/Spatial data as well as Engineering/Technology and

Geophysics/Geography, with *(Pos)PhD All WP, involved mainly in WP4, WP5 e WP6.*

5 ECOPOTENTIAL Community Profiles

Based on the results of the surveys, the evaluation of work done in previous projects with respect to the definition of community profile for different data objects (e.g. datasets, observations), and the analysis of metadata standards community profiles for ECOPOTENTIAL were suggested. Main parts of the work are based on previous work done by the LTER community. It resulted in the adoption of the site documentation for protected areas as well as the evaluation and adaptation of the LTER community profile for dataset. This metadata community profile was extended by elements describing the data quality.

This work was done based on user feedback from the different questionnaires (see chapter 4). The current report provides and overview on the metadata elements discussed. The current chapter provides an overview on the implemented metadata models for

- research sites and protected areas
- data products
- datasets

Within the DEIMS Site and Dataset Catalogue the ECOPOTENTIAL community profiles for the different topics were implemented. Details are described in chapter on the Technical implementation (chapter 6).

5.1 Observation facility [SITE]

The basic concept in DEIMS-SDR is the '**site**'. This is defined as the location where short or long term observations and/or experimentations take place. A site clusters a series of different observations plots together, which are managed by one or more communities. A site is either the area of investigation (e.g. a catchment) or an area of interest (e.g. a socio-ecological defined region). In this respect, a **protected area** can be defined as 'site' and can by this be described using the MD model SITE in DEIMS. For ECOPOTENTIAL a minimum set of information needs to be provided, but more information is welcome. For all protected areas addressed in ECOPOTENTIAL the basic information was available in the site catalogue and updated with expert knowledge from the local site managers.

5.1.1 Definition

A '***Site***' is defined as basic generic in-situ observation or experimentation facility (of any size and organisation level), where the collection of e.g. biophysical, environmental, biological or socio-ecological data is conducted (2016). By following this definition, a protected area can be defined as a monitoring site to generate data for scientific or management purposes (including experiments, e.g. exclosures). In many case in Europe long-term observation sites are located within protected areas (e.g. Gran Paradiso National Park, Italy).

5.1.2 Recommended MD elements for observation facilities

Based on these requirements the metadata model as defined for ILTER (DEIMS 2014) was reviewed. Fig. 5-1

provides an overview on the terms used to document an observation location, as e.g. a protected area. A full description of the SITE MD model is provided in the documentation section of DEIMS[24]. There also the list of mandatory fields is defined resulting from discussions within the long term ecological research community. The basic metadata model was adopted for ECOPOTENTIAL and adapted to its needs. A version 1.1 was created providing options for a generic documentation of any observation facility. The full Site Metadata Model (SMM) is documented on the DEIMS-SDR Data Model Documentation page[25]. In the following (fig 21) some core elements of the site documentation are provided.

Fig. 5-1. Terms to document an observation or experimentation facility

A. NAME AND GENERAL DESCRIPTION

Site name (DEIMS-SI: 1.1 Site name)

Definition: Name identifying the documented observation and/or experimentation facility (site, e.g. LTER Site or LTSER Platform).

[24] See https://data.lter-europe.net/deims/documentation/site

[25] see https://data.lter-europe.net/deims/documentation/site

Recommendation: The site name should be concise and precisely, e.g. by a given named place. It should not contain unexplained acronyms or abbreviations. It is recommended a maximum length of 200 characters and keeping the similarity with the 'official naming' established in the community.

Example *Kalkalpen National park – Austria*

Site identification

DEIMS-SI: 2.1 Site UUID

Definition Unique alphanumeric identifier of the site. The UUID is automatically generated by DEIMS.

Recommendation & Hints The UUID is generated automatically by the system. The UUID is used for creating the URL for the site, e.g. https://data.lter-europe.net/deims/site/8eda49e9-1f4e-4f3e-b58e-e0bb25dc32a6.

The UUID is taken as network independent unique identifier for the observation and experimentation facility. In addition the SITE CODE as network specific identifier can be added.

Example *49515dda-1198-4013-8f43-c33e107af081*

DEIMS-SI: 3.1 Site code

Definition: Unique alphanumeric identifier for the site (e.g. LTER Site or LTSER Platform) within the DEIMS-SDR ecosystem site registry. The site code is composed of a network identification, a country identification and a sequential number following the template [Network][Country]_[Sequential Number]. The site code is defined by the responsible network coordinators (e.g. LTER member network representative) and should be provided to the site manager.

Recommendation: The *Site code* as unique alphanumeric identifier is composed of a network identification, a country identification and a sequential number. For LTER-Europe, the template is: LTER_EU_XX_YY(Y); For East Asia-Pacific: LTER_AP_XX_YY(Y); for North America: LTER_NAM_XX_YY(Y); for South America: LTER_SAM_XX_YY(Y); for Southern Africa: LTER_SAF_XX_YY(Y); where XX is the 2-letter ISO country code and YY(Y) represents a sequential number. Use (Y) for a subsite. Please make sure that the number is not used for any site in your country. For example: LTER_EU_AT_001. For other networks than LTER a different prefix for the network identification is used. If you are part of the Critical Zones or Soiltrek Network, then the format for your code is: CZO_EU_XX_YY(Y); where XX is the 2-letter country code and YY(Y) represents a number. For example: CZO_FR_01 or RBV_FR_100.

Example *LTER_EU_AT_008*

Site area (DEIMS-SI: 4.1 Size)

Definition: Size of the site in hectare (ha). The size refers to the the boundary of the catchment, the observed area or the relevant area for extrapolation dependent on the observed ecosystem type (e.g. forest, lake, rivers).

Recommendation: Please provide the size of the site in hectares. 1 hectare = 1 square kilometer /100 OR

hectares = square kilometers X 100. The estimation of the size refers to the observed area of the site. This could be either the observed catchment area, the observed area defined by the bounding boundary of the observation devices, or the relevant area for extrapolation depending on the observed ecosystem type (e.g. forest, lake, and glacier). In addition to the estimation of the size a separate dataset with the site boundaries should be provided as open dataset.

Example *20850 [ha]*

Site description

DEIMS-SI: 5.1 General site description

Definition: A short textual description of the site or platform which includes the location, biophysical characteristics, a brief history, the main scientific purpose at present and major plans for future, a brief overall description of infrastructure, and a brief overall description of available data/information.

Recommendation: Please provide a short site description which could also be used on the web site. This description should cover: a brief history of the LTER Site or LTSER Platform, the main scientific purpose at present and major plans for future, a brief overall description of infrastructure, and a brief overall description of available data/information. In the section ECOSYSTEM AND ENVIRONMENTAL CHARACTERISTICS additional detailed descriptions on Geology, Hydrology, Soils and Vegetation can be provided.

Example: *Kalkalpen National Park is made up of two mountain ranges * The Reichraminger Hintergebirge is one of Austria's largest distinct forest areas - a sea of forest, which has not yet been dissected by public transportation routes and human habitation. Here, you will also find one of the longest intact stream systems of the Eastern Alps. Old shelters and overgrown trails remind us today of how wood was used and harvested in earlier times. * The Sengsengebirge is a northern outpost of the Limestone Alps. The ca. 20 km long main ridge reaches its highest point at the Hoher Nock (1,963 m). The name Sengsengebirge can be traced back to the use of the forests as a source of energy for the numerous scythe smithies once located here. Facts & Figures Established: July 25, 1997 Area: 20,850 ha Zoning: 89% nature zone, 11% conservation area Property: 88% federally owned (Austrian Forestry Service), 11% privately owned, and 1% municipal property Internationally recognized: as a national park (IUCN category II) since 1998, Ramsar protected area (wetland of global importance), and Natura 2000 area (European nature reserve) since 2004 Elevation: 385 to 1,963 m (Hohe Nock) Main types of rock: Wetterstein limestone, primary dolomite Landscape Classification of Kalkalpen National Park (Updated June 2011) 81% forest 8% mountain pine 6% alpine pastures and meadows 5% rock and scree ...*

DEIMS-SI: 5.2 Purpose of the site

Definition: Description of the monitoring or research objectives of the site, e.g. including the legal framework (e.g. UN-ECE CLTRAP, national monitoring programme) for the observation or experimental site. This includes an explanation why the observation or experimentation facility was established.

Recommendation: Please provide a short description that summarises the purpose of the site, as well as the main research or monitoring objectives.

Example: *National parks provide enduring protection to unique natural landscapes for the benefit of future*

generations According to the definition by the International Union for the Protection of Nature and Natural Objects, national parks are natural areas on water or land, which are designated to protect the integrity of one or several ecosystems and to preserve them for current and future generations. They are intended to prevent exploitation and other activities that may cause damage to the area. They are also meant to provide a basis for spirituality, research, education, recreation, and sightseeing that is environmentally and culturally compatible.

Site preview (DEIMS-SI: 6.1 Images)

Definition: Image showing a characteristic situation at the site. This can include instrumentation, observation plots, etc.

Recommendation: Please provide an image showing a characteristic situation at the site. Please select the file from a directory and press the 'UPLOAD' button to upload the image to DEIMS. The formats can be png, jpg, jpeg or gif.

Site organisation

DEIMS-SI7.1 Parent site name

Definition: Reference to the parent LTSER Platform in which the site is located. The relation could either be a spatial or thematic.

Recommendation: Please provide the reference to the LTSER Platform (by Site name) in which the site is located.

Example *LTSER Platform Eisenwurzen (EW)*

DEIMS-SI7.2 Sub site name(s)

Definition: Reference to the child elements of a LTSER Platform. The relation could either be a spatial or thematic.

Recommendation: Please provide the reference to the sites (by Site name) which belong to the LTSER Platform.

Example

Keyword set (DEIMS-SI: 8.1 EnvThes Keywords)

Definition: Provides a set of related keywords characterising the site derived from the controlled vocabulary implemented by EnvThes - thesaurus for long term ecological research, monitoring, experiments EnvThes and other environmentally related thesauri as Catalogue of Life, EUNIS Habitats and INSPIRE Spatial Data Themes. Additional concepts can be defined as free keywords.

Recommendation: The keywords from the following groups can be selected from a) Site classification and infrastructure, b) Observation devices, c) Research topics, d) Parameter, e) Method, f) Events, g) EUNIS habitats, h) Observed objects, i) System classification, and j) Scale. The field provides an autocomplete

function, which allows to select the values from the list. Begin typing a keyword. Select the one that most resembles the concept you want. You may add as many keywords as you need. The current version of EnvThes is cached in DEIMS. Therefore minor differences between the current version on the web server and in DEIMS-SDR can occur. Updates of EnvThes in DEIMS-SDR are done automatically including updates of concepts.

Example Permanent oligotrophic lakes, ponds and pools, microclimate, maps, ecosystem ecology, phenological stage, rainfall chemical analysis, organism classification, water properties, ecosystem processes, protected area, ECOPOTENTIAL.

B. SITE CONTACT DETAILS

Site manager (DEIMS-SI: 9.1 Contact: Site Manager)

Definition: Reference to the contact person responsible for the site. The person could be either the principal investigator (PI, co-PI), the information manager (IM), the site manager (SM) or a technician (TE). The details for the person/organisation needs to be specified in the persons module (PEMM).

Recommendation: Please provide the reference (Person) to the person responsible for the site. Please start typing the name of the person. If the person is not in the list, please add first the person/organisation in the Person MD Edit Form.

Example *Pöpperl, Franziska*

Site owner (DEIMS-SI: 10.1 Contact: Site Owner)

Definition: Reference to the contact person/organisation owning the site. The details for the person/organization needs to be specified in the person's module (PEMM).

Recommendation: Please provide the reference (Person) to the person responsible for the site. Please start typing the name of the person. If the person is not in the list, please add first the person/organisation in the Person MD Edit Form.

Example *Mayrhofer, Erich*

Funding organisation (DEIMS-SI: 11.1 Contact: Funding Agency)

Definition: Reference to the contact person/organisation or the responsible national authority for funding and strategic decisions about the site. This could be e.g. the Ministry of Science as responsible national authority. The details for the person/organisation needs to be specified in the person's module (PEMM).

Recommendation: Please provide the reference to the person/organisation responsible for the site. Please start typing the name of the person. If the person is not in the list, please add first the person/organisation in the Person MD Edit Form.

Example *Nationalpark OÖ Kalkalpen GesmbH*

Online information (DEIMS-SI: 12.1 Web address title and DEIMS-SI:12.2 Web address URL)

Definition: Link to website which provides a description and further information about the site. The web address is composed of the web link (8.2) and a title (8.1) which is shown.

Recommendation: Web address title - provide a human readable title for the website.

Example *http://www.kalkalpen.at/system/web/default.aspx*

C. METADATA DETAILS

Metadata provider (DEIMS-SI: 13.1 Metadata provider)

Definition: Provides the full name of person(s), who created the documentation for the site. The details for the organisation needs to be specified in the person's module (PMM).

Recommendation: By typing the name of a person, the system provides a list of people matching the query. If the person is not in the list, please add first the person/organisation in the Person MD Edit Form.

Example *Pöpperl, Franziska*

Metadata date (DEIMS-SI: 14.1 Metadata Date)

Definition: Provides date of metadata creation or last update.

Recommendations: Is automatically generated and cannot be entered manually

Example *2016-08-22*

D. GEOGRAPHIC LOCATION

Country (DEIMS-SI: 15.1 Country)

Definition: Country in which the site is located.

Recommendation: Please provide the name of the country in which the observation or experimentation facility is located. This might be different from the LTER member network. It describes in which country the site is located, even if the member network is located in another country. Please select a country from thre reference list.

Example *Austria*

Site geographic center coordinates (DEIMS-SI: 16.1 Center coordinates)
defined as DEIMS-SI:16.1.1 Latitude and DEIMS-SI:16.1.2 Longitude

Definition This is the location of the site in the geographic space, given as a center coordinates expressed in latitude and longitude in decimal degrees (projection EPSG: 4258 ETRS89 for Europe, EPSG: 4326 WGS84 for outside Europe).

Recommendations & Hints¨Please provide the center coordinates of the site in decimal degrees. Enter the center point latitude and longitude. Supports entering data as both Decimal Degrees (-122.340932) OR Degrees-Minutes-Seconds (-123° 49' 55.2" W). Please be careful to display southern hemisphere

latitude values as as "-" and western hemisphere latitudes as "-".

Example *Latitude:* *46.00936*
Longitude: 8.56796

Site geographic boundaries (DEIMS-SI: 17.1 Site Boundaries)

Definition: Delineation of the site boundaries or the bounding box.

Recommendations: Please delineate the site boundaries on the map. The minimal requirement is the bounding box of your site.

Example *bounding box represented as POLYGON ((13.8043 48.2759, 15.4358 48.2759, 15.4358 47.2718, 13.8043 47.2718, 13.8043 48.2759)))*

Elevation (DEIMS-SI: 18.1 Elevation - Average, DEIMS-SI:18.2 Elevation range - from and DEIMS-SI:18.3 Elevation range - to)

Definition: Average and Bounding elevation range for the site (observation or experimentation facility).

Recommendations: This element consists of two numbers (minimum and maximum altitudes) and the average altitude defining height above or below the sea level. As a default vertical coordinate reference system of bounding altitudes is used Mean sea level height (EPSG: 5714). Negative values are below sea level.

Example *Elevation - Average: 956*
Elevation range - FROM: 520
Elevation range - TO: 990

Site boundary shapefile (DEIMS-SI: 19.1 Shapefile)

Definition: Shapefile representing the boundaries of the site using the geographic projections ETRS89 (EPSG: 4258) for within Europe and WGS84 (EPSG: 4326) for outside Europe.

Recommendation: Please upload the zipped shapefile with geographic projection (Europe: ETRS89 (EPSG: 4258), outside Europe WGS84 (EPSG: 4326)). If you don't have a shapefile at hand you can easily create one at: http://gis.ucla.edu/apps/click2shp/ You can also easily edit existing shapefiles at http://www.mapshaper.org/. Files must be less than 10 MB. Allowed file types: zip 7z. Please provide site boundaries and internal organisation (e.g. observation plots) as documented dataset. All datasets are linked with the site.

Example *NP_Kalkalpen_ETRS89.zip*

E. ECOSYSTEM AND ENVIRONMENTAL CHARACTERISTICS

Temperature (DEIMS-SI: 20.1 Temperature - Average annual, DEIMS-SI:20.2 Temperature range (monthly minimum) FROM and DEIMS-SI:20.3 Temperature range (monthly maximum) TO)

Definition: Temperature characteristics of the site. The temperature characteristics is composed of the mean

annual temperature, the monthly average of the coldest month (e.g. January) as the annual minimum temperature and the average of the hottest month (e.g. July) as the annual maximum temperature. The mean, maximum and minimum values should be calculated on a climatic period (e.g. 10 or 30 years) based on data from the site or a near comparable climatic station.The temperature is given in [°C].

Recommendation: This element consists of three values (annual average, minimum and maximum) of the air temperature. If possible site data should be used to calculate the mean, minimum and maximum air temperature based on a 10 or 30 years period. If no long term site data are available please use data or values from a comparable climatic station near to the site.

Example *Temperature - Average annual [°C]: 7.2*
Temperature range (monthly minimum) FROM [°C]: -12.6
Temperature range (monthly minimum) TO [°C]: 28.3

Precipitation (DEIMS-SI: 21.1 Precipitation - Average annual, DEIMS-SI:21.2 Precipitation range (monthly minimum) FROM and DEIMS-SI: 21.3 Precipitation range (monthly maximum) TO)

Definition: Precipitation characteristics of the site. The precipitation characteristics is composed of the total annual precipitation (sum), the long term average of the driest month with the lowest precipitation (based on the monthly sum of precipitation) as the annual minimum precipitation and the long term average of the wettest month with the highest precipitation (based on the monthly sum precipitation) as the annual maximum precipitation. The mean, maximum and minimum values should be calculated on a climatic period (e.g. 10 or 30 years) based on data from the site or a near comparable climatic station.The precipitation is given in [mm].

Recommendation: This element consists of three values (annual average, minimum and maximum) of the precipitation. If possible site data should be used to calculate the mean, minimum and maximum precipitation based on a 10 or 30 years data period. If no long term site data are available please use data or values from a comparable climatic station near to the site.

Example *Precipitation - Average annual [mm]: 980*
Precipitation range (monthly minimum) FROM [mm]: 25
Precipitation range (monthly minimum) TO [mm]: 350

Global biome (DEIMS-SI: 22.1 ILTER Biome and DEIMS-SI:22.2 GEO-BON Biome)

Definition: Assignment of the main habitat type based on the ILTER Reference list for biomes as well as the GEO-BON reference list on biomes. Note: the list needs to be discussed in detail also on relation to the 'main ecosystem types'.
The GEO-BON Biome are defined as the main domain of the site (observation or experimental facility), classifying the site as terrestrial, aquatic, coastal or marine. The ILTER Biomes allow a more detailed classification on a global scale.

Recommendation: Please provide the main habitat type (global classification) which is characterising the site. Please select the values from the list.

Example *ILTER Biome Mixed forest*

GEO-BON Biome Terrestrial

Global ecosystem type (DEIMS-SI: 23.1 Ecosystem and Land Use)

Definition: Assignment of the main habitat type based on list of main habitat types (or ecoregions) according to Olsen et al. 2001.

Recommendation: Please provide the main habitat type according to the list. Please select the values from the list.

Example *Mixed forest*

Biogeographic region (DEIMS-SI: 24.1 Biogeographic region)

Definition: Assignment of the site to a biogeographic region (EEA, 2002) based on the spatial location and the site characteristics. Map: Biogeographic regions of Europe (EEA 2012).

Recommendation: Please provide the biogeographic region according to the list. Please select the values from the list.

Example *Continental*

Habitat type (DEIMS-SI: 25.1 EUNIS Habitat)

Definition: Assignment of the habitat type based on the EEA, EUNIS Habitat types. Only the levels 1-3 are used.

Recommendation: Please provide the habitat type according to the list. Please select the values from the list.

Example *G. Woodland, forest and other wooded land*

Additional characteristics (DEIMS-SI: 26.1 Geology, DEIMS-SI:27.1 Hydrology, DEIMS-SI:28.1 Soil, and DEIMS-SI:29.1 Vegetation)

Definition: Textual description of the specific characteristics of the site.

Recommendation: Please provide a short textual description of the specific characteristics at the site (e.g. beech forests)

Example *The site is characterised by beech forests.*

I. FOCUS, DESIGN AND SCALE OF SITE

Research topics (DEIMS-SI: 37.1 Research topics)

Definition: Description of the research, experimentation or observation focus of the site.

Recommendation: Please indicate the research topics tackled at the site. Select the research topic from the list provided in DEIMS. The list of research topics is defined by EnvThes Research Topics.

Example *Deposition*

Observed parameters (DEIMS-SI: 38.1 Parameters)

Definition: Description of the observed parameters and parameter groups at the site. The parameter is defined as property of the ecosystem or an ecosystem compartment which can be observed either by sensors or humans, e.g. pH, species number, radiation.

Recommendation: Please indicate the parameters observed at the site. Please record every parameter for the site platform in one line. More than one entry is possible. The reference list is defined by the entries in ***EnvThes Measure*** and is frequently updated.

Example *water quality, habitat cover, biodiversity*

Observation design (DEIMS-SI: 39.1 Scale of observation and DEIMS-SI:39.2 Design of observation)

Definition: Description of the scale and design of the observations at the site. The information is composed of the scale of the observation as the spatial and thematic extent of the observation at the site, e.g. if is addressing a complete catchment or sub-catchment as well as the design of the observation as the main focus of the effort, being mainly observation or in addition experimental.

Recommendation: Scale of observation addresses the spatial and thematic scope of the observation program at the site. Please select the most appropriate value from the list. Design of observation addresses the importance of observations at the site. Please select the most appropriate value from the list. Please also check the entries in the Experimental design field.

Example *Scale of observation entire catchment*
Design of observation only observation

Experimentation design (DEIMS-SI: 40.1 Scale of experimentation and DEIMS-SI:40.2 Design of experimentation)

Definition: Description of the scale and design of the experimentation at the site. The information is composed of the scale of the experiments as the spatial and thematic extent of the observation at the site, e.g. if is addressing a complete catchment or sub-catchment as well as the design of the experiments as the main focus of the efforts, being mainly experimental or in addition observation.

Recommendation: Scale of experiments addresses the spatial and thematic scope of the experiment program at the site. Please select the most appropriate value from the list. Design of experiments addresses the importance of experiments at the site. Please select the most appropriate value from the list. Please also check the entries in the Observation design field.

Example *Scale of experimentation catchment scale*
Design of experimentation mainly experimentation

J. PROTECTION STATUS AND RESOURCE MANAGEMENT

Site protection (DEIMS-SI: 42.1 Protection program, DEIMS-SI:42.2 Protection program cover and DEIMS-SI:42.3 Protection program notes)

Definition: Description of the protection status of the site. The information is composed of the protection program (e.g. IUCN National Park), the cover of protected areas within the site and additional notes on the protection. As each site can be protected by different programs and therefore more than one values can be given.

Recommendation: Please select the protection program and an estimate of the share of the protected area within the site from the list of values. In addition textual notes can be given.

Example
Protection program IUCN Nat. Park
Protection program cover 100%
Protection program notes ...

Management of resources (DEIMS-SI: 43.1 Management of resources, DEIMS-SI:43.2 Management of resources cover and DEIMS-SI:43.3 Management of resources notes)

Definition: Description of resources (e.g. timber, game animal, etc.) are frequently and regularly used within the site. This also includes agricultural management if the site is an agricultural landscape. The information is composed of the information on the existence of management efforts (e.g. yes), the cover of area which affected by management practises within the site and additional notes on the management (42.3), e.g. as the type of management.

Recommendation: Please select if management activities are present at the site and an estimate of the share of the area within the site which is affected by the management. In addition textual notes, e.g. on the type of management can be given.

Example
43.1 Management of resources yes
43.2 Management of resources cover 100%
43.3 Management of resources notes timber harvest

In addition to the described metadata elements for the site information on the Status and history of the site, the infrastructure and operation of the site, the data management and the data sharing policies, as well as the affiliation of the site in various networks and projects.

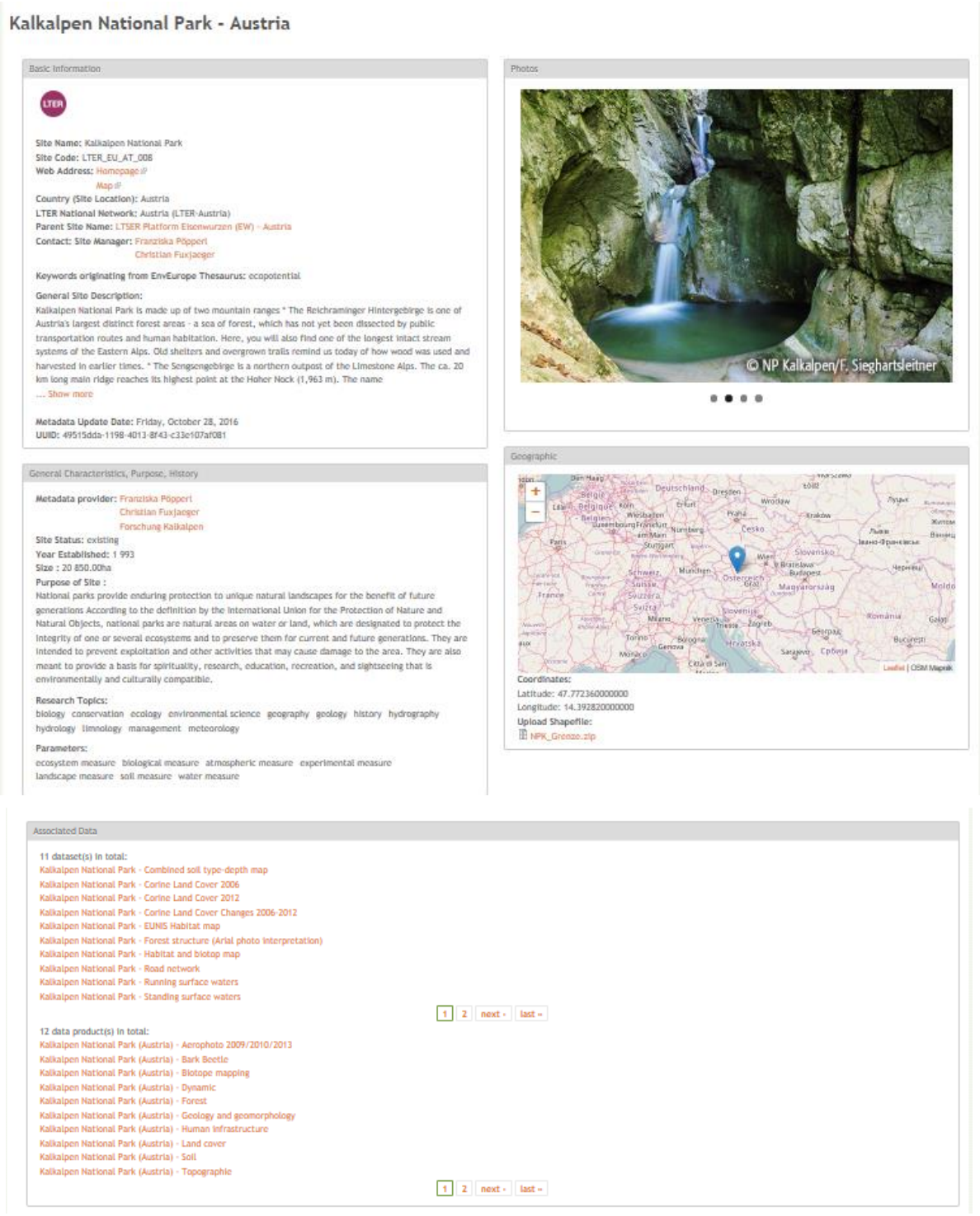

Fig. 5-2. Example of basic documentation of a Protected Area on DEIMS-SDR (see https://data.lter-europe.net/deims/site/49515dda-1198-4013-8f43-c33e107af081)

5.1.3 Mapping of site information to INSPIRE EF Schema[26]

The legal definition of INSPIRE EF is "*Location and operation of environmental monitoring facilities includes observation and measurement of emissions, of the state of environmental media and of other ecosystem parameters (biodiversity, ecological conditions of vegetation, etc.) by or on behalf of public authorities [Directive 2007/2/EC]. A generic model [for monitoring facilities], which can be used across various domains and leave the necessary freedom to thematic domains to bring in specific, needs while keeping a shared structure. So the data specification provides a common structure but not a thematic harmonisation across domains*"[27]. Environmental monitoring facilities are linked to information describing aggregations/collections of monitoring facilities and their thematic or organisational grouping and background.

In general, the aim of INSPIRE EF is to provide a standardised format to answer the following questions:

- What is being observed?
- Where is it being observed?
- When is it being observed?
- Who is carrying out the observation?

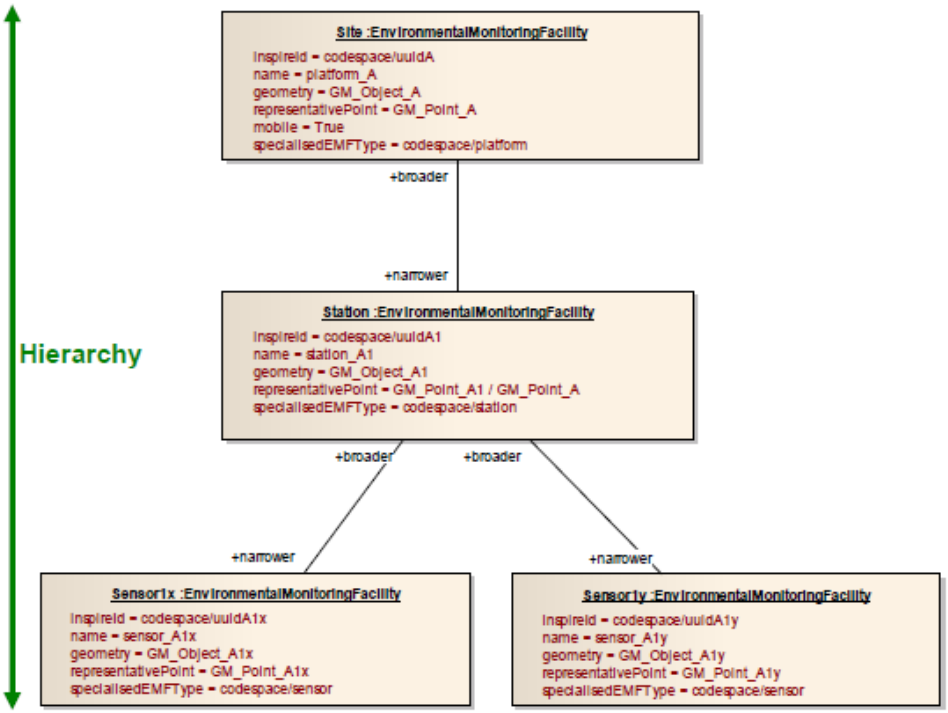

Fig. 5-3. Schema of a hierarchical cascade of environmental monitoring facilities28

The specification itself covers all kinds of environmental monitoring using fixed stations, moving equipment or remote sensing, however in the context of DEIMS-SDR it is currently limited to documenting fixed stations. INSPIRE EF has a multi-layered structure. The model provides a recursive hierarchical link between

[26] https://docs.google.com/document/d/1x87zukrpsIM42KmAO51lHUCoA-QNfNq63Fmd-nWVZfk/edit

[27] http://inspire.ec.europa.eu/id/document/tg/ef

[28] D2.8.II/III.7 INSPIRE Data Specification on Environmental Monitoring Facilities – Technical Guidelines, p.21

Environment Monitoring Facilities, thus allowing to map the hierarchy of a platform, its sites, stations and sensors.

However, these hierarchical cascades of facilities make EF information more difficult to parse and visualise compared to other metadata formats (see Fig. 5-3).

Tab. 5-1. Mapping of INSPIRE EF properties to the DEIMS-SDR MD elements

INSPIRE EF property	complete	Mandatory	multiple	DEIMS content type	Mapped DEIMS field(s)
INSPIREId		X		site	UUID
name				site	Site Name
additionalDescription				site	General Site Description
mediaMonitored		X	x	site	GEO-BON biome
legalBackground			x	dataset	Legal Act
responsibleParty	x		x	site	Site Manager, Site Owner, Founding Organisation and Metadata Provider
geometry			x	site	Coordinates and/or Site Boundaries
onlineResource				site	Web Address
purpose (has empty code list)				site	Purpose of Site
observingCapability				site	Parameters
broader				site	Parent Site Name (as link to parent site ef file)
hasObservations			x	dataset	Dataset records as links to ISO19139 files
involvedIn	x		x	data product	parameter as description, data product type as identifier, name as name, Date Range as activityTime abstract as Date Range, Owner/Creator as responsibleParty, uuid as INSPIREID, abstract as activityConditions
representativePoint				site	coordinates
measurementRegime					always “continuousDataCollection”
mobile					always “false” (boolean)
OperationalActivityPeriod	x			site	site status and Year Established (and Year Site closed)
belongsTo	x		x	site	LTER National Network or other networks

5.1.4 Possible Harvesting Mechanisms of site information

This chapter discusses possible ways to exchange site information. The preferred way to expose site information is INSPIRE EF as it is the only available standardised site metadata format. INSPIRE EF records can be exposed in various ways, however, there is no standardised approach for EF harvesting yet.

Therefore, there are a number of possible harvesting mechanisms, which include:

- An EF harvest list consisting of:
  - A unique identifier
  - The link to the EF file
  - And the update data of the EF record

  This harvest list can be used to download every single EF record available, index them locally using a script that periodically checks for any updates, and replace any updated EF record if necessary.
- Writing extensions/plugins for established metadata harvesting software, e.g. Geonetwork.
- A WFS service that stores the link to each EF record associated with a geographic entity allowing parsing these files.

Other more common metadata formats like Dublin Core, ISO 19139, ISO GMI and other metadata profiles can be harvested using CSW. Whether these formats are useful for exchanging site information is up for debate as they can only store a very limited set of information. Harvest and Transaction requests (CSW-T), which for instance can be processed by pyCSW can be used for harvesting Dublin Core, ISO 19139 and ISO GMI, but not for INSPIRE EF. The solution for site metadata harvesting strongly depends on the chosen metadata format. A more common format like ISO 19139 allows easier exchange of metadata by relying on common and easily deployable mechanism provided by Geonetwork, pySCW or others. Using a more comprehensive format that is built for storing site information, like INSPIRE EF, hinders the exchange of that information and requires newly built solutions or relying on less standardised approaches.

5.2 Data Product [DP]

The concept of '**data product**' was added in order to allow a summarised description of a collection of datasets. A 'data product' can be e.g. 'Deposition data', which contains a range of different parameters and time series. Instead of describing each single dataset or parameter the 'data product' can be described as a whole. The related datasets (by its MD) can be linked to the data product. By applying this concept, a fast overview on available data sources in a protected area can be created without providing a full description of each single dataset. This is needed in a second step in order to ensure the reusability of the data. Applying the concept of the data product, a prioritisation in the creation of metadata (if not available) and provision of datasets (e.g. in common data formats) can be made. This allows for a targeted data provision policy. If no complete dataset MD are provided via the ECOPOTENTIAL VRE, the data products forms should be used in order to provide aggregated information on the available data. The model for the data product in the current version was developed within the ECOPOTENTIAL project and was implemented in the DEIMS-SDR Site and Dataset Catalogue extending the underlying information model.

5.2.1 Definition

A ***data product*** is a product of a workflow (e.g. observations) that facilitates a series of datasets. In the specific case of ECOPOTENTIAL the 'data product' represents a summarised description of a series of data aimed at presenting the available data sources in a protected area without a full description of each single dataset. It should be seen as something between a broad class (e.g. biodiversity) and a detailed dataset (e.g. abundance of a species within 0.1 hectare plot). For instance, 'deposition data' can be a 'data product' that

contains several parameters measured with different temporal frequencies and spatial scales. The advantage of using this approach is that instead of describing each single dataset or parameter the 'data product' can be described as a whole. The related datasets can be linked to the data product through its metadata. In this way the protected areas as data holders do not have to document with detailed metadata every dataset they are collecting. In addition, the modellers or remote sensing experts can find out different types of data they can use. Once a specific datasets was identified as needed for modelling, for example, the associated metadata have to be documented by its owner.

5.2.2 Recommended MD elements for data products

The data model on data products consists of a series of terms describing the content and accessibility of the underlying data. Fig X provides an overview on the terms. A description of the terms is given in this chapter. In addition, the documentation of the terms can be found in the DEIMS-SDR Data Model Documentation section on Data Products[29].

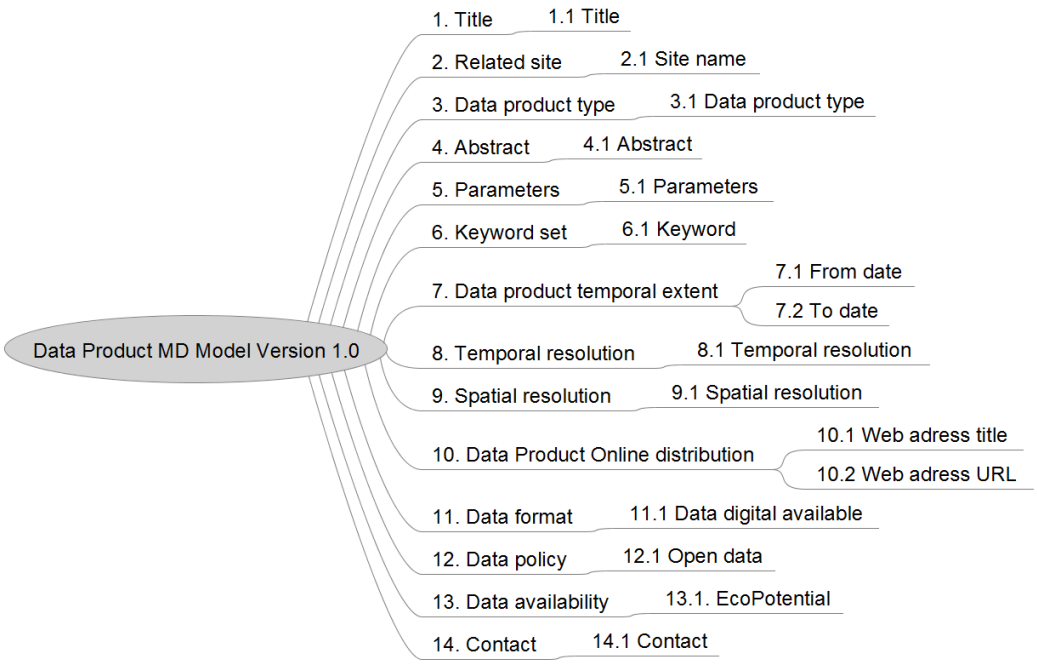

Fig. 5-4. Draft model of elements for the description of data products

A. IDENTIFICATION

Data product title (DEIMS-DP: 1.1 Title)

Definition: Provides a name of the data product that is being documented as it is known within the community described in detail by the following elements. "Title" is a characteristic and often unique and is the most informative element of a metadata record and usually with the highest priority as search engines index this element.

[29] see https://data.lter-europe.net/deims/content/Data-Product-Metadata-Model

Recommendation: The data product title has to be concise and precisely describing the point. It should not contain unexplained acronyms or abbreviations. It is recommended a maximum length of 200 characters and keeping the similarity with the original title of the dataset in the sense of the 'official naming' established in the community. If the dataset is part of a larger project, it is recommended to indicate the Project at the end of the title, in brackets. In case of Project names, abbreviations are allowed, as long as the rest of the title follows the guidelines above and the abbreviation is spelled out immediately in the abstract.
The title should follow the following template:
SITE NAME _ DATA PRODUCT TYPE _ DATE RANGE (optional)

Format	Text(255)
Multiplicity	[1], required
Example	*DP_Zöbelboden-IP3_WetDeposition_2014-2017*
Reference List	n/a

Related site (DEIMS-DP: 2.1 Site name)

Definition Reference to the site (observation or experimentation facility) where the data were gathered.

Recommendation: Please select the relevant sites from the list provided by DEIMS. The list provides all published sites in the DEIMS-SDR Site Registry. By typing the site name the system provides a list of research sites matching the query. If none available, please first document the site metadata using the editing form available and then return back to dataset metadata definition.
If the Data Product is defined as XML record, the field contains the link to the Site UUID.

Format	Reference to [Content Type] SITE
Multiplicity	[1..n], required
Example	*https://data.lter-europe.net/deims/site/8eda49e9-1f4e-4f3e-b58e-e0bb25dc32a6*
Reference List	Reference to [Content Type] SITE

B. GENERAL INFORMATION

Data product type (DEIMS-DP: 3.1 Data product type)

Definition: Reference to the thematic content of the data product.

Recommendation: Classification of the data product according to data product types specified in the taxonomy. Please select a relevant category from the list, which describes the thematic context of the data product. If new categories are needed, please use the 'Provide feedback' link to post the request.

Format	Reference to [Taxonomy] Data Product
Multiplicity	[1], required
Example	*Meteorology*
Reference List	Reference to [Taxonomy] Data Product

Meteorology
|-- Air humidity
|-- Air pressure
|-- Air temperature

|-- Radiation
|---- Global radiation
|---- Photosynthetic active radiation (PAR)
...

Air chemistry
|-- Carbon flux
…

Data product abstract (DEIMS-DP: 4.1 Abstract)

Definition: Short textual description of the data product documented as a summary of the resource.

Recommendation: Please provide a short textual description on the scope, methods, policies applied to the data product. Please summarise the most important information in the first 100 words.

Format Text
Multiplicity [1], required
Example: *"The data product contains monthly precipitation data collected at different weather stations at Zöbelboden. An automatic weather station is used to collect the data. The main variables are Air temperature at different heights and precipitation. The sampling interval is every 15 minutes aggregated to 30 minutes values. An open data policy is applied to meteorological data. The data product contains 10 different datasets from stations located in the study area. They can be accessed …"*
Reference List n/a

Observed parameters (DEIMS-DP: 5.1 Parameters)

Definition: Description of the observed parameters and parameter groups for the data product. The parameter (or measure) is defined as property of the ecosystem or a ecosystem compartment which can be observed either by sensors or humans, e.g. pH, species number, radiation.

Recommendations: Please indicate which parameters are provided by the data product. Please record every parameter. More than one entry is possible. The reference list is defined by the entries in EnvThes Measure and is frequently updated.

Format Reference [Taxonomy] Parameter
Multiplicity [0..n], optional but recommended
Example Precipitation, Air temperature, Atmospheric pressure
Reference List Reference [Taxonomy] Parameter based on EnvThes Measure
…

Biodiversity
|- Biodiversity of animals
|- Biodiversity of microbes
|- Biodiversity of plants
|- Biodiversity, overal
|- Biodiversity, genetic
Biomass
|- Biomass, aquatic
|- Biomass, terrestrial
...

Keyword set (DEIMS-DP: 6.1 EnvThes Keywords)

Definition: Provides a set of related keywords characterising the site derived from the controlled vocabulary implemented by EnvThes - thesaurus for long term ecological research, monitoring, experiments EnvThes and other environmentally related thesauri as Catalogue of Life, EUNIS Habitats and INSPIRE Spatial Data Themes.

Recommendation: The keywords from the following groups can be selected: a) Site classification and infrastructure, b) observation devices, c) research topics, d) observed parameters, e) methods, f) events, g) EUNIS habitats, h) observed objects, i) observed system compartments, and j) scale of the data.

The field provides an autocomplete function, which allows to select the values from the list. Begin typing a keyword. Select the one that most resembles the concept you want. You may add as many keywords as you need. The current version of EnvThes is cached in DEIMS. Therefore minor differences between the current version on the web server and in DEIMS-SDR can occur. Updates of EnvThes in DEIMS-SDR are done automatically including updates of concepts.

Format Reference [Taxonomy] LTER Controlled Vocabulary
Multiplicity [0..n], optional but recommended
Example *Meteorology, Mixed forest, microclimate, Mountain forest, ecosystem ecology, rainfall chemical analysis, LTER Site*
Reference List EnvThes captured in [Taxonomy] LTER Controlled Vocabulary

C. DATA RESOLUTION

Data product temporal extent (DEIMS-DP: 7.1 From date and DEIMS-DP:7.2 To date)

Definition: Defines the time period covered by the content of the data product (and the related datasets). This period may be expressed as a time (an individual date) or date ranges (interval of dates/From-To) or a mix of individual and interval dates.

Recommendation: Please provide the most appropriate range for the data collected. If the collection is still going on, please leave the to date blank.

Format	Date
Multiplicity	[1], required
Example	*01-01-1991* to *31-12-2013*
Reference List	n/a

Data product temporal resolution (DEIMS-DP: 8.1 Temporal resolution)

Definition: Description of the temporal resolution of the data. This describes the update frequency of the data, e.g. weekly or only one observation.

Recommendation: Please select the appropriate value from the list. If there are deviations or important issues to be mentioned, please describe that in addition in the abstract.

Format	Reference [Taxonomy] Temporal Resolution (Data Products)
Multiplicity	[1], required
Example	*minutes*
Reference List	Reference [Taxonomy] Temporal Resolution (Data Products)

N/A
other temporal resolution
single measurement
less than seconds
seconds
minutes
hourly
half daily
daily
weekly
bi-weekly
monthly
bi-monthly
half yearly
yearly
bi-annual
every 2-5 years
...

Data product spatial resolution (DEIMS-DP:9.1 Spatial resolution)

Definition Description of the spatial resolution of the data. This describes the spatial design for the data collection, e.g. area covering observations or grid points.

Recommendations & Hints Please select the appropriate value from the list. If there are deviations or important issues to be mentioned, please describe that in addition in the abstract.

Format Reference [Taxonomy] Spatial Resolution (Data Products)
Multiplicity [1], required
Example *single point*
Reference List Reference [Taxonomy] Spatial Resolution (Data Products)

N/A
single point
transect sampling horizontal
transect sampling vertical
regular sampling
random sampling
stratified random sampling
full area coverage
catchment
plot
no spatial design
_unknown

D. AVAILABILITY

This section describes metadata elements which provide an overview on the availability of the data. It also could include a link to an online resource locator providing either specific datasets or further information.

Data product online distribution (DEIMS-DP: 10.1 Web address title and DEIMS-DP:10.2 Web address URL)

Definition: Web address is the "navigation section" of a metadata record pointing users to the location (URL) where the datasets grouped in the data product can be retrieved directly, or provides information about how to acquire a datasets.

Recommendation: Setting up the correct resource locators is important for the connection between the data and the services that provide access to them or for providing additional information concerning the

resource. If Web address for dataset is available, the Dataset Locator shall be a valid URL providing one of the following: a link to a web page with further instructions; a link to a web service capabilities document; a link to a client application (web data portal) that directly accesses dataset. If a dataset is available offline, it may be uploaded into the system and made online available with access and use constraints and IPR defined previously.

Format	Web address title			Text(255)
	Web address URL URL			
Multiplicity	[0..1], optional			
Example	*Web address title: LTER Europe Sensor Observation Service (SOS)*			
	Web address URL:			
	http://sp7.irea.cnr.it/tomcat/envsos/sos?REQUEST=getcapabilities&service...			
Reference List	n/a			

Data product format (DEIMS-DP: 11.1 Data digitally available)

Definition: Indication if the data are available in digital format.
Recommendation: Please indicate, if the data are available in digital format.

Format	Reference
Multiplicity	[1], required
Example	*yes*
Reference List	n/a, yes, no

Data product policy (DEIMS-DP:12.1 Data openly available)

Definition: Indication if the data are available as open data.

Recommendation: Please indicate, if the data are available as open data.

Format	Reference
Multiplicity	[1], required
Example	*yes*
Reference List	n/a, yes, no

Data availability (DEIMS-DP: 13.1 Data available for ECOPOTENTIAL)

Definition: Indication if the data are freely available within the ECOPOTENTIAL H2020 project.

Recommendation: Please indicate, if the data are freely available within the ECOPOTENTIAL H2020 project.

Format	Reference
Multiplicity	[1], required
Example	*yes*
Reference List	n/a, yes, no

E. CONTACT

This section provides information on contact points. Note: the inclusion of MD Creator and MD Edit Date needs to be discussed for the next release of the DP MD model.

Dataset creator and Contact point (DEIMS-DP: 14.1 Responsible party)

Definition: Provides the full name of person(s), who created the dataset and/or who serve as contact points.

Recommendations & Hints: By typing the name of a person, the system provides a list of people matching the query. If none available, it's strongly recommended to first create a person metadata using the editing form available and then return back to dataset metadata definition.

Format: Reference [Content Type] Person
Multiplicity: [1..n], required
Example: *Kobler, Johannes*
Reference List: Reference [Content Type] Person

General notes

For each data product a MD record is created and referenced to the site/protected area. In a later phase dataset MD can also be linked to the data product.

5.3 Dataset [DS]

The dataset metadata schema describes the data files and data services provided by any data provider. Is is based on an initial selection of metadata elements for long term data established within the EnvEurope project (EnvEurope 2011). The dataset metadata model is the compromise between the efforts for metadata provision, the needs for the data discovery and the requirements for the data re-usability.

All selected metadata elements have been mapped to a corresponding metadata element in ISO19115/19139 and EML. The mapping is implemented in DEIMS-SDR for the metadata export (see chapter on Technical implementation).

5.3.1 Definition

A '**Dataset'** is a single data file or a series of data files, which are described by a metadata (MD) record. This can be either a physical file (e.g. excel, csv, shape) or a data service (e.g. OGC WFS, OGC SOS). For datasets metadata as specified in the DEIMS-SDR community profile (link to INSPIRE and EML) or the INSPIRE MD Specification need to be provided.

5.3.2 Recommended MD Elements for In-Situ Data

Currently the community profile of the Dataset MD Model defined in DEIMS-SDR is adopted in ECOPOTENTIAL. The community profile is based on metadata elements defined by the INSPIRE metadata

specification and EML. A full description of the metadata elements can be found within the DEIMS-SDR data model documentation[30].

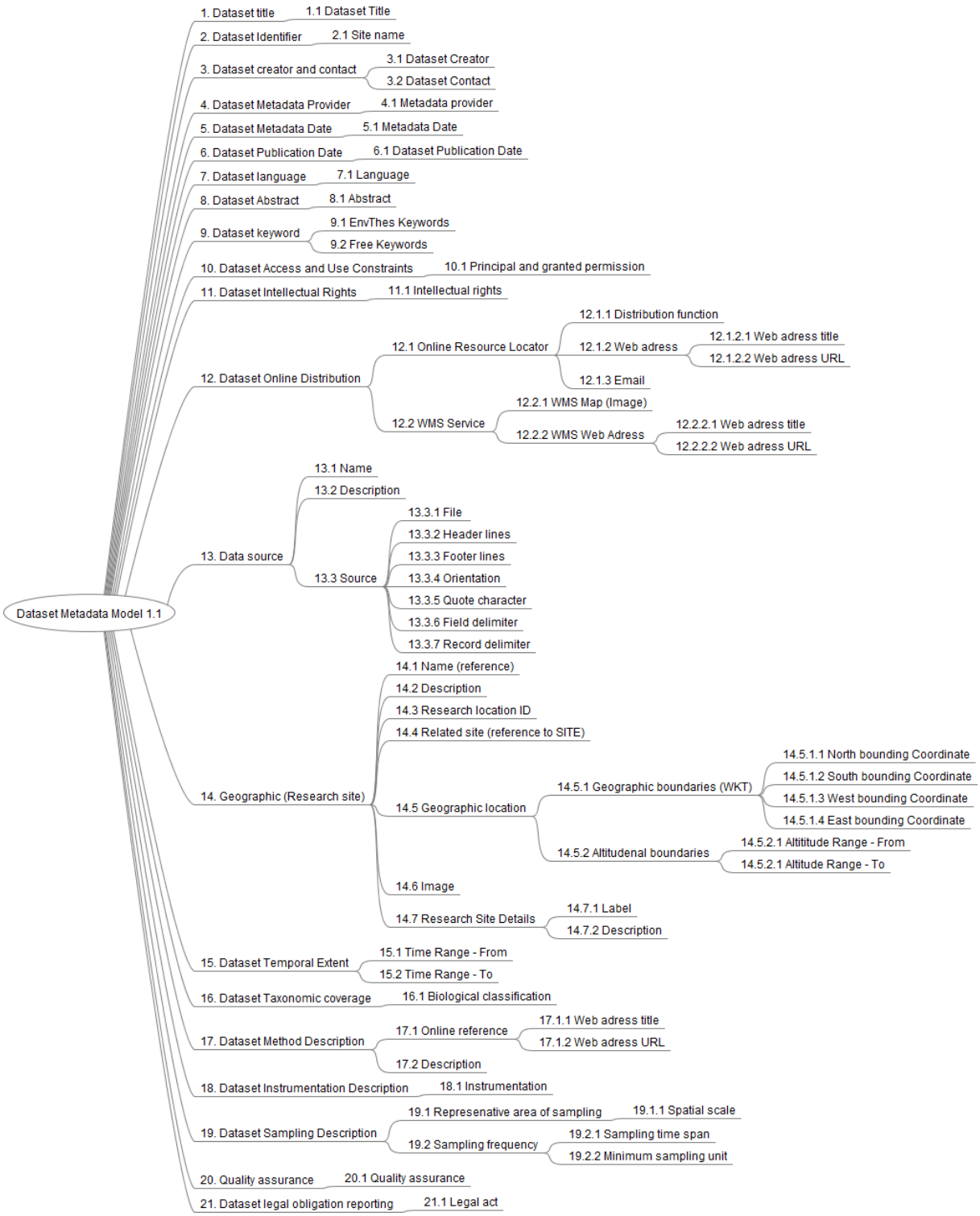

Fig. 25. Overview on the metadata elements of the Dataset Metadata Model (DSMM)

The DEIMS-SDR community profile adopted by ECOPOTENTIAL contains the following metadata fields for a

[30] See https://data.lter-europe.net/deims/documentation/dataset

dataset.

Dataset title (DEIMS-DS: 1.1 Title)

Definition: Provides a name of the dataset that is being documented as is known within the community described in detail by following elements. Title is a characteristic, and often unique and is the most informative element of a metadata record and usually with the highest priority as search engines go to this element.

Recommendation: Title has to be concise and precisely describing the point. It should not contain unexplained acronyms or abbreviations. It is recommended a maximum length of 200 characters and keeping the similarity with the original title of the dataset in the sense of the 'official naming' established in the community. If the dataset is part of a larger project, it is recommended to indicate the Project at the end of the title, in brackets. In case of Project names, abbreviations are allowed, as long as the rest of the title follows the guidelines above and the abbreviation is spelled out immediately in the abstract.

Example *Precipitation measured near Lake Santo Parmense and Lake Scuro Parmense (1951-2010)*

Dataset identifier (DEIMS-DS: 2.1 Site Name)

Definition: Value uniquely identifying a dataset. Its final form is composed of the Site Code retrieved from the Site Name and a unique identifier defined be proposed template.

Recommendation: The identifier shall contain standardized codes of the Network and the Site that are derived from the Site Name. Additionally it should identify the dataset within your working environment (network, organization, laboratory, parameters observed, time coverage, etc.) with internal codes used. By typing the site name the system provides a list of research sites matching the query. If none available, it's strongly recommended to first create a site metadata using the editing form available and then return back to dataset metadata definition.

Example Zone Atelier Seine – France

Dataset creator and contact points (DEIMS-DS: 3.1 Dataset Owner/Creator and DEIMS-DS: 3.2 Dataset contact point)

Definition: Provides the full name of person(s), who created the dataset or who serve as contact points.

Recommendation: By typing the name of a person, the system provides a list of people matching the query. If none available, it's strongly recommended to first create a person metadata using the editing form available and then return back to dataset metadata definition.

Example *Peterseil, Johannes*

Dataset metadata provider (DEIMS-DS: 4.1 Metadata provider)

Definition: Provides the full name of person(s), who created the documentation of a dataset.

Recommendation: By typing the name of a person, the system provides a list of people matching the query.

If none available, it's strongly recommended to first create a person metadata using the editing form available and then return back to dataset metadata definition.

Example *Frenzel, Mark*

Dataset metadata date (DEIMS-DS: 5.1 Date)

Definition: Provides date of metadata creation or last update.

Recommendation: Is automatically generated and cannot be entered manually.

Example *2014-12-18*

Dataset publication date (DEIMS-DS: 6.1 Date of publication)

Definition: Represents the date when the actual dataset was published on the LTER DEIMS. Any other maintenance activity is documented as a date of last revision.

Recommendation: If the dataset has not been yet published due to e.g. quality validation checks, access rights definition, further content updates expected, etc. leave this field empty. A publication date can be typed as YYYY-MM-DD of taken from the pop-up calendar.

Example *2016-08-18*

Dataset language (DEIMS-DS: 7.1 Language)

Definition: The language in which the textual parts of the dataset are written.

Recommendation: The names of parameters and their units collected within the dataset. Any other language used in textual information shall be referenced here as well. If the dataset does not contain any textual information (e.g. only codes and digits), the language should be defaulted to the value of the metadata language, which is defined as a default value - English. Start typing the language name and the system will offer you options to select.

Example *English*

Dataset abstract (DEIMS-DS: 8.1 Abstract)

Definition: Provide a brief overview of the dataset being documented.

- Recommendation: The dataset abstract is a succinct description that shall include:
- A brief summary with the most important details that summarise the data aggregated in this dataset.
- Coverage: linguistic transcriptions of the extent or location in addition to the bounding box.
- Main attributes - e.g. values of the parameters X,Y within the time frame T1-T2, etc.
- Data sources, Legal references.
- Importance of the work.

Do not use unexplained acronyms and summarise the most important details in the first sentence or first 100 characters.

Example: *Dataset provides monthly precipitation data collected at different weather stations that are closest to Lake Santo Parmense and Lake Scuro Parmense. We provide a long-term dataset (1951-2010) collected at two stations near the town of Bosco di Corniglio. In particular, from 1951 to 1998 precipitation data were collected at the weather station of Bosco Centrale. Between 1999 and july 2000 the station was out of order (no data available) and from 2001 the Bosco Centrale weather station was substituted by the Bosco di Corniglio station. The two weather stations are less than 1 km apart and they are about 6 km far from Lake Santo Parmense and 7 km far from Lake Scuro Parmense. We also provide a shorter dataset of precipitation data collected from 1994 to 2010 at the weather station of Lagdei. This station is closer to our sampling sites (about 1 km from Lake Santo Parmense and 5 km from Lake Scuro Parmense), but a shorter time series is available. All weather stations are managed by the Environmental Protection Agency of the Emilia-Romagna region (ARPA Emilia-Romagna).*

Keyword set (DEIMS-DS: 9.1 EnvThes Keywords and DEIMS-DS: 9.2 Free Keywords)

Definition: Provides a set of related keywords describing the content of the dataset derived from the controlled vocabulary implemented by EnvThes - thesaurus for long term ecological research, monitoring, experiments EnvThes and other environmentally related thesauri as EUNIS Habitats and INSPIRE Spatial Data Themes. Additional concepts can be defined as free keywords.

Recommendation: The keywords from the following groups can be selected: a) EUNIS habitats, b) Organism, c) Site, d) Research Topics, e) Parameter, f) Method, g) Events, h) System classification, i) Object of interest, and j) INSPIRE Spatial Data Themes.

Example *Permanent oligotrophic lakes, ponds and pools, microclimate, maps, ecosystem ecology, phenological stage, rainfall chemical analysis, organism classification.*

Dataset access and use constraints (DEIMS-DS: 10.1 Principal and granted permission)

Definition: Provide a list of rules defining permissions granted for a dataset.

Recommendation: It is recommended that the 3-year rule be implemented and that the data owner must take specific actions to change this.

Example *Research >> Free for access*

Dataset intellectual rights (DEIMS-DS: 11.1 Intellectual rights)

Definition: Intellectual Rights provides a list of rights management statements for the dataset, or reference a URL (web address) that provides such information. Rights information encompasses Intellectual Property Rights (IPR), copyright, and various property rights. Moreover these rights might include detailed requirements for use, requirements for attribution, or other requirements the owner would like to impose.

Recommendation: Select an option from the list, if any matching. If none matching, use an option Other IPR and provide free textual description, or if IPR information available on an online source, paste the URL pointing to this source. For multiple selection use the CTRL button.

Example *Mutual agreement on reciprocal sharing of data*

Dataset online distribution (DEIMS-DS: 1 2.1 Online Locator)

DEIMS-DS: 12.1.1 Distribution function, DEIMS-DS: 12.1.2 Distribution URL, DEIMS-DS: 12.1.2.1 Web address Title, DEIMS-DS: 12.1.2.2 Web address URL, DEIMS-DS: 12.1.3 Email

Definition: Web address is the "navigation section" of a metadata record pointing users to the location (URL) where a dataset can be retrieved directly, or provides information about how to acquire a dataset.

Recommendation: Setting up the correct resource locators is important for the connection between the data and the services that provide access to them or for providing additional information concerning the resource. If Web address for dataset is available, the Dataset Locator shall be a valid URL providing one of the following: a link to a web page with further instructions; a link to a web service capabilities document; a link to a client application (web data portal) that directly accesses dataset. If a dataset is available offline, it may be uploaded into the system and made online available with access and use constraints and IPR defined previously.

Example *Example Data Service (SOS)*
Web address function: Access to the dataset by LTER Europe FTP, SOS or Linked Data service
Web address title: LTER Europe Sensor Observation Service (SOS)
Web address URL:
http://sp7.irea.cnr.it/tomcat/envsos/sos?REQUEST=getcapabilities&service..

Dataset online distribution (DEIMS-DS: 12.2 WMS Related)

DEIMS-DS: 12.2.1 WMS Map, DEIMS-DS: 12.2.2 WMS Map Web Address, DEIMS-DS: 12.2.2.1 Web address Title, DEIMS-DS: 12.2.2.2 Web address URL

Definition: The WMS Map Web address is the "navigation section" to an of a metadata record pointing users to the location (URL) where a (optional) OGC WMS representation of the dataset can be retrieved directly.

Recommendation: The WMS Map Web Address provides an additional link to a OGC WMS representation of the dataset. This can specifically be done, or using the basic 'Online Resource Locator'.

Example *Web address title: OGC WMS Example Map*
Web address URL:
http://giswebservices.massgis.state.ma.us/geoserver/wms?VERSION=1.1.1&RE...

Data sources (DEIMS-DS: 13.1 Name, DEIMS-DS:13.2 Description, DEIMS-DS:13.3 Source)

Definition: A data source is defined as a physical data upload of a dataset to DEIMS. A data source is described by a number of fields, describing the structure of the data file. For a dataset MD record [0..n] data files can be physically uploaded and described.

Recommendation: The concept of the data source is used, when data files are physically uploaded to DEIMS. This is not recommended. Please used the 'online distribution link' instead, pointing to an online

resource, e.g. file uploaded to B2SHARE.

Example *data file*

Dataset geographic location (DEIMS-DS14.1 Name of Research Location, DEIMS-DS: 14.2 Description, DEIMS-DS: 14.3 Research Location ID, DEIMS-DS:14.4 Related Site, DEIMS-DS:14.5 Geographic Location, DEIMS-DS:14.6 Images, DEIMS-DS:14.7 Details)

Definition: The Geographic reference for the dataset is done by the entity type Research site, which is the location, where specific observations are done. By grouping the information the entity type RESEARCH SITE can be reused. A RESEARCH SITE are the observation plots within a SITE (e.g. LTER SITE or LTSER PLATFORM).
The RESEARCH SITE is the extent of the dataset in the geographic space, given as a bounding box. Defining the coordinates of a boundary rectangle representing the dataset area on a map allows the discovery by geographical area. It consists of: Northern bound coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north), Southern bound coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north), Western bound coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east) and Eastern bound coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east).
In addition the bounding altitudes, images and additional details can be provided.

Recommendation: Please select a Research site from the list, or create a new one with 'Create Research Site' if missing. For the Research Site the bounding box shall be as small as possible and shall be expressed in decimal degree with a precision of at least two decimals. The coordinates of the bounding boxes shall be expressed in any geodetic coordinate reference system with the Greenwich Prime Meridian. You can define bounding box either by drawing a polygon on a map or inserting manually the bounding coordinates. To activate drawing functionality and to (move, delete, reshape) use the editing toolbar in the right top corner of the map.
If you would like to define more bounding coordinates, please add more RESEARCH SITES for the dataset MD record.

Example *Zöbelboden_IP1, Zöbelboden_IP2, Zöbelboden_IP3*

Dataset temporal extent (DEIMS-DS: 15.1 From date and DEIMS-DS: 15.2 To date)

Definition: Defines the time period covered by the content of the dataset. This period may be expressed as a time (an individual date) or date ranges (interval of dates/From-To) or a mix of individual and interval dates.

Example *01/01/2014 to 12/31/2014*

Dataset taxonomic coverage (DEIMS-DS: 16.1 Biological Classification)

Definition: Provides information about the taxonomic classification of the organisms represented in the dataset.

Recommendation: This field is applicable only for biotic data, e.g. when biotic diversity has been chosen in the Keyword tab. Depending on the content of the dataset, provide information about the most common level of taxonomy aggregation (e.g. plants: family, marine invertebrates: phylum or class, etc.)
Recommendation is to use common catalogue of species, for example, GBIF, or EUNIS.

Example *Fagus sylvatica L.*

Dataset method description (DEIMS-DS: 17.1 Method Online Reference and DEIMS-DS: 17.2 Description)

Definition: Provides repeated sets of elements that document a series of procedures followed to produce any dataset object.

Recommendation: If the method description is available online (e.g. previous research resulting in new methodologies, guidelines, specifications, and standards) provide the Method title and URL pointing to the description. This information shall include information about procedure steps, software used within individual steps, source data and any quality measures taken. All information included here should help future data user to evaluate and understand more about the dataset content, thus allow the user to determine whether he/she would be able to combine the dataset within his/her workflows. If not online resource available provide the Method title and comprehensive description. References can be found in Method and related concepts available in EnvThes provided.

Example *Web address title: permanent plots*
Web address URL: http://vocabs.lter-europe.net/EnvThes/USLterCV_409
Description: long-term sample locations used for measuring/estimating biophysical parameters of the environment.

Dataset instrumentation description (DEIMS-DS: 18.1 Instrumentation)

Definition: Provides information about any instruments used in the data collection or quality control and quality assurance.

Recommendation: Instrumentation is a textual description of the used devices including the parameter observed. The description should provide information about: a) parameter, b) device type, c) device brand and type number, d) producer: company or country, e) additional notes
Use the instruments already available in the system through autocomplete functionality if applicable.

Example *Reversing thermometer associated with sampling bottle.*

Dataset sampling description (DEIMS-DS: 19.1 Representative area of sampling and DEIMS-DS: 19.2 Sampling frequency)

Definition: Provides information about sampling part of the method as measurement frequency, and spatial scale.

Recommendation: Select the value for spatial scale, sampling time and minimum sampling unit from the available lists. If available lists do not contain required value, use the last option Other and define new value in the box appeared below.

Example Spatial scale: 10x50M (Line Transects)
Sampling time span: Weekly
Minimum sampling unit: 30min

Quality assurance (DEIMS-DS: 20.1 Quality assurances)

Definition: Provides information on QA/QC procedures applied for the data.

Recommendation: Please provide information on QA/QC procedures applied for the data and quality information in the data.

Example *Values are quality checked using automatic outlier control (R-Script) as well as visual inspection of data. The QA/QC procedure is applied ...*

Reference List n/a

Dataset legal obligation reporting (DEIMS-DS: 21.1 Legal act)

Definition: Provides information whether the dataset has been reported to the local, regional or national bodies to fulfil the obligations from particular legal regulations.

Recommendation: Select from list of EU or other relevant directives. For national or regional directives use option Other and provide references. Even if the element is optional it is recommended to provide information linking to related policies through the regulations.

Example *Water Framework Directive (00/60/EEC)*.

5.3.3 Recommended MD Elements for EO Data

Within WP4, several EO images are processed and several products are derived according to the request of modellers and Protected Area managers for their processing and ingestion on further processes and models. Those products, as approved by unanimous consent of WP4 partners in a face-to-face meeting in Texel (General Assembly 2016, 26-30th June Texel, The Netherlands) must have an ancillary metadata file, in one of the DAB supported format (e.g. Dublin core, ISO19115).

Metadata should include standard elements (regarding its content) but its format must be essentially adapted to the main requirements of the ECOPOTENTIAL project community, in order to increase its usefulness when data sharing especially among project partners. As a result, ECOPOTENTIAL has identified and conformed some metadata elements to their own needs defining a new dimension fitted to existent standard elements which surely goes through existent gaps and helps in further metadata developments. As much as possible, these metadata files will be automatically derived from processing systems and will include a description on how the data has been derived and advice on the constraints on the use of the products that must be taken into account by any user. Metadata should also include a description of quality products to be distributed along with the data itself in a consistent and both human and machine readable ways. The description of general and quality metadata for WP4 derived products is related to work developed in tasks Task 4.6 on EO product quality, validation, intercomparisons and documentation and Task 4.7 on EO data preparation for the protected sites ready to use in WP6, WP7 and WP8. This will be described in more detail in future deliverable D4.6 EO data quality elicitation and documentation. Even though future work is still needed, general guidelines have been produced in WP4 to define minimum metadata elements to be documented and its content, in order to produce useful and consistent metadata records. When possible, the recommendation is to create ISO 19139 XML metadata files. Guidelines on content are described below:

Dataset title

Definition: characteristic, and often unique, name by which the cited dataset is known.

Recommendation: the title should briefly describe not only all the products within the dataset (*e.g.* the acronym of the vegetation index and its error estimation) but also the spatial and temporal extent of the dataset.

Origin: INSPIRE 1205/2008 B.1.1 Resource title + ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:citation//gmd:CI_Citation//gmd:title.

Example: *Dataset title: NDVI and NDVI error for Northern Limestone (2003/07/16)*

Dataset date

Definition: reference date for the cited dataset. Includes date and date type.

Recommendation: at least dataset creation date should be documented. Please note that this is not the date of the image/product (documented on the temporal extent) but the date when the dataset/product is created.

Origin: ISO 19115 md:MD_Metadata/gmd:identification Info/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:date.

Example: **date:** 2016-04-11
dateType: Creation

Dataset unique identifier

Definition: value uniquely identifying the dataset within a namespace. The value domain of this metadata element is a mandatory character string code, generally assigned by the data owner, and a character string namespace uniquely identifying the context of the identifier code (for example, the data owner).

Origin: INSPIRE 1205/2008 B.1.5 Resource unique identifier + ISO 19115 gmd: MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation//gmd:CI_Citation// gmd:identifier/gmd:RS_Identifier (code and codeSpace).

Recommendation: The code should identify the product (*e.g.* the name of the vegetation index) and the codeSpace the organization that has provided it.

Example: **code:** *NDVI_2003_07_16*
codeSpace: *http://www.grumets.uab.cat/*

Abstract

Definition: This is a brief narrative summary of the content of the resource.

Origin: INSPIRE 1205/2008 B.1.2 Resource abstract + ISO 19115 gmd: MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:abstract.

Recommendation: The abstract should identify the product with the name and acronym, and moreover should give details on data type and scale factor applied (if needed) for all the products within the dataset (*e.g.* NDVI and NDVI error in the current example)..

Example: *Normalized Difference Vegetation Index (NDVI) (signed integer file, scale factor 0.001) and an estimation of NDVI error (based on error propagation) on the same units.*

Dataset point of contact

Definition: This is the description of the organisation responsible for the distribution, creation and maintenance of the dataset. This description shall include: the name of the organisation as free text, a contact e-mail address as a character string.

Recommendation: Even tough according to INSPIRE only organization name and an email are needed, the recommendation is to include organisation name and also individual name and personal email (not only the general for the organization) of the dataset processor, in order that the final user can contact them to ask for more details if needed. These people usually are the same than the metadata originators.

Origin: INSPIRE 1205/2008 B.9 Responsible organization + ISO 19115 gmd: MD_Metadata/gmd: identificationInfo/gmd:MD_DataIdentification/gmd:pointOfContact.

Example: ***Role:*** *Processor*
Individual name: *Alaitz Zabala*
Organization name: *Geography Department - Universitat Autònoma de Barcelona*
Electronic mail address: *alaitz.zabala@uab.cat*

Descriptive keywords

Definition: The keyword value is a commonly used word, formalised word or phrase used to describe the subject. While the topic category is too coarse for detailed queries, keywords help narrowing a full text search and they allow for structured keyword search. If the keyword value originates from a controlled vocabulary (thesaurus, ontology), for example GEMET, the citation of the originating controlled vocabulary shall be provided. This citation shall include at least the title and a reference date (date of publication, date of last revision or of creation) of the originating controlled vocabulary.

Recommendation: some theme keywords within GEMET thesaurus should be described. Usually *remote sensing* and *vegetation cover* are the ones that apply to most EO products derived within WP4.

Origin: INSPIRE 1205/2008 B.3 Keyword + ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:descriptiveKeywords/gmd:MD_Keywords/gmd:keyword, gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords/ gmd:MD_Keywords/gmd:type and gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:descriptiveKeywords/gmd:MD_Keywords/gmd:thesaurusName.

Example:

Keyword: *remote sensing*
Type*: Tema*
Thesaurus title*: GEMET Thesaurus version 1.0*
Thesaurus date: *10-06-2008*
Organization name*: European Environment Information and Observation*
Organization Role*: Originator*
Organization URL address: *htpp://www.eionet.europa.eu/gemet*

Dataset topic category

Definition: The topic category is a high-level classification scheme to assist in the grouping and topic-based search of available spatial data resources.

Origin: INSPIRE 1205/2008 B.2.1 Topic category + ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:topicCategory/gmd:MD_TopicCategoryCode.

Recommendation: Topic category should always be described. Typically in products derived within WP4 *biota* applies.

Example: *Biota*

Geographic description

Definition: identifier used to represent a geographic area, describing the extent of the resource.

Origin: ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/ gmd:extent/gmd:EX_Extent/gmd:geographicElement/gmd:EX_GeographicDescription/ gmd:geographicIdentifier/gmd:MD_Identifier/gmd:code.

Recommendation: geographic description using a identifier for the place should be used, typically describing the protected area name in products derived within WP4.

Example: *Northern Limestone*

Spatial representation

Definition: digital mechanism used to represent spatial information.

Origin: ISO 19115 MD_SpatialRepresentation is an abstract and complex element including MD_VectorSpatialRepresentation and MD_GridSpatialRepresentation and, within the latter, MD_Georetified and MD_Georeferenceable: *i.e.* gmd:MD_Metadata/gmd:spatialRepresentationInfo/ gmd:MD_Georectified/gmd:axisDimensionProperties/gmd:MD_Dimension.

Recommendation: within WP4 products, MD_Georetified spatial representation applies. To fully describe this section, at least two axisDimensionProperties should be described, namely "column" and "row", and for each of them dimensionSize and resolution (including units) should be described.

Moreover, an id identifier is given to the full section, namely *spatialRepres*, in order to use it when describing the spatial representation of the CoverageResult for the DQ_QuantitativeAttributeAccuracy quality element (if any).

Example: ***gmd:****MD_Georectified@id: spatialRepres*
dimensionName *: row*
dimensionSize*: 2736*
resolution*:30 m*
dimensionName *: column*
dimensionSize*: 2951*
resolution*:30 m*

Corner points

Definition: earth location in the coordinate system defined by the Spatial Reference System and the grid coordinate of the cells at opposite ends of grid coverage along two diagonals in the grid spatial dimensions.

Origin: ISO 19115 gmd:MD_Metadata/gmd:spatialRepresentationInfo/gmd:MD_Georectified/ gmd:cornerPoints/gml:Point/gml:pos

Recommendation: within WP4 products, gmd:cornerPoints are used to describe the precise spatial extent in the spatial reference system of the dataset. Two corner points are described, with specific gml:id values: NW_corner and SE_corner.

Example: ***cornerPoint@gml:id****: NW_corner*
cornerPoint pos*: 411510 5329740*
cornerPoint@gml:id*: SE_corner*
cornerPoint pos*: 500010 5247690*

Temporal extent

Definition: The temporal extent defines the time period covered by the content of the dataset.

Origin: INSPIRE 1205/2008 B.5 Temporal reference + ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:extent/gmd:EX_Extent/gmd:temporalElement/ gmd:EX_TemporalExtent/gmd:extent/gml:TimeInstant/gml:timePosition.

Recommendation: date of the distributed image or that is the source of the distributed product is described

as an individual date. If the product is derived as a result of several EO images, several individual dates may be described.

Example: *2003-07-16T10:02:00*

Reference system

Definition: identifier and codespace for reference system.

Origin: ISO 19115 gmd:MD_Metadata/gmd:referenceSystemInfo/gmd:MD_ReferenceSystem/ gmd:referenceSystemIdentifier/gmd:RS_Identifier/gmd:code.

Recommendation: reference system using EPSG identifiers is recommended.

Example: *http://www.opengis.net/def/crs/EPSG/0/32633*

Content information

Definition: description of the content of a resource.

Origin: ISO 19115 MD_ContentInformation is an abstract and complex element including MD_FeatureCatalogue Description and MD_CoverageDescription *i.e.* gmd:MD_Metadata/ gmd:contentInfo/gmd:MD_CoverageDescription.

Recommendation: within WP4 products, MD_CoverageDescription content information applies. To fully describe this section attributeDescription, contentType and several dimensions should be described.

attributeDescription is created to describe each dimension name and data type.

When describing each dimension, data type and nodata value is described.

Moreover, an id identifier is given to each dimension section, in order to use it when describing the content description of the CoverageResult for the DQ_QuantitativeAttributeAccuracy quality element (if any).

Example: ***gmd:attributeDescription****:* 1: NDVI value (integer); 2: NDVI error (integer)
gmd:MD_RangeDimension@id*: NDVI_value*
dimension name*: NDVI value*
dimension attribute type name*: signed integer, NoData Value: -32768*
gmd:MD_RangeDimension@id*: NDVI_error*
dimension name*: NDVI error*
dimension attribute type name*: signed integer, NoData Value: -32768*

Distribution information

Definition: information about the distributor of and options for obtaining the resource.

Origin: ISO 19115 MD_Distribution is an aggregated class including several information.

Recommendation: within WP4 products, a single distributor is described in this section, namely the organization generating the product. Within the distributor section several subections are described:

- distributor contact: including organisation name and URL
- distribution format name, including an id identifier in order to use it when describing the format of the CoverageResult for the DQ_QuantitativeAttributeAccuracy quality element (if any). Typically TIFF format is used to distribute products.

- digital transfer options: describing each file distributed within the dataset.

Example: ***distributor organisation name**:* CREAF-UAB
***distributor organisation URL**: http://www.grumets.uab.cat*
***distributor format gmd:MD_Format@id**: distribFormat*
***distributor format name**: Tagged Image File Format (TIFF)*
***distributor transfer options, units of distribution**: Two TIFF files, both signed integer with an scale factor 0.001. One for NDVI and the other for the NDVI error*
***distributor transfer options, URL:**: NDVI_2003_07_16.tif*
***distributor transfer options, URL:**: NDVI_error_2003_07_16.tif*

Lineage

Definition: general explanation of the data producer's knowledge about the lineage of a resource, including events or source data used in constructing the data specified by the scope.

Origin: INSPIRE 1205/2008 B.6.1 Lineage + ISO 19115 gmd:MD_Metadata/gmd:dataQualityInfo/ gmd:DQ_DataQuality/gmd:lineage/gmd:LI_Lineage/gmd:statement.

Recommendation: within WP4 products, the lineage description should include how the data has been derived, the algorithm or equations used, in order to be informative for modellers. In include also scale factor, if needed, and general description of the steps for creating the data (*e.g*. clip, cloud masking, etc.). Typically the scope of this element is "layer", and it is include in a single gmd:dataQualityInfo/gmd:DQ_DataQuality jointly with source description.

Example: *NDVI computation (using Landsat-5 TM bands: ([band4]-[band3])/([band4]+[band3])), scale factor 0.001. Clip to area of interest. Cloud masking. Additionally, the NDVI error (on the same units) is estimated for every sensor and band, and has to be modulated according to the incidence angle (i). The bigger i, the more specular effects in reflectance ant thus the higher errors expected.*

Source description

Definition: information about the source data used in creating the data specified by the scope.

Origin: ISO 19115 gmd:MD_Metadata/gmd:dataQualityInfo/gmd:DQ_DataQuality/gmd:lineage/ gmd:LI_Lineage/gmd:source/gmd:LI_Source/gmd:description.

Recommendation: within WP4 products, the images distributed or the images used to derive the product should be described taking into account to describe sensor, instrument, date, product level, path and row. Typically the scope of this element is "layer", and it is include in a single gmd:dataQualityInfo/gmd:DQ_DataQuality jointly with lineage description.

Example: *Landsat-5 TM 2003-07-16 (surface reflectance), WRS2 path=191 row=27*

Quantitative Attribute Accuracy

Definition: closeness of the value of a quantitative attribute to a value accepted as or known to be true.

Origin: ISO 19157 gmd:MD_Metadata/gmd:dataQualityInfo/gmd:DQ_DataQuality/gmd:report/ gmd:DQ_QuantitativeAttributeAccuracy/gmd:result.

Recommendation: within WP4 products, whenever it is possible, a result describing the quantitative accuracy

of the attribute should be provided as a quantitative result for the whole product or, even better, as a coverage result. If a coverage result is selected (*e.g.* having an "NDVI error" band), then spatial representation, content description and format should be described as links to the previous metadata sections. Semantically it should be related to standard descriptions of quality measures such as the ones on UncertML or QualityML (but this is still work in progress in task 4.6).

Example 1: NDVI RMS error: 0.2

Example 2: **QE_CoverageResult**
MD_SpatialRepresentationTypeCode: Grid
gmi:resultSpatialRepresentation: #spatialRepres"
gmi:resultContentDescription: #NDVI_error
gmi:resultFormat: #distribFormat"

Metadata contact

Definition: This is the description of the organisation responsible for the creation and maintenance of the metadata. This description shall include: the name of the organisation as free text, a contact e-mail address as a character string.

Recommendation: Even though according to INSPIRE only organization name and an email are needed, the recommendation is to include organisation name and also individual name and personal email (not only the general for the organization) of the metadata generator, in order that the final user can contact them to ask for more details if needed. These people usually are the same than the dataset processors.

Origin: INSPIRE 1205/2008 B.10.1 Metadata point of contact + ISO 19115 gmd:MD_Metadata/gmd:contact/ gmd:CI_ResponsibleParty

Example: ***Role:*** *Originator*
Individual name: *Alaitz Zabala*
Organization name: *Geography Department - Universitat Autònoma de Barcelona*
Electronic mail address: *alaitz.zabala@uab.cat*

Other recommended metadata elements

Other metadata elements are recommended to be described in the metadata records produced within WP4, but with no special recommendation beyond metadata element description in the standards. A list of these elements can be found below.

Metadata file identifier: ISO 19115 gmd:MD_Metadata/gmd:fileIdentifier.

Metadata language: INSPIRE 1205/2008 B.10.3 Metadata language + ISO 19115 gmd:MD_Metadata/gmd:language/gmd:LanguageCode.

Metadata character set: ISO 19115 gmd:MD_Metadata/gmd:characterSet/gmd:MD_CharacterSetCode.

Metadata hierarchy level and hierarchy level name: INSPIRE 1205/2008 B.1.3 Resource type + ISO 19115 /gmd:MD_Metadata/gmd:hierarchyLevel/gmd:MD_ScopeCode and /gmd:MD_Metadata/ gmd:hierarchyLevelName/gco:CharacterString.

Metadata creation date: INSPIRE 1205/2008 B.10.2 Metadata date + ISO 19115 gmd:MD_Metadata/ gmd:dateStamp/gco:Date.

Metadata standard name and version: ISO 19115 gmd:MD_Metadata/gmd:metadataStandardName and /gmd:MD_Metadata/gmd:metadataStandardVersion.

Spatial representation type: ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:spatialRepresentationType/ gmd:MD_SpatialRepresentationTypeCode. Usually Grid is the value that applies for products distributed within WP4.

Spatial resolution: INSPIRE 1205/2008 B.6.2 Spatial resolution + ISO 19115 gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:spatialResolution/gmd:MD_Resolution/ gmd:distance/gco:Distance. resolution distance equivalent to pixel size (including it units) should be documented here.

Dataset language: INSPIRE 1205/2008 B.1.7 Resource language + ISO 19115 gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:language/gmd:LanguageCode.

Geographic bounding box: INSPIRE 1205/2008 B.4.1 Geographic bounding box + ISO 19115 gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:extent/ gmd:EX_Extent/gmd:geographicElement/gmd:EX_GeographicBoundingBox.

5.3.4 Recommended MD elements for documentation of Data Quality

The ECOP researchers who responded (54) to the online questionnaire (fig 4-5) described in 3.3 (Method applied to evaluate relevant quality elements):

i. classify their position relating spatial data life cycle, mainly as data user or present equilibrium about double nature as data provider/supplier and data user/consumer;
ii. indicate their limited knowledge about (spatial) data quality/domain of spatial/geographic data concepts and related ISO standards as well as, INSPIRE, EML, DEIMS, Dublin and Darwin Core metadata standards/application schemas;
iii. At same time, the user show knowledge relating data quality elements according ISO19157:2013 and limited experience in implementing quantitative/qualitative procedures of spatial data quality management along spatial data life cycle and consider or communicate in quality assurance (QA)/quality control (QC) in contrast of explicit high interests in know/use spatial data quality elements, in know/use spatial data quality assessment methods and tools as well as, in use/participate in spatial data quality management process;
iv. consider decisive data quality elements (ISO19157:2013) in order to discover and select input data for applying socio-ecological/environmental models and workflows, to explore the results of practical/ecological meaning of output models data and to communicate with end user/technical-political decision makers;
v. Indicate as important, critical and relevant spatial data elements/indicators to incorporate into a metadata profile to support/facilitates data quality assessment and management in scientific collaborative network management.

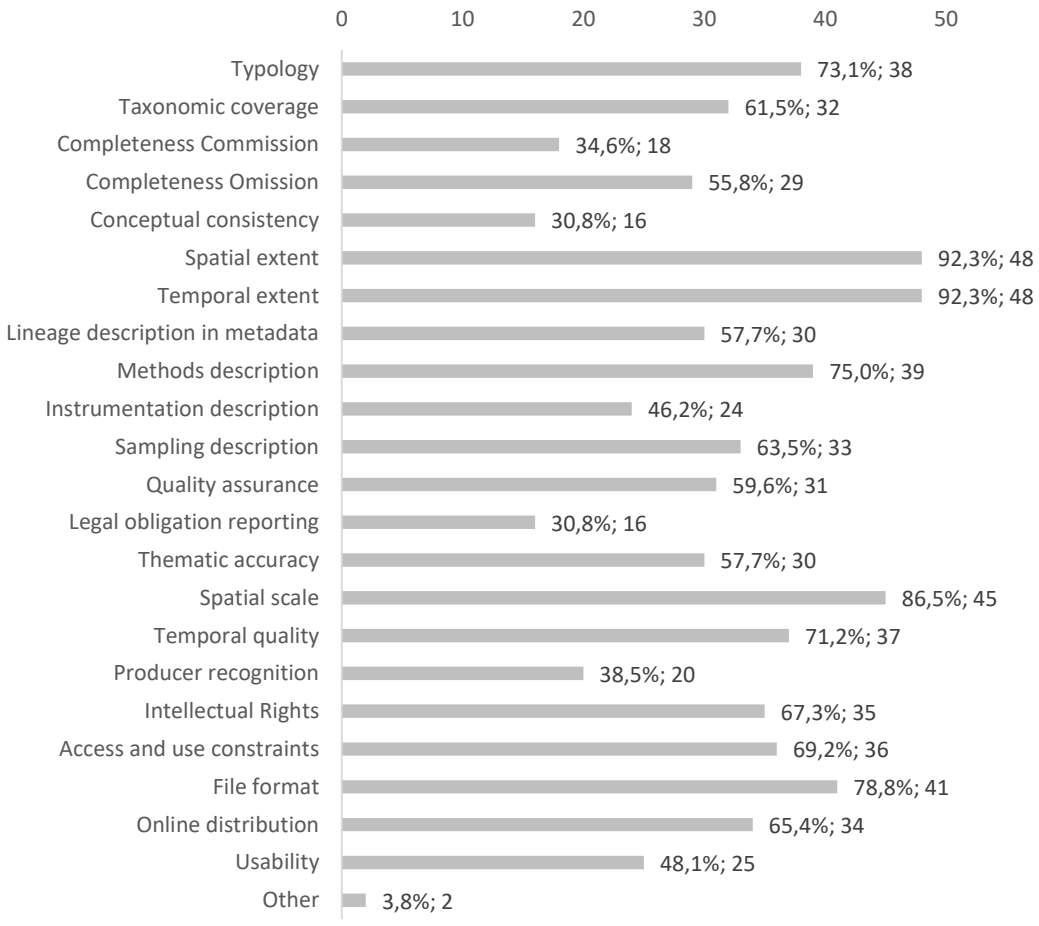

Fig. 5-5. Relevant spatial elements/indicators to incorporate into a spatial metadata profile in order to support/facilitates data quality assessment and management in scientific collaborative network management.

The online questionnaire results will be of high relevance for future activities namely when:

(1) defining the fields/attributes of quality elements to include in the selected metadata profile (Task 5.3);
(2) specifying and developing methods (external evaluation) and tools (routines and information/ technological application) of spatial data quality assessment; and
(3) devising proposals for implementation of quality management processes (Task 5.5).

Important progress was made in the activities to support the identification, analysis and specification of requirements for the user-oriented quality evaluation routines, supported on metadata. To evaluate external data quality of spatial datasets based on metadata, it is necessary that metadata entries contain information about the content, quality parameters, access and use conditions, and other characteristics of datasets that can be used for external (meta) quality evaluation and also for knowledge discovery, indexing and searching. As such, efforts have been made to ensure that the metadata profile used in other tasks will allow the application of user-oriented quality evaluation routines. This assessment was made in cooperation with Task 5.5, and new fields (quality elements) to be included in-situ and other spatial metadata profile were discussed in order to effectively allow quality evaluation based on metadata. A comparative analysis of different metadata profiles indicates an insufficient and omissions relating of quality elements fields (Tab. 5-2). The analysis indicates the importance of increasing and detailing the quality elements fields related to descriptors of quality elements according to ISO 19157 in order to improve and guarantee quality control in the production and maintenance of quality throughout the data life cycle (Tab. 5-3).

Tab. 5-2. Description and nature of fields included ISO 19115, INSPIRE/ISO19139, DEIMS-SDR and ISO191158 metadata profiles

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
Metadata	This is the description of the organisation responsible for the creation and maintenance of the metadata. The date which specifies when the metadata record was created or updated. This is the language in which the metadata elements are expressed. [Dataset Metadata Provider] [Dataset Metadata date]	Text Date Text Date	Mandatory	Mandatory	Mandatory	---
Resource title	This is a characteristic, and often unique, name by which the resource is known. This field refers to the title of a specific dataset [e.g. a dataset of distribution information for the population of bats should be referred as “bats distribution data”]. The titles should be short (in length) and objective. [Dataset Title]	Text Text	Mandatory	Mandatory	Mandatory	---
Resource abstract	This is a brief narrative summary of the content of the	Text Text	Mandatory	Mandatory	Mandatory	---

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
	resource with no more than 200 characters. [Dataset Abstract]					
Resource type	Scope to which metadata applies "MD_ScopeCode" (Data type – see annex B.5.25 of ISO 19115)	List	Mandatory	Mandatory	---	Mandatory
Resource locator	Location (address) for on-line access using a Uniform Resource Locator address or similar addressing scheme [Dataset Online distribution-function]	URL List	Optional	Conditional	Optional	---
Unique resource identifier	Value uniquely identifying an object within a namespace "MD_Identifier".	URI	---	Mandatory	---	---
Topic category	The topic category is a high-level classification scheme to assist in the grouping and topic-based search of available spatial data resources "MD_TopicCategory", according to the ISO 19115.	List	Mandatory	Mandatory	---	---
Keywords	Commonly used word(s) or formalised word(s) or phrase(s) used to describe the subject. Note: related to identification of the INSPIRE Spatial Data themes (thematic category) according to Directive INSPIRE. The keyword set is related keywords describing the content of the dataset derived from the controlled vocabulary implemented by EnvThes-thesaurus for long term ecological research, monitoring, experiments EnvThes and other environmentally related thesauri as EUNIS Habitats and INSPIRE Spatial Data Themes.	Text List	---	Mandatory [Keywords]	Mandatory [EnvThes Keywords] Optional [Free keywords]	---
Spatial resolution	Spatial resolution refers to the level of spatial detail of the dataset. It shall be expressed as a set from zero to many resolution distances (typically for gridded data and imagery-derived products) or equivalent scales (typically for maps or map-derived products). An equivalent scale is generally expressed as an integer value expressing the scale denominator. A resolution distance shall be expressed as a numerical value associated with a unit of length. [Spatial scale]	Numeric Numeric	Optional	Mandatory	Mandatory	---

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
Temporal extent	The temporal extent defines the time period covered by the content of the resource. [Dataset Temporal extent]	Date	Optional	Mandatory [Conditional]	Mandatory	---
Date of publication	This is the date of publication of the resource when available, or the date of entry into force (Date of publication; Date of last revision; Date of creation). [Dataset Publication Date]	Date	Mandatory	Mandatory [Conditional]	Optional	---
Geographic bounding box	This field refers to the geographical scope of the dataset, particularly whether the dataset covers all or just a portion of the study area. The bounding box shall be expressed with westbound and eastbound longitudes, and southbound and northbound latitudes in decimal degrees, with a precision of at least two decimals. [Geographic – Research site] ->->-> **Dataset Geographic extent**	Numeric	Conditional	Mandatory	Mandatory	---
File format	This field refers to the type of file of the dataset (distribution format).	Text	Optional	---	---	---
Author	This field refers to the institution or individual that produced the dataset (Responsible organisation). [Producer recognition]	Text List	Mandatory	Mandatory	Optional	---
Property	This field refers to the property of the dataset being necessary to state if there are any conditions applying to its access and use. [Dataset Access and Use constraints]	List List	---	Mandatory	Optional	---
Spatial Reference System	This field refers to the geographical reference system of the dataset.	List	Optional	---	---	---
Lineage (Phases of the life cycle considered are: specification, production, delivery, use and update)	Lineage describes the history of a dataset and recount the life cycle of a dataset from collection and acquisition through compilation and derivation to its current form. This general, non-quantitative information is illustrative for users and can help assessing the quality of a dataset, especially in cases where it is used for a particular application that differs from the intended application. [Dataset Methods description] [Dataset Instrumentation] [Dataset Sampling description]	Text Text Text List/ Text	Optional	Mandatory	Mandatory Optional Mandatory	---

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
Maintenance information	This field refers to the information about the scope and frequency of updating	Text	---	Mandatory	---	---
Data quality	Quality evaluation processes are used **in different phases of a product life cycle, having different objectives in each phase.** The process for evaluating data quality is a sequence of steps to produce a data quality result, namely: i) **data quality unit(s) identification**: composed of a **scope** and **quality element(s),** all data quality elements relevant to the data for which quality is to be described should be used; ii) **data quality measures identification**: if applicable a measure should be specified for each data quality element; iii) **data quality evaluation procedures selection**: a data quality evaluation procedure consists of applying one or more evaluation methods; iv) **determine the output of the data quality evaluation** (quantitative, qualitative/descriptive, conformance): this result is the output of applying the evaluation; v) **reporting data quality** (metadata). **Data quality evaluation process** (1) **Description of data quality unit(s**) **Scope** [MD_Scope] ISO19115-1:2013; List Reference: a) a dataset series; b) a dataset; c) a subset of data defined by one or more of the following characteristics: 1) types of items (sets of feature types, feature attributes, feature operations or feature relationships); 2) specific items (sets of feature instances, attribute values or instances of feature relationships); 3) geographic extent; 4) Temporal extent (the time frame of reference and accuracy of the time frame). (2) **Data quality evaluation**	Text List List List Numeric Text/List Text Text List Text	---	Mandatory [Scope] [Specification] [Conformity-Degree] [Lineage]		Mandatory

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
	element(s) [DQ_Element(s)]*** (3) **Data quality measures** [DQM_Description_DataType] (MD_Identifier_DataType ISO19115-1:2013); (4) **Data quality evaluation method** [EvaluationMethodType] CodeList: directInternal; directExternal; indirect. (5) **Result** (Meta) quality elements are a set of quantitative and qualitative statements about a quality evaluation and its result. The knowledge about the quality and the suitability of the evaluation method, the measure applied and the given result may be of the same importance as the result itself. **Data quality result (metaquality)** (Meta)quality may be described using the following elements (report): (1) **Confidence** – trustworthiness of a data quality result. (2) **Representativity** – degree to which the sample used has produced a result which is representative of the data within the data quality scope. (3) **Homogeneity** – expected or tested uniformity of the results obtained for a data quality evaluation. *(see ISO19157:2013 - Geographic information - Data quality; and XML implementation)*					
Quality assurance	Quality assurance (QA), is defined with "all the planned and systematic activities implemented within the quality system, and demonstrated as needed, to provide adequate confidence that as entity will fulfil requirements for quality". This analysis is done according to a requisition (product, legal or normative obligation, etc. ...) Recommendations: Several spatial data producers normally have an obligation of mean but no obligation of result. Consequently, QA is not	List	---	---	Optional Optional	Mandatory

Metadata	Description	Type of field	ISO19115	INSPIRE /ISO19139	DEIMS	ISO19157
	sufficient since it cannot guarantee the production of quality data. *(see ISO19158:2012 - Geographic information - Quality assurance of data supply)* [Dataset Legal obligation reporting- Legal act] [Quality Assurance]					

The analysis of the data tables indicates the importance and possibility of suggesting / adding new quality elements to the metadata profiles in order to complement the missing elements as well as to value / support the data quality evaluation exercises/processes (Tab. 5-3).

*Tab. 5-3. Examples the data quality elements, sub-elements and measures (according ISO19157:2013) (*** the selection of data quality elements should consider the scope)*

DQ_Element(s)	Examples DQ_SubElement(s)	Examples Measures (DQM_Description_Datatype)
Completeness	Completeness commission	Number of excess items/ Integer/ error count
	Completeness omission	Number of missing items/ Integer/ error count
	…	…
Consistency	Conceptual consistency	Number of items not compliant with the rules of the conceptual schema/ Integer/ error count
	Domain consistency	Number of items not in conformance with their value domain/ Integer/ error count
	Format consistency	Number of physical structure conflicts/ Integer/ error count
	…	…
Positional Accuracy	Vertical positional uncertainties	Root mean square error (RMSE)/ Measure
	Horizontal positional uncertainties	Root mean square error of planimetry (RMSEP)/ Measure
	Gridded Data Positional Accuracy	Select measure
	…	…
Thematic Accuracy	Thematic classification correctness	Incorrectly classified features/ %/ error indicator
		Misclassification matrix/ Integer – Matrix (n*n)
		Kappa coefficient/ Integer
	…	…
Temporal Quality	Accuracy of a time measurement	Time accuracy at 50% significance level (interval defined by an upper and a lower limit)/ Measure
	Temporal consistency	Chronological error/ Boolean/ error indicator
	…	…
Usability	Usability element	Degree of adherence of a dataset to a specific set of requirements (specifications)
	…	…

Tab. 5-4. Proposal the metadata fields/elements for spatial data quality evaluation procedure (task 5.3 and ask 5.5).

Main Characteristics Quality	Data Quality Element (ISO19157)	Adaptation of Data Set Metadata element for ECOPOTENTIAL Community Profile
Definition		Keywords
		Resource type (scope)
	Typology	Topic category
		Taxonomic coverage
		Date of publication (metadata)
		Maintenance information
Data quality (internal)	Completeness	Commission
		Omission
	Logical Consistency	Conceptual consistency
		Domain consistency
		Format consistency
		Topological consistency
Coverage	Spatial/Geographic	Geographic bounding box
	Temporal	Temporal extent
	Geographic (Research site)	Geographic (Research site)
Lineage	Lineage (process step, …)	Lineage
		Methods description
		Instrumentation description
		Sampling description
		Quality assurance
		Legal obligation reporting
Data quality (internal)	Positional Accuracy	Absolute external positional accuracy
		Relative internal positional accuracy
		Gridded data positional accuracy
	Temporal Quality	Accuracy of a time measurement
		Temporal consistency
		Temporal validity
	Thematic Accuracy	Thematic classification correctness
		Non-quantitative attribute correctness
		Quantitative attribute accuracy
Precision		Spatial resolution
		Spatial Reference System
Legitimacy		Author
		Intellectual Rights
Accessibility		Property
		File format
		Resource locator
Data quality (internal)	Usability	Usability element

In addition to the in-situ spatial datasets/databases used, the ECOP community considers in on-line questionnaire other relevant several reference and thematic datasets in particular satellite and aerial images and other databases produced from remote sensing/earth observation products (Tab. 5-5).

Technical Guidance for the implementation of INSPIRE dataset and service metadata based on ISO/TS 19139:2007 relates to the implementation of the requirements related to metadata for spatial data sets and series and spatial data services (including network services) included in [Regulation 1205/2008] and [Regulation 1089/2010] on the INSPIRE framework. INSPIRE will make available relevant, harmonized and quality geographic information to support the formulation, implementation, monitoring and evaluation of policies and activities, which have a direct or indirect impact on the environment. This Technical Guidance recommends data quality critical quality elements according INSPIRE spatial data themes (dataset/databases nature) (Tab. 5-6) as well it´s important to define associated concepts (Tab. 5-7).

These data quality elements should support, consider and result in data quality evaluation procedures. These procedures provide information in addition to the scope, conformity[31] and lineage[32]. This information may include information on the source data used and the main transformation steps that took place in creating the current data set or data set series) according data life cycle.

[31] Comment: The metadata shall, in conformance to [INSPIRE Directive], include a statement on the degree of conformity with the specifications against which its conformity has been evaluated.

[32] Comment: Lineage is defined as the processing history of result data set or data set series can provide valuable information about the applicability of the data for a particular use.

Tab. 5-5 – Spatial Data Quality Evaluation

Main Characteristics	Quality elements/ Indicators	Description / Value domain / Comparison type
Definition	Keywords	The keyword set is related keywords describing the content of the dataset derived from the controlled vocabulary implemented by EnvThes-thesaurus for long term ecological research, monitoring, experiments EnvThes and other environmentally related thesauri as EUNIS Habitats and INSPIRE Spatial Data Themes.
	Resource type (scope)	
	Topic category	Topic category defining the main data set theme
	Taxonomic coverage	Provides information about the taxonomic classification of the organisms represented in the dataset.
	Date of publication	This is the date of publication of the resource when available, or the date of entry into force.
	Maintenance information	
Completeness	Completeness Commission	Number of excess items/ Integer/ error count
	Completeness Omission	Number of missing items/ Integer/ error count
Consistency	Conceptual consistency	Number of items not compliant with the rules of the conceptual schema/ Integer/ error count
	Domain consistency	Number of items not in conformance with their value domain/ Integer/ error count
	Format consistency	Number of physical structure conflicts/ Integer/ error count
	Topological consistency	
Coverage	Geographic bounding box	Bounding box defining spatial coverage/ Geometric bounding box/ % of cover
	Temporal extent	Time interval defining temporal coverage/ Start date – End date/ % of cover
	Geographic (Research site)	The Geographic reference for the dataset is done by the entity type Research site, which is the location, where specific observations are done. By grouping the information the entity type RESEARCH SITE can be reused. A RESEARCH SITE are the observation plots within a SITE (e.g. LTER SITE or LTSER PLATFORM).
Lineage	Lineage description in metadata	Description of data production methods and overall quality/ Text-character-string/ Boolean comparison
	Methods description	Provides repeated sets of elements that document a series of procedures followed to produce any dataset object.
	Instrumentation description	Provides information about any instruments used in the data collection or quality control and quality assurance.
	Sampling description	Provides information about sampling part of the method as measurement frequency, and spatial scale.
	Quality assurance	Provides information on QA/QC procedures applied for the data.
	Legal obligation reporting	Provides information whether the dataset has been reported to the local, regional or national bodies to fulfil the obligations from particular legal regulations.

Main Characteristics	Quality elements/ Indicators	Description / Value domain / Comparison type
Precision	Thematic accuracy	Incorrectly classified features/ %/ error indicator
		Misclassification matrix/ Integer – Matrix (n*n)
		Kappa coefficient/ Integer
	Spatial scale	Equivalent scale or spatial resolution defining the level of detail / Integer – Double/ Intersection test
	Spatial Reference System	This field refers to the geographical reference system of the dataset.
	Positional accuracy	Mean value of positional uncertainties
		Mean value of positional uncertainties excluding outliers
		Number of positional uncertainties above a given threshold
		Vertical positional uncertainties / Root mean square error (RMSE)/ Measure
		Horizontal positional uncertainties / Root mean square error of planimetry (RMSEP)/ Measure
		Gridded Data Positional Accuracy/ select measure
	Temporal quality	Accuracy of a time measurement - Time accuracy at 50% significance level (interval defined by an upper and a lower limit)/ Measure
		Temporal consistency - Chronological error/ Boolean/ error indicator
Legitimacy	Author (Producer recognition)	Data producer recognition type/ Text-string/ String comparison
	Intellectual Rights	Intellectual Rights provides a list of rights management statements for the dataset, or reference a URL (web address) that provides such information. Rights information encompasses Intellectual Property Rights (IPR), copyright, and various property rights. Moreover, these rights might include detailed requirements for use, requirements for attribution, or other requirements the owner would like to impose.
Accessibility	Property (Access and use constraints)	Conditions applying to access and use (Directive 2007/2/EC) / Text-string/ String comparison
	File format	Distribution file format/ Text-string/ String comparison
	Resource locator (online distribution)	Web address is the "navigation section" of a metadata record pointing users to the location (URL) where a dataset can be retrieved directly, or provides information about how to acquire a dataset.
Data quality (internal)	Usability Element	Degree of adherence of a dataset to a specific set of requirements (specifications)

Tab. 5-6 – Proposal of metadata fields/elements for spatial data quality evaluation according INSPIRE themes (INSPIRE metadata elements concerning data quality)

Main Characteristics Quality	**Data Quality Element** (ISO19157)	**Metadata element**	**INSPIRE Spatial Data themes** (see Table 5)																																	
			RS	GG	GN	AU	AD	CP	TN	HY	PS	EL	LC	OI	GE	SU	BU	SO	LU	HH	US	EF	PF	AF	PD	AM	NZ	AC	MF	OF	SR	BR	HB	SD	ER	MR
Definition	Typology	Keywords			x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	X	x	x	x
		Resource type (scope)			x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Topic category			x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Taxonomic coverage																				x								x	x		x	x		
		Date of publication			x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Maintenance information			x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Data quality (internal)	Completeness	Commission				x	X		x	x	x	x	x			x	x		x		x				-											
		Omission			x	x	X	x	x	x	x	x	x	x		x	x	x	x		x		x	x	-					x						
	Logical Consistency	Conceptual consistency			x	x	X	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x	x	x
		Domain consistency			x	x	X	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x	x	x
		Format consistency							x			x	x								x				-											
		Topological consistency							x	x																										
Coverage	Spatial/Geographic	Geographic bounding box			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Temporal	Temporal extent			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Geographic (Research site)	Geographic (Research site)			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Lineage	Lineage (process step, …)	Lineage			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Methods description			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Instrumentation description			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Sampling description			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Quality assurance			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Legal obligation reporting			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Data quality (internal)	Positional Accuracy	Absolute external positional accuracy			x	x	x	x	x	x	x	x	x			x	x		x	x					-					x	x					
		Relative internal positional accuracy								x			x												-											

		Gridded data positional accuracy								x		x											-											
	Temporal Quality	Accuracy of a time measurement																																
		Temporal consistency									x												-											
		Temporal validity									x			x				x			x	x	-											
	Thematic Accuracy	Thematic classification correctness					x				x			x			x				x	x	-											
		Non-quantitative attribute correctness			x		x	x			x						x						-											
		Quantitative attribute accuracy						x			x							x					-											
Precision		Spatial resolution	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Spatial Reference System	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Legitimacy		Author	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Intellectual Rights	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Accessibility		Property	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		File format	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		Resource locator	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Data quality (internal)	Usability	Usability element													x													x	x					

Tab. 5-7. Legend INSPIRE Spatial Data Themes.

INSPIRE		
Annex I	01. Coordinate reference systems	RS
	02. Geographical grid systems	GG
	03. Geographical names	GN
	04. Administrative units	AU
	05. Addresses	AD
	06. Cadastral parcels	CP
	07.Transport networks	TN
	08. Hydrography	HY
	09. Protected sites	PS
Annex II	01. Elevation	EL
	02. Land cover	LC
	03. Orthoimagery	OI
	04. Geology	GE
Annex III	01. Statistical units	SU
	02. Buildings	BU
	03. Soil	SO
	04. Land use	LU
	05. Human health and safety	HH
	06. Utility and governmental services	US
	07. Environmental monitoring facilities	EF
	08. Production and industrial facilities	PF
	09. Agricultural and aquaculture facilities	AF
	10. Population distribution - demography	PD
	11. Area management/ restriction/ regulation zones and reporting units	AM
	12. Natural risk zones	NZ
	13. Atmospheric conditions	AC
	14. Meteorological geographical features	MF
	15. Oceanographic geographical features	OF
	16. Sea regions	SR
	17. Bio-geographical regions	BR
	18. Habitats and biotopes	HB
	19. Species distribution	SD
	20. Energy resources	ER
	21. Mineral resources	MR

The literature review allowed to update the current state of knowledge about important issues to be considered in the development of the data quality evaluation procedures in ECOPOTENTIAL, namely:

(1) novel challenges in spatial data quality assessment are raised by the increasing amount of spatial data production, and by handling and sharing of data with different types and sources, different frequency of acquisition, different spatial/temporal scales, different levels of accuracy, different processing methods or techniques (e.g. Earth observations, field measurements, data analysis and modelling);
(2) spatial data are used in very different application contexts (e.g. ecosystem service assessments, conservation of natural heritage);
(3) neglecting data quality fosters the risk of misuse and misinterpretation by users and can cause misleading results; and

Co-funded by the European Union

(4) data quality must be an essential criterion to identify datasets that satisfy the requirements of a particular application for a specific user (e.g. when developing statistical or process-based models to support PA management).

Therefore, measuring, assessing, managing and communicating data quality is important throughout the data life cycle. Data quality can be categorized in two broad and complementary approaches:

(1) Internal quality, related with the intrinsic characteristics of the data as described at the producer level and usually detailed in metadata; and
(2) External quality, which refers to the level of similarity between the characteristics of the data and the user's needs in a specific context of application, i.e. fitness for use.

Both approaches have received attention in geographical information science (ISO 19157), with an emphasis so far on the evaluation of internal quality. The work in progress in task 5.5 is focused on external quality evaluation (based on user's requirements), but beforehand internal quality information (data product specification) must be able to make effective and useful external quality assessment based on metadata. As such, the external quality of datasets is assessed for each specific application context, by measuring the match between the characteristics of the data, as detailed in metadata entries (often described at producer level), and the characteristics of the data as required by users (quality expectations).

Important progress was made in the activities to support the identification, analysis and specification of requirements for the user-oriented quality evaluation routines, supported on metadata. To evaluate external data quality of spatial datasets based on metadata, it is necessary that metadata entries contain information about the content, quality parameters, access and use conditions, and other characteristics of datasets that can be used for external (meta) quality evaluation and for knowledge discovery, indexing and searching. As such, efforts have been made to ensure that the metadata profile used in other tasks will allow the application of user-oriented quality evaluation routines. This assessment was made in cooperation with Task 5.3, and new fields (quality elements) to be included in the DEIMS-SDR Metadata profile were discussed in order to effectively allow quality evaluation based on metadata. The work in progress in task 5.5 is focused on external quality evaluation (based on user's requirements), but beforehand internal quality information (data product specification) must be able to make effective and useful external quality assessment based on metadata. As such, the external quality of datasets is assessed for each specific application context, by measuring the match between the characteristics of the data, as detailed in metadata entries (often described at producer level), and the characteristics of the data as required by users (quality expectations).

6 Technical implementation

In order to collect the metadata of existing in-situ data we started from the information provided by protected areas in several documents developed during the project. These documents were related with protected area description, data availability and ecosystem services. One of the most consistent source for metadata collection were the storylines that provided a common framework for displaying the in-situ data. It should be kept in mind that the storylines are specific to some research questions.

6.1 Metadata authoring

In order to collect the desired metadata the DEIMS-SDR system was used (see chapter 3.2.5 LTER Europe & DEIMS-SDR). For that purpose, we created a dedicated user account ('*ecopotential*') in order to upload information about the sites, data products and datasets.

If a site had not been documented on DEIMS-SDR yet a new site record was created. According to the dataset metadata model of DEIMS-SDR datasets can be created and linked to the corresponding site records. When a user creates a data product or dataset metadata record, a related site and a number of datasets (in the case data products) can be linked to those records[33].

The information provided in all records affiliated with a site is used to generate an INSPIRE EF compliant metadata record that features parts of the provided information. The schema applied for the EF application schema currently does not allow exposing all available information on DEIMS.

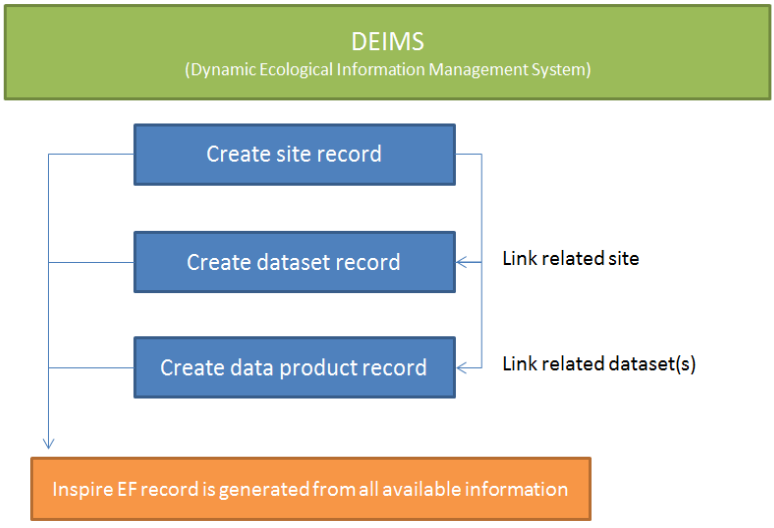

Fig. 6-1. Metadata workflow in DEIMS-SDR

In order to collect the metadata of existing in-situ data we started from the information provided by protected areas in several documents developed during the project. These documents were related with protected area description, data availability and ecosystem services. One of the most consistent sources for

[33] For a detailed list of steps to enter metadata , please see: https://docs.google.com/document/d/1F828tgVtt3HhbGtIMa0YCcco_Vy-KQEjnH4eVP5eoqE/edit

metadata collection was the storylines that provided a common framework for displaying the in-situ data. It should be kept in mind that the storylines are specific to some research questions.

The information was checked for each protected area and grouped by similarity into an upper level called data products. The similarity took into account the aim of the protected area, the dominant ecosystem type, the type of earth observation, and the type of measurement. For each data product, a description was issued in a structured way being in fact the metadata of the data product. The structure of the metadata is based on the following information:

- Title: title of the data product;
- Site / Protected area: reference to the protected area(s) in which the data were generated;
- Data product classification: classification of the data product based on a taxonomy;
- Abstract: textual description of the content of the data and the methods used;
- Parameter (optional): more detailed information on the parameters provided in the data product;
- Keywords: providing keywords (based on EnvThes) to annotate the data product and allow the user a more targeted search and discovery;
- Time period: in which the data are generated;
- Temporal resolution: temporal resolution of the data;
- Spatial resolution: spatial resolution of the data collection;
- Data digital available: are the data available in digital format;
- Link: link to the data portal;
- Data publically available: [yes/no];
- Data available for ECOPOTENTIAL: [yes/no];
- Contact: contact person
-

6.2 Discovery catalogue for In-situ data

The ECOPOTENTIAL In-situ metadata catalogue[34] is a web service to publish and search in-situ metadata from different protected areas and institutions involved in the project (fig 28). It is based on the OGC CSW[35] (Catalogue Service for the Web), supports both CSW 2.0.2 and 3.0.0 versions and is using a common metadata model to harmonise the metadata in a database.

[34] http://eco.starlab.es/csw

[35] http://www.opengeospatial.org/standards/cat

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- pycsw 2.1-dev -->
- <csw30:Capabilities xsi:schemaLocation="http://www.opengis.net/cat/csw/3.0 http://schemas.opengis.net/csw/3.0/cswGetCapabilities.xsd"
version="3.0.0" updateSequence="1498743850" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:ows20="http://www.opengis.net/ows/2.0" xmlns:ows11="http://www.opengis.net/ows/1.1" xmlns:ows="http://www.opengis.net/ows"
xmlns:inspire_ds="http://inspire.ec.europa.eu/schemas/inspire_ds/1.0"
xmlns:inspire_common="http://inspire.ec.europa.eu/schemas/common/1.0" xmlns:gml="http://www.opengis.net/gml"
xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:fes20="http://www.opengis.net/fes/2.0" xmlns:dct="http://purl.org/dc/terms/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:csw30="http://www.opengis.net/cat/csw/3.0"
xmlns:csw="http://www.opengis.net/cat/csw/2.0.2">
  - <ows20:ServiceIdentification>
      <ows20:Title>EcoPotential WP5 In-situ Metadata provider</ows20:Title>
      <ows20:Abstract>OGC CSW server implementation written in Python</ows20:Abstract>
    - <ows20:Keywords>
        <ows20:Keyword>catalogue</ows20:Keyword>
        <ows20:Keyword>discovery</ows20:Keyword>
        <ows20:Keyword>metadata</ows20:Keyword>
        <ows20:Type codeSpace="ISOTC211/19115">theme</ows20:Type>
      </ows20:Keywords>
      <ows20:ServiceType codeSpace="OGC">CSW</ows20:ServiceType>
      <ows20:ServiceTypeVersion>2.0.2</ows20:ServiceTypeVersion>
      <ows20:ServiceTypeVersion>3.0.0</ows20:ServiceTypeVersion>
      <ows20:Profile>http://www.opengis.net/cat/csw/apiso/1.0</ows20:Profile>
      <ows20:Fees>None</ows20:Fees>
      <ows20:AccessConstraints>None</ows20:AccessConstraints>
    </ows20:ServiceIdentification>
  - <ows20:ServiceProvider>
      <ows20:ProviderName>Starlab Barcelona</ows20:ProviderName>
      <ows20:ProviderSite xlink:href="http://www.starlab.es/" xlink:type="simple"/>
    - <ows20:ServiceContact>
        <ows20:IndividualName>Prados, Jordi</ows20:IndividualName>
        <ows20:PositionName>Software Engineer</ows20:PositionName>
      - <ows20:ContactInfo>
        - <ows20:Phone>
            <ows20:Voice>missing</ows20:Voice>
            <ows20:Facsimile>missing</ows20:Facsimile>
          </ows20:Phone>
        - <ows20:Address>
            <ows20:DeliveryPoint>Av. Tibidabo 47bis</ows20:DeliveryPoint>
            <ows20:City>Barcelona</ows20:City>
            <ows20:AdministrativeArea>missing</ows20:AdministrativeArea>
            <ows20:PostalCode>missing</ows20:PostalCode>
            <ows20:Country>Spain</ows20:Country>
            <ows20:ElectronicMailAddress>jordi.prados@starlab.es</ows20:ElectronicMailAddress>
```


Fig. 6-2. ECOPOTENTIAL Catalogue - GetCapabilities request

The common metadata model consists of different metadata standards: Dublin Core Metadata, ISO 19115 and ISO 19119, but this model can be extended by creating different metadata extensions to add more fields and functionalities to the metadata model. To harmonise different types of metadata formats, a semi-automatic process is used to collect metadata from the different sources (e.g. files, databases, web services), identify common fields and insert them in the catalogue database.

7 Statistics of the available datasets

Tab. 7-1 summarises the datasets and data products documented in DEIMS-SDR. The table contains information on the completeness of the documentation for the protected area with regard to the recommended fields from the ILTER network. In addition, the number of datasets and data products is provided. The figures encompass all information, which is entered in DEIMS-SDR. Information from other systems (e.g. local metadata catalogues) is not included in this numbers. This information will be provided through the ECOPOTENTIAL Discovery Catalogue provided by WP10 in the runtime of the project.

Tab. 7-1. Overview on in-situ data documented by ECOPOTENTIAL protected areas in DEIMS-SDR

Protected area	Country	LTER metadata profile completeness [%]	Number of available datasets on DEIMS	Number of available data products on DEIMS	SITE-UUID
Kalkalpen National Park	Austria	97	11	12	https://data.lter-europe.net/deims/site/49515dda-1198-4013-8f43-c33e107af081
LTER Zöbelboden	Austria	100	24	9	https://data.lter-europe.net/deims/site/8eda49e9-1f4e-4f3e-b58e-e0bb25dc32a6
Camargue Biosphere Reserve	France	97			https://data.lter-europe.net/deims/site/2104c05c-3b60-46f8-9ebd-da86695f2ede
Cap Corse MPA	France	62			https://data.lter-europe.net/deims/site/dfb77c1a-9e86-4c76-bcbd-8b202ce1f156
Réunion National Park	France	55			https://data.lter-europe.net/deims/site/c8779bc7-0033-4513-8da4-cfde210ffb4d
NP Bayerischer Wald	Germany	93	5		https://data.lter-europe.net/deims/site/993ed2fc-1cb0-4810-a619-8bcf78b6ecee
Samaria National Park	Greece	100	12		https://data.lter-europe.net/deims/site/42fff684-eaf8-455b-a105-6affb609d64d
LTSER Northern Negev	Israel	100			https://data.lter-europe.net/deims/site/871a90b2-e372-456a-93e3-518ad1e11239
Gran Paradiso National Park	Italy	100		2	https://data.lter-europe.net/deims/site/e33c983a-19ad-4f40-a6fd-1210ee0b3a4b
Murgia Alta	Italy	66	41		https://data.lter-europe.net/deims/site/bbd428c6-c4b8-4f47-9298-2b969d38a664
Curonian lagoon biosphere polygon	Lithuania	100			https://data.lter-europe.net/deims/site/dac34969-3d78-4eb7-baec-a7bca4aa40b4
Curonian Spit National Park	Lithuania	100			https://data.lter-europe.net/deims/site/f1f3bc3b-ae91-4730-aa74-529ae097a0e1
Lithuanian Coastal Site (LT-04 Nagliai, Curonian Spit NP)	Lithuania	76	3		https://data.lter-europe.net/deims/site/efdb6a54-a42f-4dc4-8091-dc7bfb937bcf

Protected area	Country	LTER metadata profile completeness [%]	Number of available datasets on DEIMS	Number of available data products on DEIMS	SITE-UUID
Nemunas Delta Regional Park	Lithuania	100			https://data.lter-europe.net/deims/site/e3b03313-adad-4910-a78a-b9a685174c69
Ohrid and Prespa	Macedonia	76	2	1	https://data.lter-europe.net/deims/site/fd82e174-fe05-4cfa-b30a-82aee361d258
LTSER Dutch Wadden Sea Area*	Netherlands	100	See WaLTER catalogue		https://data.lter-europe.net/deims/site/fcbf044b-20e2-4e87-894d-10bfe6e3a7ac
Hardangervidda National Park	Norway	66		9	https://data.lter-europe.net/deims/site/e932fc1f-416e-45a7-836c-3a9ebbbf1baa
Tatra National Park	Poland	83		14	https://data.lter-europe.net/deims/site/2cfd89fe-e561-4551-9ae3-78ad8a8a4d72
LTsER-Montado	Portugal	100			https://data.lter-europe.net/deims/site/ecfcc7e2-82e5-4ec1-adee-a3f9f815fc0d
Montado in Alentejo Natura 2000 sites	Portugal	79		25	https://data.lter-europe.net/deims/site/10e6a0e2-a593-43d0-b915-31ec6e2d840e
Peneda-Gerês	Portugal	66			https://data.lter-europe.net/deims/site/b1b47b83-fac5-4ca9-be17-65e367a70c58
Danube Delta Biosphere Reserve	Romania	100	1	7	https://data.lter-europe.net/deims/site/cdf8ba96-4a30-4eb7-96cc-1c81e8320989
Kruger National Park	South Africa	69	4		https://data.lter-europe.net/deims/site/8fb35fc0-08bc-480a-b779-afd01e89cac2
Doñana Long-Term Socio-ecological Research Platform	Spain	100	17	4	https://data.lter-europe.net/deims/site/bcbc866c-3f4f-47a8-bbbc-0a93df6de7b2
La Palma Island	Spain	72		8	https://data.lter-europe.net/deims/site/e4b06c68-9a26-4b9a-9fe6-dc65a87c0b63
Sierra Nevada / Granada (ES- SNE)	Spain	100	8	9	https://data.lter-europe.net/deims/site/e51cee43-dc12-4545-8e5b-dad35431e3f7
Abisko National Park	Sweden	69			https://data.lter-europe.net/deims/site/f0b67e0d-d3f0-483b-adf0-99ca7bcdcc26
Caribbean LME	Transnational	55			https://data.lter-europe.net/deims/site/77b7caf2-4c92-45c3-8820-f870ba9a7af3
Mediterranean LME	Transnational	62			https://data.lter-europe.net/deims/site/cee1cc87-da10-4726-bcb0-32cded15c31a
Pelagos Sanctuary	Transnational	55			https://data.lter-europe.net/deims/site/0fb65931-5eeb-4645-bf0b-6170adbd1131
Tatra Mountains Biosphere Reserve PL-SK	Transnational	79	See Tatra NP Poland	See Tatra NP Poland	https://data.lter-europe.net/deims/site/c4bbebb0-f5e4-43c3-9cf8-6bb946e7173d
TOTAL			**128**	**100**	

* note: data documented in other systems than DEIMS-SDR are not included in the list (e.g. data from WaLTER)

8 Discussion

8.1 Metadata and Essential Biodiversity Variables

A proper metadata documentation is a key prerequisite in order to make biodiversity data accessible and discoverable. Metadata describes its associated data, and this is particularly important for scientific data analysis and research in general to evaluate if a dataset is suitable for research, e.g. in studies and scientific projects. To monitor biodiversity at various scales, frameworks are needed, and particularly a set of key variables to monitor biodiversity change. The Essential Biodiversity Variables (EBVs) serve as intermediate link between primary observations, indicators and assessment possibilities (Geijzendorffer et al. 2015). In order to generate EBVs, a clear documentation of the data is needed in order to use data and particularly sufficient metadata information needs to be available to evaluate whether data fulfils the basic requirements. EBVs and the underlying data need a sufficient temporal, taxonomic and spatial coverage, as they measure change in an aspect of biodiversity for example in species distribution, phenology, and taxonomic diversity across multiple species and ecosystems (Brummit et al. 2016). There are many datasets available via various platforms and data hosts, such as GBIF, with varying metadata standards as shown in the earlier chapters. For example, GBIF alone provides access to over 783 million observation records (07/2017) that are free to use according to the creative commons principles. However, not all data are suitable for generating EBVs, so metadata needs to give information about the timespan covered by the data, monitoring protocols used, regions that were surveyed and other key data. Projects such as ECOPOTENTIAL and others have an important function in establishing links between data providers and researchers and people from the policy arena that use this kind of data. The projects need to give recommendations on how to generate metadata, for example by using appropriate online tools and repositories such as DEIMS-SDR, EuMon or others and improve existing metadata documenting routines in projects. Another important point to enhance biodiversity data usage is the documentation of the licensing information in the metadata. In order to use data for scientific research or policy reporting, a proper documentation of licensing information is needed. In order to provide open access, a straightforward way is to use creative commons licenses, for example CC0 where no usage restrictions exist or CC-BY and CC-BY-NC that require a attribution (citation) of the data respectively require citation and do not allow a use for commercial purposes.

8.2 DEIMS-SDR and marine sites

The roots of DEIMS-SDR are in terrestrial research sites and datasets. However, the ECOPOTENTIAL project also includes marine sites such as the Caribbean Large Marine Ecosystem (LME) and the Pelagos Sanctuary for Mediterranean Marine Mammals. While, these sites can be described in DEIMS-SDR some mismatches and omissions are apparent. For example, precipitation is generally not recorded nor relevant for marine areas. Also, the currently provided vocabulary for many describing parameters are very limited for marine sites in comparison to terrestrial sites. On top of this, datasets can also originate from cruises and (moving) buoys, which may or may not be part of a research site. Although mobile monitoring is supported by the INSPIRE Environmental Monitoring Facility (EF) Application Schema, this scenario is currently not very well supported by DEIMS-SDR. While the current capabilities of DEIMS-SDR are sufficient for the next steps in the ECOPOTENTIAL project, further improvements to reduce these mismatches will be researched and if possible

implemented in the coming months. These changes could for example include adding parts of the BODC vocabulary, including additional marine physical parameters, habitats and biogeographies.

8.3 Linking metadata for science and research

Traditionally, in-situ studies and research are mostly restricted to single, local, specific data sets. However, when a larger perspective is applied (e.g. at the European scale, or when a series of Protected Areas are compared) data sets have to be comparable to each other. Furthermore, many data have been recorded in the past under very different circumstances. In order to benefit from these, their scope and other important information needs to be made accessible. The fact that meta-analyses and Big Data approaches are rare on the basis of ecological in-situ information illustrates the deficits that must be addressed. Evidently, the usability of existing and recorded data through other projects, researchers and practitioners than the ones that have recorded the data can be improved by informing clearly about the contents of data sets. Such information is needed beyond published data sets (e.g. in supplements to papers) because huge amounts of data are available, but have not made it into international publications or published data in the internet. Metadata are a crucial part of data sets, particularly when series of different data sets need to be combined. The question is, what kind of meta-information should be made available. Here, the fundamental needs must be detected, and redundancy as well as too much complexity must be avoided in order to create a user-friendly and thus efficient access to the relevant knowledge about data sets. Meta data need to be informative and standardized. However, there is no common standard defined up to now. We worked towards addressing these needs, at the same time aiming at simplicity and incorporating existing standards with the ECOPOTENTIAL Discovery Catalogue.

The Questionnaire that was released in ECOPOTENTIAL was aiming to identify the importance of and types of metadata that users are interested in. Respondents to the questionnaire are associated mostly with WP5, WP7, WP6 and WP 4 (59.1%, 50%, 31.8% and 27.3%, respectively). This represents both users and providers of data, and includes ecologists, remote sensing specialists and modelers. Of the respondents, the large majority expressed that they needed metadata for their work (81.8%), and that the absence of metadata was a problem for their work (77%). However, the question about whether in-situ data lacking metadata was useful received the same number of positive as negative replies. This seems to be contradictory, but implies that even though metadata is not critical for all users in all situations, it is considered necessary, and improves the applicability of unknown datasets. Indeed, the current state of metadata for many available datasets is relatively poor, and often requires users to invest extra effort into finding relevant information required to understand and use a dataset. The improvement of metadata is therefore a necessary step to both enable the use of undescribed datasets, but also to make the use of these datasets more efficient.

The types of spatial elements and indicators considered to be the most essential in a spatial metadata profile were the spatial and temporal extent, the spatial scale, typology, method descriptions, temporal quality and file format. These top choices for types of metadata were closely followed by details on taxonomic coverage and sampling descriptions, but access and use constraints, online distribution, and intellectual property rights were also considered very important. It can be difficult to pinpoint which metadata elements are the most essential to include in a metadata profile, since users do not always find the same information useful, depending on their specific interests. However, from the questionnaire, it is clear that although all of the suggested 22 elements were considered quite important, few respondents indicated that they were missing

Co-funded by the European Union

elements (3.8%). This result is promising as it indicates that we have identified the elements most essential to users within the project. What remains is to extend this type of study to other user groups, so that this process can be further standardized and suited to a wider variety of scientists.

We also wish to benefit from previous work and the search for standards in projects and consortia such as EU BON and GEO BON. These links are required as far as biodiversity information is concerned. However, not all metadata and data sets are focused on biodiversity. There is a large diversity of in-situ information that is important for ecosystem functions and services such as soil traits, climatic conditions, vegetation structures, biomass, productivity, and much more, that is not contained in biodiversity data bases. Very close to the aims of ECOPOTENTIAL is the approach of international Long Term Ecosystem Research ILTER. There, a common data base structure has been developed and implemented (DEIMS-SDR), upon which we drew for the development of our catalogue (see chapter 6). However, this approach is not focused on earth observation, but rather on measuring and understanding the ecological processes within ecosystems. Of course there is a link to spatial information, but LTER goes beyond and into much more detail. On one hand, this might give the chance for accurate documentation, on the other; it also poses a risk of over complexity. The particular challenge a project like ECOPOTENTIAL faces is the creation of metadata standards for in-situ data, which can be applied to the different type of data – from biodiversity information to climatic conditions or biomass –, while at the same time facilitating the link to earth observation data. Furthermore, these standards need to be usable for a diverse group of people, such as scientists from different fields, staff from Protected Areas or experts working at the science-policy interface. To ensure the usage of metadata standards, they need to be clearly structured, balancing between the needed amount of detail to ensure adequate documentation and the need for simplicity.

Generally, metadata on local or regional in-situ data sets that are or have been recorded must be linked to science and research in order to gain insights and to reach beyond descriptive monitoring approaches. For this purpose, the access to information needs to be structured and made more efficient. Definitively, the use of data cannot be left to haphazard stumbling upon specific data or extensive data mining comprising many hours of work. We need an overview, but this overview can only be granted if clear criteria that inform about data quality are defined. Creating a clear structure for metadata is the only way to improve the overview on which information is contained in a data set in order to increase the use of a huge variety of data sets in scientific approaches and research projects.

9 References

Braak, K., Robertson, T. (2016) Supporting the publication of sampling-event data through the GBIF IPT. EUBON (European Biodiversity Observation Network) Factsheet http://biodiversity.eubon.eu/web/guest/documents

Brummitt, N., Regan, E.C., Weatherdon, L.V., Martin, C.S., Geijzendorffer, I.R., Rocchini, D., Gavish, Y., Haase, P., Marsh, C.J., Schmeller, D.S. (2016). Taking stock of nature: Essential biodiversity variables explained. Biological Conservation.

Geijzendorffer, I., Regan, E., Pereira, H., Brotons, L., Brummit, N., Haase, P., Martin, C., Mihoub, J.-B., Secades, C., Schmeller, D., Stoll, S., Wetzel, F., Walters, M. (2015). Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective. Journal of Applied Ecology 16, 137-149.

Güntsch, A., Fichtmüller, D., Petersen, M., Hoffmann, J., Holetschek, J., Dröge, G. (2016): Towards next-generation ABCD. TDWG Biodiversity Information Standards Conference Paper.

Holetschek, J., Dröge, G., Güntsch, A., and Berendsohn, W.G. (2012): The ABCD of primary biodiversity data access, Plant Biosystems, 146, 771-779, 2012

Kliment, T., Oggioni, A. (2011) EnvEurope (LTER-Europe) Metadata Specification for Dataset Level. Deliverable A.1.1.2b, EnvEurope project report LIFE08 ENV/IT/000399.

Köljalg, U. (2016) Data mobilisation and curation with PlutoF (https://plutof.ut.ee/). EUBON (European Biodiversity Observation Network) Factsheet http://biodiversity.eubon.eu/web/guest/documents

Lengyel, S., Déri, E., Varga, Z., Horváth, R., Tóthmérész, B., Henry, P.-Y., Kobler, A., Kutnar, L., Babij, V., Seliškar, A., Christia, C., Papastergiadou, E., Gruber, B., Henle, K., (2008). Habitat monitoring in Europe: a description of current practices. *Biodiversity and Conservation* 17, 3327-3339.

Michener, W.K., Allard, S., Budden, A., Cook, R.B., Douglass, K., Frame, M., Kelling, S., Koskela, R., Tenopir, C., Vieglais, D.A., (2012). Participatory design of DataONE—Enabling cyberinfrastructure for the biological and environmental sciences. *Ecological Informatics* 11, 5-15.

Michener, W.K., Brunt, J.W., Helly, J.J., Kirchner, Th.B., Stafford, S.G. (1997) Nongeospatial metadata for the Ecological Sciences. Ecological Applications 7:330-342

Penev, L., (2016) Toolbox for Scholarly Publishing and Dissemination of Biodiversity Data (ARPHA-BioDiv). EUBON (European Biodiversity Observation Network) Factsheet http://biodiversity.eubon.eu/web/guest/documents

Petersen, M., Glöckler, F., Kiessling, W., Döring, M., Fichtmüller, D., Laphakorn, L., Baltruschat, B., Hoffmann, J. (submitted): History and development of ABCDEFG: a data standard for geosciences, submitted.

Robertson, T., Doring, M., Guralnick, R., Bloom, D., Wieczorek, J., Braak, K., Otegui, J., Russell, L., Desmet, P., (2014). The GBIF integrated publishing toolkit: facilitating the efficient publishing of biodiversity data on the internet. PLoS One 9, e102623.

Sahoo S.S., Ngyuen, V., Bodenreider, O., Parikh, P., Minning, T. Sheth, A.P. (2011) A unified framework for managing provenance information in tranlational research. BioInformatics 12:461. https://doi.org/10.1186/1471-2105-12-461

Smirnova L, Mergen P, Groom Q, De Wever A, Penev L, Stoev P, Pe'er I, Runnel V, Camacho A, Vincent T, Agosti D, Arvanitidis C, Bonet F, Saarenmaa H. (2016) Data sharing tools adopted by the European Biodiversity Observation Network Project. Research Ideas and Outcomes 2: e9390. https://doi.org/10.3897/rio.2.e9390

Smirnova L., Mergen P., Groom Q., De Wever A., Penev L., Stoev P., Pe'er I., Runnel V., Camacho A., Vincent T., Saarenmaa H., Agosti D., Arvanitidis C. (2016b) Data sharing tools. Deliverable 2.2 EC-ENV-2012.6.2-2-308454, EUBON - European Biodiversity Observation Network. http://www.eubon.eu/getatt.php?filename=EU%20BON_D2.2_Data%20sharing%20tools_13350.pdf.

Tessarolo G, Ladle R, Rangel T, Hortal J. (2017) Temporal degradation of data limits biodiversity research. Ecol Evol. 2017;00:1–8.

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, et al. (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLOS ONE 7(1): e29715. https://doi.org/10.1371/journal.pone.0029715

9.1 List of tables

Tab. 2-1. INSPIRE Metadata elements (based on http:// INSPIRE.ec.europa.eu/documents/ INSPIRE-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1) 11
Tab. 2-2. Minimum set of DwC terms for resource metadata 13
Tab. 4-1. Mapping of INSPIRE EF properties to the DEIMS-SDR MD elements 53
Tab. 4-2. Description and nature of fields included ISO 19115, INSPIRE/ISO19139, DEIMS-SDR and ISO191158 metadata profiles 79
Tab. 4-3. Examples the data quality elements, sub-elements and measures (according ISO19157:2013) (*** the selection of data quality elements should consider the scope) 84
Tab. 4-4. Proposal the metadata fields/elements for spatial data quality evaluation procedure (task 5.3 and ask 5.5) 85
Tab. 4-5 – Spatial Data Quality Evaluation 87
Tab. 4-6 – Proposal of metadata fields/elements for spatial data quality evaluation according INSPIRE themes (INSPIRE metadata elements concerning data quality) 89
Tab. 4-7. Legend INSPIRE Spatial Data Themes. 91
Tab. 6-1. Overview on in-situ data documented by ECOPOTENTIAL protected areas in DEIMS-SDR 96

9.2 List of figures

Fig. 1-1 The protected areas that participate in ECOPOTENTIAL. 8
Fig. 2-1. Information model for DCAT (see https://www.w3.org/TR/vocab-dcat/) 11
Fig. 2-2. Documentation of the protected area NP Kalkalpen (AT) using SMM in DEIMS-SDR (landing page)16
Fig. 2-3. GEOSS Architecture 19
Fig. 2-4. GEOSS Discovery and Access Broker 20
Fig. 2-5. Architecture of LifeWatch e-Infrastructure 21
Fig. 2-6 INSPIRE Data Discovery Interface (see http://inspire-geoportal.ec.europa.eu/discovery/) 23
Fig. 2-7. Data One Search Interface for obtaining information on Earth and Environmental Data 24
Fig. 2-8. DEIMS-SDR Site and Dataset Catalogue 25
Fig. 2-9. Schematic architecture for the integration of data from the LTER network (reference eLTER H2020 project) 25
Fig. 2-10. eLTER Data Integration Portal (prototype) 26
Fig. 2-11. eLTER Data Integration Portal (prototype) - time series viewer for sensor data 26
Fig. 2-12. Overview over the Darwin Core sample-based dataset (source: Braak and Robertson 2016). 27
Fig. 2-13. Screenshot of the PlutoF welcome page 28
Fig. 2-14. The EuMon database on European habitat and species monitoring (available under http://eumon.ckff.si/monitoring/) 29
Fig. 3-1. Question 1 - Results on the involvement in the different WP (multiple answers possible); total = 22 answers. 33
Fig. 3-2. Question 2 - Results on the importance of metadata; total = 22 answers. 33
Fig. 3-3. Question 3 - Results on relevant metadata elements; total = 22 answers. 34
Fig. 3-4. Question 4 - Results on the effects of missing metadata; total = 22 answers. 34
Fig. 3-5. Question 5 - Results on the importance of metadata on in-situ dataset; total = 22 answers. 35
Fig. 3-6. Question 6 - Results on relevant metadata schemas (multiple answers possible) total = 22 answers. 35
Fig. 4-1. Terms to document an observation or experimentation facility 40
Fig. 4-2. Example of basic documentation of a Protected Area on DEIMS-SDR (see https://data.lter-europe.net/deims/site/49515dda-1198-4013-8f43-c33e107af081) 51
Fig. 4-3. Schema of a hierarchical cascade of environmental monitoring facilities 52
Fig. 4-4. Draft model of elements for the description of data products 55
Fig. 4-5. Relevant spatial elements/indicators to incorporate into a spatial metadata profile in order to support/facilitates data quality assessment and management in scientific collaborative network management. 78
Fig. 5-1. Metadata workflow in DEIMS-SDR 93
Fig. 5-2. ECOPOTENTIAL Catalogue - GetCapabilities request 95

10 Annexes

10.1 Annex 1 List of documented data products in DEIMS-SDR

Documentation on in total 100 data products provided by ECOPOTENTIAL protected areas sites are available on DEIMS-SDR. Details can be viewed via DEIMS-SDR (https://data.lter-europe.net/deims/) following the link provided in the column data product UUID.

Data product title	Data product type	Data range	Data product UUID
Ancient irrigation channels	Water use	01.07.2016	https://data.lter-europe.net/deims/activity/ef8d741f-36ac-4e31-b09c-92ed6a195b1f
Biodiversity Data Project	Biodiversity and species richness (incl. population)	2006-05-15 to 2016-10-15	https://data.lter-europe.net/deims/activity/7cc75937-123f-4260-a42a-dc7dd5ba145b
Biomass herbaceous layer	Herbaceous biomass and growth (e.g. grass)	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/b9ef811b-f17a-405a-9e60-822d2f37c61f
Birds Diversity	Bird	2015-01-01 to 2015-12-31	https://data.lter-europe.net/deims/activity/447394b5-c7fc-439a-9f32-842cb99aa1bd
Burnt area	Fire	1990-01-01 to 2014-12-31	https://data.lter-europe.net/deims/activity/4869cbdf-3a3c-4673-875d-96d5cc06b4fa
Canopy height	Vegetation structure and species cover	1979-01-01 to 2009-01-12	https://data.lter-europe.net/deims/activity/6734988f-419b-47dc-81ea-3a6a8768280a
Cartography	Remote sensing data products	15.09.2016	https://data.lter-europe.net/deims/activity/854109f3-e03f-4a2b-b0fc-bacdcf13cade
Climate	Meteorology	1969-01-01 to 1998-01-01	https://data.lter-europe.net/deims/activity/6434444a-dc0d-47ee-b12a-0d84b85ce356
Corine land Cover	CORINE	01.01.2006	https://data.lter-europe.net/deims/activity/bab1c65e-79dc-484e-884b-0590ce6686b4
Corine Land Cover Change - 1990/2000/2006/2012	CORINE	2006-01-01 to 2006-12-31	https://data.lter-europe.net/deims/activity/b6862b0a-9a5b-4e3e-a47c-dab33bdc6aa8
Crop yield	Provisioning services	01.01.2007	https://data.lter-europe.net/deims/activity/79e0bf2c-376a-48ea-a868-38f1a9ddc6fe
Danube Delta alkalinity, TOC, pH, conductivity, water temperature	Wetland physiscs (temperature, depth, hydroperiod, surface, turbidity)	2000-10-05 to 2014-10-05	https://data.lter-europe.net/deims/activity/d0d0fe0d-f1f6-4616-8d6f-841958bfb743
Danube Delta bird census	Wetlands	2000-10-05 to 2014-10-05	https://data.lter-europe.net/deims/activity/139f7512-cb28-4ed2-a926-ba89a2a77c18
Danube Delta nutrients	Weltand chemistry	2000-10-05 to 2014-10-05	https://data.lter-europe.net/deims/activity/75e1d91e-561b-4f3c-8e99-d12f1a85832c
Danube Delta river water level	Hydrobiology of rivers	1932-10-05 to 2014-10-05	https://data.lter-europe.net/deims/activity/ca9c7842-283f-4fb0-8871-3dd096f8e41b

Data product title	Data product type	Data range	Data product UUID
Danube Delta suspended solids	Riverwater physics (temperature, velocity, depth, suspended solids)	1840-10-04 to 2002-10-04	https://data.lter-europe.net/deims/activity/6673b81c-a2ee-4a63-8428-93f380117d8e
Danube Delta water depth	Lakewater physics (incl. temperature, depth, hydroperiod, turbidity)	1995-10-04 to 2009-10-04	https://data.lter-europe.net/deims/activity/99b4ef31-3f3f-4f41-8528-4037e24810af
Danube Delta Water turbidity	Riverwater physics (temperature, velocity, depth, suspended solids)	2003-10-04 to 2004-10-04	https://data.lter-europe.net/deims/activity/c5573eec-c586-438b-b05b-b41366d60215
Doñana_Groundwater	Groundwater levels	1974-01-01 to 2016-09-30	https://data.lter-europe.net/deims/activity/6038c371-6e8f-4e7d-9e23-d78e79ab5107
Doñana_Herbivore_counts	Mammal	1986-01-01 to 2016-10-11	https://data.lter-europe.net/deims/activity/f3e51a1d-063c-4653-8b65-2d1736f7c234
Doñana_Meteorology	Meteorology	1978-01-01 to 2016-09-29	https://data.lter-europe.net/deims/activity/0b9c0123-4c5c-407c-8cba-f70b9bf80597
Doñana_waterbird_counts	Bird	1973-01-01 to 2016-10-11	https://data.lter-europe.net/deims/activity/50249e7c-756c-4af9-a7a5-56d56525e907
Forest Cover Map - 2000/2006	Other Land cover (please specify)	2000-01-01 to 2006-12-31	https://data.lter-europe.net/deims/activity/efb141e7-a28d-4c75-a152-aeefe3a7b1cd
Geology	Geology and Geomorphology	15.09.2016	https://data.lter-europe.net/deims/activity/bc2c09f9-ae7c-4ec5-bc76-0a57690936a3
Grazing	Provisioning services	01.01.1956	https://data.lter-europe.net/deims/activity/47a5db6e-ab73-4d1b-a617-ce2f13f4736b
Grazing	Provisioning services	2007-01-01 to 2007-12-31	https://data.lter-europe.net/deims/activity/404ecf1a-d2de-4de6-840a-3d0dda4f871c
Hardangervidda (Norway) DEM	Elevation model	30.08.2016	https://data.lter-europe.net/deims/activity/f5b5d666-e7cd-4c73-b031-bfd80ebbc40d
Hardangervidda (Norway) Orthophotos	Airborne Images corrected	23.08.2016	https://data.lter-europe.net/deims/activity/c8e9c31e-20c6-4104-9854-55803cfef0e4
Hardangervidda (Norway) Precipitation	Precipitation	1957-01-01 to 2016-08-01	https://data.lter-europe.net/deims/activity/8e7420a8-f46a-4ebe-b445-6c2d3f3e4aff
Hardangervidda (Norway) Reindeer GPS locations	Focal species	2001-01-01 to 2017-01-06	https://data.lter-europe.net/deims/activity/c67e88ad-84c1-41d2-ba8c-23619585bb14
Hardangervidda (Norway) Reindeer Hunting Data	Hunting data	1979-01-01 to 2016-01-01	https://data.lter-europe.net/deims/activity/103578d4-9fe6-4899-bee6-8e61a5cf665c
Hardangervidda (Norway) Reindeer Population Structure	Focal species	1995-01-01 to 2015-01-01	https://data.lter-europe.net/deims/activity/55149f82-563b-4386-a445-b81e5eb9cba1
Hardangervidda (Norway) Snow data	Snow cover and duration	1958-01-01 to 2016-01-09	https://data.lter-europe.net/deims/activity/cb4d05cd-96e1-4ef1-85e1-210cbe105ee7
Hardangervidda (Norway) Temperature	Air temperature	1957-01-01 to 2016-08-01	https://data.lter-europe.net/deims/activity/646613b6-e71b-4d53-99d0-9c392e913243

Data product title	Data product type	Data range	Data product UUID
Hardangervidda (Norway) Vegetation cover & diversity	Vegetation structure and species cover	2016-07-04 to 2016-08-18	https://data.lter-europe.net/deims/activity/21d441a9-9230-4d83-bcf8-d1eaae6b817a
Hydrology	Hydrology and water budget	15.09.2016	https://data.lter-europe.net/deims/activity/2af0f7cb-9340-42fc-b916-fe6e0fa24da3
Kalkalpen National Park (Austria) - Aerophoto 2009/2010/2013	Original images	2016-01-01 to 2016-09-28	https://data.lter-europe.net/deims/activity/5a18453e-8016-4d02-8bd3-e6a7802b6698
Kalkalpen National Park (Austria) - Bark Beetle	Infestation	2016-01-01 to 2016-09-28	https://data.lter-europe.net/deims/activity/c632b61a-9437-4d77-b5b7-d1c344b40031
Kalkalpen National Park (Austria) - Biotope mapping	Vegetation structure and species cover	2016-01-01 to 2016-09-28	https://data.lter-europe.net/deims/activity/45cd080e-2661-4a71-9648-fd8196501a9b
Kalkalpen National Park (Austria) - Dynamic	Disturbance and desaster events	2016-01-01 to 2016-09-28	https://data.lter-europe.net/deims/activity/217f59ce-9120-4881-b724-bb46035fc380
Kalkalpen National Park (Austria) - Forest	Vegetation structure and species cover	2016-01-01 to 2016-09-26	https://data.lter-europe.net/deims/activity/bc6cc12f-1882-4cad-a0c9-b0979b768144
Kalkalpen National Park (Austria) - Geology and geomorphology	Geology and Geomorphology	26.09.2016	https://data.lter-europe.net/deims/activity/b1d83441-2c49-4d79-9f35-a11aa3f8ecce
Kalkalpen National Park (Austria) - Human infrastructure	Social systems characteristics	26.09.2016	https://data.lter-europe.net/deims/activity/f54a0688-5e1f-4b1a-9bd8-cb2f9133e976
Kalkalpen National Park (Austria) - Land cover	Current land cover and habitats	1990-01-01 to 2016-12-31	https://data.lter-europe.net/deims/activity/932158a9-42a9-47f8-ada4-543a5f3a52d9
Kalkalpen National Park (Austria) - Soil	Soil map	01.01.2004	https://data.lter-europe.net/deims/activity/9f44f6c7-33b6-461e-a0c8-7792e1965394
Kalkalpen National Park (Austria) - Topographie	Topography	2011-01-01 to 2014-12-31	https://data.lter-europe.net/deims/activity/e12f8253-608f-47e6-9125-cf47831b9358
Kalkalpen National Park (Austria) - Water	Running water (River and streams)	26.09.2016	https://data.lter-europe.net/deims/activity/ca3a98c7-82ac-43a5-b72d-b45ce2817bec
Kalkalpen National Park (Austria) - Wild animal management	Vegetation structure and species cover	2016-01-01 to 2016-09-28	https://data.lter-europe.net/deims/activity/d72bb1f7-cd95-4dc2-b3a1-dd7f43ea4c8b
Land use / land cover	Historic land cover and habitats	01.01.1956	https://data.lter-europe.net/deims/activity/47861c47-ad4a-4111-9968-2315b6304d75
Land use / land cover	Current land cover and habitats	2007-01-01 to 2007-12-31	https://data.lter-europe.net/deims/activity/cedb0d21-8588-4803-8338-c41cc51e5e44
Land Use Land Cover	Historic land cover and habitats	15.09.2016	https://data.lter-europe.net/deims/activity/923b1c40-a40d-4353-9bd6-5d63f2d59180
LTER Zöbelboden Austria aerial data	Airborne Images	1995-07-15 to 2016-07-15	https://data.lter-europe.net/deims/activity/8689b125-ee46-4d09-9e46-640f9c5c6eab
LTER Zöbelboden Austria catchment runoff	Hydrology and water budget	1995-01-01 to 2015-12-31	https://data.lter-europe.net/deims/activity/62d588a3-bf27-470a-b999-546a66bb8852

Data product title	Data product type	Data range	Data product UUID
LTER Zöbelboden Austria deposition	Precipitation chemistry	1993-01-01 to 2015-12-31	https://data.lter-europe.net/deims/activity/4efaa2f2-6f4a-4f75-b95c-c3ffb13594a5
LTER Zöbelboden Austria forest disturbance map	Disturbance and desaster events	1993-07-15 to 2014-07-15	https://data.lter-europe.net/deims/activity/a6dcdb32-c654-40d9-b8c7-10ecd057bfc4
LTER Zöbelboden Austria forest structure	Terrestrial systems characteristics	1992-07-15 to 2014-07-15	https://data.lter-europe.net/deims/activity/1f23526e-c156-48c6-b795-9b6b0e0358c4
LTER Zöbelboden Austria land cover and habitats	Land Cover and habitats	1995-07-15 to 2016-07-15	https://data.lter-europe.net/deims/activity/bc7ad5e9-d3cf-46e6-8559-c0bc970b4397
LTER Zöbelboden Austria meteorology	Meteorology	1993-01-01 to 2015-12-31	https://data.lter-europe.net/deims/activity/94016f3b-2e6b-4f95-a759-1b0a40126dcd
LTER Zöbelboden Austria soil, soil water, soil respiration data	Soil	1992-07-15 to 2015-12-31	https://data.lter-europe.net/deims/activity/f63f4d99-785e-4071-83d8-e04fd2aecbba
LTER Zöbelboden Austria vegetation structure and species cover	Vegetation structure and species cover	1993-07-15 to 2014-07-15	https://data.lter-europe.net/deims/activity/44a81d47-bf02-46df-af74-00b11f9fe69c
Meteorological parameters for modelling (Mohid) - Pressure	Air pressure	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/072b7705-e559-4421-b80e-47afc7a9f30b
Meteorological parameters for modelling (Mohid) - Solar radiation	Global radiation	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/18312d84-305a-4f9c-a2a1-8d392e44b9b1
Meteorological parameters for modelling (Mohid) - Temperature	Air temperature	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/23697cd9-994a-4edf-8aaa-7139f779fb7f
Meteorological parameters for modelling (Mohid) - Wind speed	Wind speed and direction	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/f86e57ab-e148-4791-92bf-3090caca6cf3
Ohrid and Prespa - Macedonia - Standing water (Lakes)	Standing water (Lakes)	1998-01-01 to 2016-12-31	https://data.lter-europe.net/deims/activity/53060e05-356e-49ee-8233-5fec7b883b8b
Portuguese Land Cover Map	Historic land cover and habitats	2007-01-01 to 2007-12-31	https://data.lter-europe.net/deims/activity/cd4327f4-a662-451d-aabc-2500cbe7f633
Portuguese Land Cover Map - COS Level 2	Other Land cover (please specify)	2007-01-01 to 2007-12-31	https://data.lter-europe.net/deims/activity/9645cfe4-96ef-4c0f-84fb-4f7f5e261e4f
Protected Area	Nature conservation and management	15.09.2016	https://data.lter-europe.net/deims/activity/1102375e-8a5e-4971-9915-8f3c54eee71e
Root depth	Phenology	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/1f4ef427-d860-47f0-b48a-f207163f0702
Runoff	Runoff and discharge	2007-01-01 to 2007-01-12	https://data.lter-europe.net/deims/activity/c85f4cba-d7dc-482c-a0f7-c53c0ebe2428
Runoff	Runoff and discharge	1956-01-01 to 1956-01-12	https://data.lter-europe.net/deims/activity/773ed596-300f-4378-b80a-bc427ff825d1
Shrub control - area under intervention	Vegetation structure and species cover	2011-01-01 to 2014-12-31	https://data.lter-europe.net/deims/activity/1ff0c248-d79c-4f84-b77c-b115a80e075b

Data product title	Data product type	Data range	Data product UUID
Shrub cover	Vegetation structure and species cover	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/d5deced2-3347-4807-b46c-2ea75d1c158a
Soil characteristics and Carbon fluxes	Soil	2016-08-01 to 2016-09-30	https://data.lter-europe.net/deims/activity/7de3b04f-d012-4153-b371-1f9a427f7947
Soil moisture (surface, 5-10 cm)	Soil moisture	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/d9e87990-c4bb-4d79-89bb-1785a40b5ef6
Soil Organic Matter	Soil	2008-01-01 to 2012-12-31	https://data.lter-europe.net/deims/activity/41c1df09-228e-4af1-9cb9-f5aaccec970b
Soil water content	Soil	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/f9db9967-4ec0-46c4-84c3-ec90b2f8478d
Specific Leaf Area	Phenology	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/58cbedb8-0681-4dae-91be-382d900e17df
Tatra National Park (Poland) - bark beetle infestations	Infestation	30.08.2016	https://data.lter-europe.net/deims/activity/153c505a-444d-4cd3-b810-d72e71ecf74f
Tatra National Park (Poland) - conservation zoning	Governance and planning	30.08.2016	https://data.lter-europe.net/deims/activity/2e016ef1-9dbc-4b6d-aa13-26732656c03a
Tatra National Park (Poland) - forest types	Current land cover and habitats	30.08.2016	https://data.lter-europe.net/deims/activity/5df4f2b2-6f2b-4376-970e-b29fb31bb77e
Tatra National Park (Poland) - human activity (tourism, infrastructure, land mgmnt. etc.)	Economic data	30.08.2016	https://data.lter-europe.net/deims/activity/4c7eadaf-c2fe-466e-a996-aa4aebec1b9d
Tatra National Park (Poland) - hydrographic data	Aquatic systems characteristics	30.08.2016	https://data.lter-europe.net/deims/activity/946c3b13-bb5a-49bd-9432-c1627248138c
Tatra National Park (Poland) - MaB zoning	Governance and planning	30.08.2016	https://data.lter-europe.net/deims/activity/e2a6031f-d3c1-49c4-b324-f274112cebe5
Tatra National Park (Poland) - mountain meadows	historical land use	30.08.2016	https://data.lter-europe.net/deims/activity/c8aeef6f-2234-47fe-bf49-4d01e07a2035
Tatra National Park (Poland) - Natura 2000 habitats	Current land cover and habitats	30.08.2016	https://data.lter-europe.net/deims/activity/8022d5a9-1e22-4d94-b353-ced9fa88adab
Tatra National Park (Poland) - on-going nature monitoring - animals	Focal species	30.08.2016	https://data.lter-europe.net/deims/activity/a4929900-75dc-4f10-a140-821822baa27b
Tatra National Park (Poland) - PLANNED nature monitoring - ANIMALS	Focal species	31.08.2016	https://data.lter-europe.net/deims/activity/95d48f5b-6ec4-47b3-9267-65799c7b69c3
Tatra National Park (Poland) - PLANNED nature monitoring - Natura 2000 habitats	Focal species	30.08.2016	https://data.lter-europe.net/deims/activity/7d4fde96-3310-408b-a7cd-63e8b266fd79
Tatra National Park (Poland) - PLANNED nature monitoring - PLANTS	Focal species	31.08.2016	https://data.lter-europe.net/deims/activity/005f1703-0cc7-4bc3-bb09-3991558a0056
Tatra National Park (Poland) - plant communities	Current land cover and habitats	30.08.2016	https://data.lter-europe.net/deims/activity/f4987dea-f0e0-4e78-920c-6afcafb16b34

Data product title	Data product type	Data range	Data product UUID
Tatra National Park (Poland) - wind damages to forests	Other disturbance	30.08.2016	https://data.lter-europe.net/deims/activity/ebd2a452-4825-4ce1-bcb3-b5c5ae325036
Time series of predicted values (Mohid) - Transpiration	Air humidity	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/89310028-8242-4b39-9a7e-603c3f7b5b2c
Topography	Topography	15.09.2016	https://data.lter-europe.net/deims/activity/8b1a0517-b359-4a65-8d09-c0f155821735
Tree cover	Vegetation structure and species cover	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/d8bacd00-edef-495d-b1bb-b9be58a12455
Vegetation biomass	Herbaceous biomass and growth (e.g. grass)	1979-01-01 to 2009-12-31	https://data.lter-europe.net/deims/activity/b48a0b33-604f-42a0-b409-f103b18e0185
Vegetation Data	Terrestrial systems characteristics	15.09.2016	https://data.lter-europe.net/deims/activity/0e1274a4-04e3-4020-b722-7001dc42dc50
Vegetation water content - herbaceous layer	Plant and foliage chemistry	2016-04-01 to 2016-08-01	https://data.lter-europe.net/deims/activity/79bc20c1-f9fe-4270-aca9-739d21962110
Vegetation water content - tree layer	Plant and foliage chemistry	2016-04-01 to 2016-11-30	https://data.lter-europe.net/deims/activity/e9119ae8-411d-43ac-9e4b-62688bd4ef87
Weather data from meteorological stations Sierra Nevada	Meteorology	2008-01-10 to 2016-07-28	https://data.lter-europe.net/deims/activity/738b29d0-f57c-420c-9880-7cb2a8b724b9

10.2 Annex 2 List of documented datasets in DEIMS-SDR

Documentation on in total 128 datasets provided by ECOPOTENTIAL protected areas sites are available on DEIMS-SDR. Details can be viewed via DEIMS-SDR (https://data.lter-europe.net/deims/) following the link provided in the column dataset UUID.

Dataset title	Date range	Dataset UUID
Agios Ioannis Weather Station	2008-03-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/9eaa13e3-27b1-41e5-96da-a02964c3f993
Alta Murgia Architectural, Archaeological and Landscape Constraints		https://data.lter-europe.net/deims/dataset/8c638766-a1c6-4548-a016-c69f3d8fed98
Alta Murgia Areas of Faunistic Importance		https://data.lter-europe.net/deims/dataset/aa4c65a9-a229-47a5-9347-39dbf37fd693
Alta Murgia boundaries	20.05.2016	https://data.lter-europe.net/deims/dataset/0f2a7e7e-571d-4e89-880e-3416e11c40a7
Alta Murgia geological map		https://data.lter-europe.net/deims/dataset/faf14062-655e-4c1a-96ce-53ddef1eac80
Alta Murgia geomorphologic map		https://data.lter-europe.net/deims/dataset/bf9994b3-2e73-4b42-9bf2-1806de2e0dba
Alta Murgia hydrogeological map		https://data.lter-europe.net/deims/dataset/6d362fd3-b724-4a4a-8482-b7d699648ccf
Alta Murgia interference and criticality		https://data.lter-europe.net/deims/dataset/701c3f96-2ab9-448e-b53d-ff175b2c0c6a
Alta Murgia land use map		https://data.lter-europe.net/deims/dataset/fafa794c-050d-4dda-8905-c9ba76c48272
Alta Murgia landscape units		https://data.lter-europe.net/deims/dataset/1be9e42a-d64b-4195-838c-44e91383982a
Alta Murgia National Park zoning		https://data.lter-europe.net/deims/dataset/0bf986ae-2bee-4e96-a182-b07f3a27c631
Alta Murgia natural habitats of Community interest		https://data.lter-europe.net/deims/dataset/053b7cb0-89f1-40e0-bdf3-33b35d7dff16
Alta Murgia pedological map		https://data.lter-europe.net/deims/dataset/e5c6f2f4-3a4e-4b05-b797-49ea5402b7b5
Alta Murgia physiognomic-structural map		https://data.lter-europe.net/deims/dataset/fe72137e-39f1-4d1e-b05a-a7fccda18a3a
Alta Murgia Point ecosystem units		https://data.lter-europe.net/deims/dataset/8712f49a-a10d-4579-893f-5cd972f0385d
Alta Murgia vegetation		https://data.lter-europe.net/deims/dataset/380b5cf4-47d3-4b5a-93a1-f37b441a2f3c

Dataset title	Date range	Dataset UUID
Alta Murgia Vegetation sensitivity map		https://data.lter-europe.net/deims/dataset/e2feee91-c045-4f27-95c0-718802b2d219
Animal species and other data sets	1989-01-02 to 2006-08-31	https://data.lter-europe.net/deims/dataset/4e6fa885-a412-44fe-8832-bca3fee2291f
Arthropoda monitoring in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/4d3bbc18-1855-4424-ab7c-337904b5de31
Askyfoy Weather Station	2008-03-01 to 2016-07-14	https://data.lter-europe.net/deims/dataset/2a7422c6-15fb-4684-88b9-adc111a700e6
Atlas of areas at desertification risk	2007-06-01 to 2012-01-19	https://data.lter-europe.net/deims/dataset/9f02e850-cafb-46b5-b12f-ecf59d74697c
Basic climate at Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/60d7eda2-864f-11e2-b534-005056ab003f
BIOHOLZ dead wood experiment	2012-07-02 to 2012-08-02	https://data.lter-europe.net/deims/dataset/d98d6108-2b1d-4dff-bcef-042dfd1de4e7
Bird monitoring in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/0249224b-6b09-4889-bd6a-f7c5bb60b793
Breeding habitats for amphibians: phenology and environmental characterization of temporary ponds in Doñana	2004-12-31 to 2014-12-30	https://data.lter-europe.net/deims/dataset/4a077742-05f9-11e5-870c-005056ab003f
Chemistry of Curonian Spit national park	2010-12-31 to 2011-12-31	https://data.lter-europe.net/deims/dataset/8ac9d586-9864-11e2-b534-005056ab003f
Chlorophyll A at Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/d57e7d04-866a-11e2-b534-005056ab003f
Climate Station Waldhauser	1972-01-01 to 2015-12-31	https://data.lter-europe.net/deims/dataset/72b84f1f-633f-4e47-a5d5-0eb8fa16794a
CORINE land cover (2012)	2009-12-22 to 2014-12-28	https://data.lter-europe.net/deims/dataset/7daa6288-c486-4bdf-883a-79e99523a101
Counts of Zerynthia rumina in Doñana	2014-01-27 to 2016-11-08	https://data.lter-europe.net/deims/dataset/d48839bb-4b5b-4427-bcb1-f17ca335407a
Cover of CORINE land cover 2006 at Doñana LTSER Platform (2006)	2005-12-31 to 2006-12-30	https://data.lter-europe.net/deims/dataset/8f0dfd76-872e-11e2-b534-005056ab003f
Digital Land Model (DTM)	2006-08-01 to 2007-03-13	https://data.lter-europe.net/deims/dataset/b3b1c47a-4094-4f1c-ac61-130910e4d977
Direct measurements aboveground grasses biomass at Doñana LTSER Platform (2011-2012)	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/01975cc6-90a2-11e2-b534-005056ab003f
Ecological protection zones	24.06.2015	https://data.lter-europe.net/deims/dataset/91473748-3209-4edc-9076-eb267a33341a
Ecopedological map	2003-01-01 to 2012-01-23	https://data.lter-europe.net/deims/dataset/4bdedc6e-6155-4551-942d-d5d510d3a3d0

Dataset title	Date range	Dataset UUID
Fire risk maps (AIB)	2009-01-09 to 2011-10-12	https://data.lter-europe.net/deims/dataset/46fa436e-32e3-4de5-906a-f73f1ebbbb70
Floristic monitoring in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/fa1cf6e5-d48b-4b69-8fad-6ee390866aa4
Forest condition test	03.12.2013	https://data.lter-europe.net/deims/dataset/a02dd77a-977d-11e2-b534-005056ab003f
Forest ecosystems of Curonian Spit national park	2000-12-31 to 2012-12-31	https://data.lter-europe.net/deims/dataset/e10766ac-9869-11e2-b534-005056ab003f
Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) in Doñana National Park for aquatic and terrestrial environments from MODIS fAPAR product	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/7ca632c8-9608-11e2-b534-005056ab003f
Geodetic networks	2006-08-02 to 2007-03-13	https://data.lter-europe.net/deims/dataset/cefe90da-d049-4e51-b32d-c1bb9b51db90
Geodiversity map of Samaria/White Mountains National Park	2015-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/d3685f14-767c-40bf-8be8-c952e3de0c57
Geolithological map	2009-09-02 to 2010-12-08	https://data.lter-europe.net/deims/dataset/b25acce9-651f-4f2b-9230-206430bf38ff
Geological map	2009-09-02 to 2010-12-08	https://data.lter-europe.net/deims/dataset/d5445222-fb16-43c7-9d02-1d7dcddfe7c3
Global Navigation System (GNSS) Puglia Network	2007-01-01 to 2014-10-24	https://data.lter-europe.net/deims/dataset/1c01392b-f9b0-4e11-b7a3-046a27cf8b5c
Habitat monitoring in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/30d414e3-8351-4864-b54b-3c341234a0ce
Herpetofauna monitoring in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/07fd0ed7-4347-4a85-9cf5-29ab2c27160b
Hydrogeological constraint and Idromorphological disposition planes		https://data.lter-europe.net/deims/dataset/92e8d45d-224f-4848-ae3a-d65b0a9f3afa
Hydrographic grid	2008-08-07 to 2011-05-05	https://data.lter-europe.net/deims/dataset/efc5e946-0246-4d4c-ba4f-413d54f5c4cb
IBA (Important Bird Area)	2011-11-25 to 2017-04-18	https://data.lter-europe.net/deims/dataset/57a395f6-d941-42d9-aa1d-365858063cf5
Internal water map	2007-10-31 to 2011-10-05	https://data.lter-europe.net/deims/dataset/29cedc92-58ed-4b81-9827-c439e01b01ff
Kalkalpen National Park - Combined soil type-depth map	1990-01-01 to 2004-12-31	https://data.lter-europe.net/deims/dataset/915b6ea1-d4e3-4b29-8e0f-33ca90db5718
Kalkalpen National Park - Corine Land Cover 2006	2005-01-01 to 2007-12-31	https://data.lter-europe.net/deims/dataset/fbaff6f6-a378-4df9-8bed-416433149456
Kalkalpen National Park - Corine Land Cover 2012	2011-01-01 to 2012-12-31	https://data.lter-europe.net/deims/dataset/d25d745d-f887-48fd-9df0-80f9b961f3f5

Dataset title	Date range	Dataset UUID
Kalkalpen National Park - Corine Land Cover Changes 2006-2012	2005-01-01 to 2012-12-31	https://data.lter-europe.net/deims/dataset/c79a7d6a-8643-4108-aa6c-5198bd68d9af
Kalkalpen National Park - EUNIS Habitat map	2012-01-01 to 2012-12-31	https://data.lter-europe.net/deims/dataset/ee2912df-6c4a-467f-97df-5427f4f5986b
Kalkalpen National Park - Forest structure (Arial photo interpretation)	01.01.1999	https://data.lter-europe.net/deims/dataset/723fc9a3-d4ee-4409-ad4f-2e8d21555b72
Kalkalpen National Park - Habitat and biotop map	2009-10-09 to 2013-12-31	https://data.lter-europe.net/deims/dataset/f358ffa6-eaa5-42ed-952f-9d41b1d15fdf
Kalkalpen National Park - Road network		https://data.lter-europe.net/deims/dataset/33ea62e4-a7b1-435d-98e7-49d2e31cb3da
Kalkalpen National Park - Running surface waters		https://data.lter-europe.net/deims/dataset/f58b8401-1082-49ad-851c-196a87284c01
Kalkalpen National Park - Standing surface waters		https://data.lter-europe.net/deims/dataset/e0f479dd-a907-44c8-8c0a-ebcdd247fa2e
Kalkalpen National Park - Wetlands		https://data.lter-europe.net/deims/dataset/3acfa10b-daa9-47ce-893f-96e0ead9de11
Lake Ohrid and Lake Prespa biological parameters	1998-01-01 to 2016-12-31	https://data.lter-europe.net/deims/dataset/a63966d8-5d49-49ed-a53c-816e2905d746
Lake Ohrid and Lake Prespa physical and chemical parameters	1998-01-01 to 2016-12-31	https://data.lter-europe.net/deims/dataset/de7c0d6d-c7f9-42b0-a883-23ae73b7588c
Land Use (2011)	2006-08-01 to 2007-03-13	https://data.lter-europe.net/deims/dataset/9d809531-faae-4b51-a04e-4bcd12d65cd2
Landscape Thematic Territorial Urban Planning (PUTT/p)	15.12.2000	https://data.lter-europe.net/deims/dataset/d0e86dcb-af64-4eb5-99c4-ef08e3c21f3f
Landslides catalogue	2010-09-25 to 2010-12-08	https://data.lter-europe.net/deims/dataset/7278f1e1-6fbf-43c8-ac0a-d182fa791d57
Leaf area index (LAI)	2009-04-22 to 2014-12-03	https://data.lter-europe.net/deims/dataset/4670477f-8bdf-4885-93a7-de6806ce089c
Leaf nitrogen (%)	2009-08-13 to 2014-12-03	https://data.lter-europe.net/deims/dataset/4bc1b86b-73eb-4442-8319-cf895d6d7b8d
List of vascular plants in Sierra Nevada	2012-06-04 to 2012-10-10	https://data.lter-europe.net/deims/dataset/9be5a99e-4543-11e3-aeb9-005056ab003f
LTER Zöbelboden, Austria, Air chemistry, 2001-2011	2000-12-31 to 2011-12-30	https://data.lter-europe.net/deims/dataset/75a7f938-7c77-11e3-8832-005056ab003f
LTER Zöbelboden, Austria, Air chemistry, 2012	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/cd1fb6f8-5e57-11e3-aa73-005056ab003f
LTER Zöbelboden, Austria, Epiphytic Lichens, 1993-2010	1993-07-14 to 2010-07-14	https://data.lter-europe.net/deims/dataset/63b2325e-4eca-11e4-a597-005056ab003f

Dataset title	Date range	Dataset UUID
LTER Zöbelboden, Austria, Litterfall chemistry, 1993-2011	1992-12-31 to 2011-12-30	https://data.lter-europe.net/deims/dataset/ebde825e-7c7a-11e3-8832-005056ab003f
LTER Zöbelboden, Austria, Litterfall chemistry, 2012	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/9e3def98-7c7b-11e3-8832-005056ab003f
LTER Zöbelboden, Austria, meteorology daily 1995 to 2016	1995-10-04 to 2016-12-31	https://data.lter-europe.net/deims/dataset/80e801f4-220d-49d1-89d0-85eafa88c6b4
LTER Zöbelboden, Austria, Meteorology, 2001-2011	2000-12-31 to 2011-12-30	https://data.lter-europe.net/deims/dataset/8d2c597c-7c78-11e3-8832-005056ab003f
LTER Zöbelboden, Austria, Meteorology, 2012	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/1ddef9fc-7c79-11e3-8832-005056ab003f
LTER Zöbelboden, Austria, Method documentation - Air meteorology	2017-04-01 to 2017-04-30	https://data.lter-europe.net/deims/dataset/14f88f20-ccd3-4bc7-871c-366ad28af223
LTER Zöbelboden, Austria, Method documentation - Precipitation chemistry	2017-04-01 to 2017-04-30	https://data.lter-europe.net/deims/dataset/0a9bb28d-83f2-487a-ac67-0b39313e16cc
LTER Zöbelboden, Austria, Method documentation - Throughfall chemistry	2017-04-01 to 2017-04-30	https://data.lter-europe.net/deims/dataset/d9b8f5f4-f92f-4e95-ad23-f4951196dec2
LTER Zöbelboden, Austria, Precipitation chemistry, 1993-2012	1993-07-14 to 2012-07-14	https://data.lter-europe.net/deims/dataset/310548be-6e41-11e4-adae-005056ab003f
LTER Zöbelboden, Austria, Precipitation chemistry, 2013	2013-01-01 to 2013-12-31	https://data.lter-europe.net/deims/dataset/1ebafb26-1c04-4f18-9690-aae8378fd04f
LTER Zöbelboden, Austria, Precipitation chemistry, 2014	2014-01-01 to 2014-12-31	https://data.lter-europe.net/deims/dataset/c2f0bc25-0d5b-4b0e-9ad9-56b9dda4402e
LTER Zöbelboden, Austria, Precipitation chemistry, 2015	2015-01-01 to 2015-12-31	https://data.lter-europe.net/deims/dataset/cfc1a860-840a-41f2-a208-cf66ab25dfa4
LTER Zöbelboden, Austria, Reference documentation	2017-04-01 to 2017-04-30	https://data.lter-europe.net/deims/dataset/bee71659-3d19-4cca-81a6-f1dad53a45b6
LTER Zöbelboden, Austria, Soil water chemistry, 1993-2012	1993-08-14 to 2012-12-14	https://data.lter-europe.net/deims/dataset/bc717ac6-5dcf-11e4-a597-005056ab003f
LTER Zöbelboden, Austria, Station Codes	2017-04-01 to 2017-04-30	https://data.lter-europe.net/deims/dataset/956b3718-53fe-4e00-9455-c4132b1195c7
LTER Zöbelboden, Austria, Throughfall chemistry, 2013	2013-01-01 to 2013-12-31	https://data.lter-europe.net/deims/dataset/d099e87a-200c-4707-8172-f03363eb3398
LTER Zöbelboden, Austria, Throughfall chemistry, 2014	2014-01-01 to 2014-12-31	https://data.lter-europe.net/deims/dataset/24114cab-8bdc-472a-9131-291330aa5aad
LTER Zöbelboden, Austria, Throughfall chemistry, 2015	2015-01-01 to 2015-12-31	https://data.lter-europe.net/deims/dataset/5ceeb7a8-db86-444b-aef5-24e81c4f4e2f
LTER Zöbelboden, Austria, Throughfall, 1993-2012	1993-08-14 to 2012-12-14	https://data.lter-europe.net/deims/dataset/bf4bdb26-5387-11e4-a597-005056ab003f

Dataset title	Date range	Dataset UUID
LTER Zöbelboden, Austria, vegetation data 1993 - 2014	1993-07-01 to 2014-08-31	https://data.lter-europe.net/deims/dataset/32721c46-717c-49fb-824f-24fd8cddf159
LTER Zöbelboden, Austria, Vegetation structure and species cover, 1992-2010	1992-07-14 to 2010-07-14	https://data.lter-europe.net/deims/dataset/1d609f60-6f05-11e4-a976-005056ab003f
LTER_EU_RO_001_Danube Delta	1974-12-31 to 1994-11-30	https://data.lter-europe.net/deims/dataset/74b37944-4315-11e3-aeb9-005056ab003f
Macrophytes (relative abundance) at Doñana LTSER Platform (2012)	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/520351de-8728-11e2-b534-005056ab003f
Monitoring of Gastropoda in Samaria National Park	2013-01-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/faa308cf-9bec-4da2-b21b-8aab80c0dc00
Monitoring of Mammals in Samaria National Park		https://data.lter-europe.net/deims/dataset/b81211b8-f1ea-40e9-8c40-11e256c66847
Multitemporal Territorial Indicators	2013-01-01 to 2014-10-24	https://data.lter-europe.net/deims/dataset/193409fa-6ac1-4cdd-a06c-224bcb89dd5a
Nitrate and Phosphate concentrations at Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/ebd28d10-8733-11e2-b534-005056ab003f
Official list of protected areas (EUAP)	10.04.2011	https://data.lter-europe.net/deims/dataset/9f0bb299-a007-4064-82d5-a82f6a8a3ed8
Orthophoto 2011	2011-01-01 to 2011-06-30	https://data.lter-europe.net/deims/dataset/df690d97-17e7-422d-ac58-b318da911310
Physical-chemical characteristics of wetlands at Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/de68cc7a-8715-11e2-b534-005056ab003f
Phytoclimatic map	2009-09-10 to 2010-12-08	https://data.lter-europe.net/deims/dataset/0a7aa1f3-ee59-4f85-acb0-63740cb2c820
Phytoplankton abundance at 2 lagoons in Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/7b002d52-95f4-11e2-b534-005056ab003f
Place names and administrative boundaries	29.06.2017	https://data.lter-europe.net/deims/dataset/65f1b791-1bd1-4c0c-8c48-98c16c4f4473
Railway infrastructure map	2007-10-01 to 2010-12-08	https://data.lter-europe.net/deims/dataset/e62d0606-6016-40d0-b04d-4f1914fe2afc
Red Crayfish Consumer Biomass at Doñana LTSER Platform (2012)	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/5c107b4c-87d8-11e2-b534-005056ab003f
Regional Technical Cartography (CTR)	2006-08-01 to 2007-03-13	https://data.lter-europe.net/deims/dataset/79ed587c-9d14-420b-80cf-0b5ed3e6e912
Samaria Weather Station	2013-03-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/8c71916e-14dd-466b-97ae-77a5698a6b83
SCI, CSZ and SPZ maps	2012-10-01 to 2017-06-01	https://data.lter-europe.net/deims/dataset/3cb1a869-4654-4256-b45b-83bd67d1e5d5

Dataset title	Date range	Dataset UUID
Secondary producers: Fish Biomass at Doñana LTSER Platform (2012)	2011-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/bda2f628-87e7-11e2-b534-005056ab003f
Shore bird census and human uses at the beach of Doñana Natural Space	2011-12-31 to 2013-12-30	https://data.lter-europe.net/deims/dataset/2a0762f2-4630-11e3-aeb9-005056ab003f
Sierra Nevada annual average temperature (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/f197051c-cdfd-11e2-a655-005056ab003f
Sierra Nevada annual precipitation (2011-2012)		https://data.lter-europe.net/deims/dataset/f17d203e-cdfd-11e2-a655-005056ab003f
Sierra Nevada daily snow cover MODIS (2011-2012)		https://data.lter-europe.net/deims/dataset/f12fc4d8-cdfd-11e2-a655-005056ab003f
Sierra Nevada global radiation (2011-2012)		https://data.lter-europe.net/deims/dataset/f163645a-cdfd-11e2-a655-005056ab003f
Sierra Nevada mean month temperature (2011-2012)		https://data.lter-europe.net/deims/dataset/f14a62a2-cdfd-11e2-a655-005056ab003f
Sierra Nevada weekly snow cover MODIS (2011-2012)		https://data.lter-europe.net/deims/dataset/f11561a6-cdfd-11e2-a655-005056ab003f
Soil_moisture_LTER_EU_SP_001_Doñana_LTSER (2008-2016)	2008-06-19 to 2016-10-29	https://data.lter-europe.net/deims/dataset/39bef67f-cd66-4578-af93-f34963ad7bc5
Species List of Doñana National Park (2005)	2004-12-31 to 2005-12-30	https://data.lter-europe.net/deims/dataset/46246986-2fff-11e3-8b78-005056ab003f
Tree cover (%) and biomass (tons/ha)	2008-04-01 to 2012-05-06	https://data.lter-europe.net/deims/dataset/d4098c96-b000-4ba8-bec1-192d384bfa4a
Vascular Plant Coverage at Doñana LTSER Platform (2011-2012)	2010-12-31 to 2012-12-30	https://data.lter-europe.net/deims/dataset/1c7fdfe0-b703-11e2-a655-005056ab003f
Vascular plant coverage in Sierra Nevada	2011-12-31 to 2012-11-27	https://data.lter-europe.net/deims/dataset/6b4f1234-453e-11e3-aeb9-005056ab003f
Wind throw monitoring slopes	1990-08-03 to 2013-07-23	https://data.lter-europe.net/deims/dataset/033b8c90-34d7-463b-93d3-a4df001c5d35
Wind throw monitoring uplands	1992-07-16 to 2010-06-24	https://data.lter-europe.net/deims/dataset/2f5ee83c-83d1-4bb8-88f2-348b97298117
Wind throw monitoring valleys	1988-08-08 to 2008-08-01	https://data.lter-europe.net/deims/dataset/3d4515c2-c6ec-4da0-b82c-4e93826ba5b5
Xyloskalo Weather Station	2008-03-01 to 2017-06-30	https://data.lter-europe.net/deims/dataset/11e4e5e3-b78c-4d17-adc3-7baef094e31e