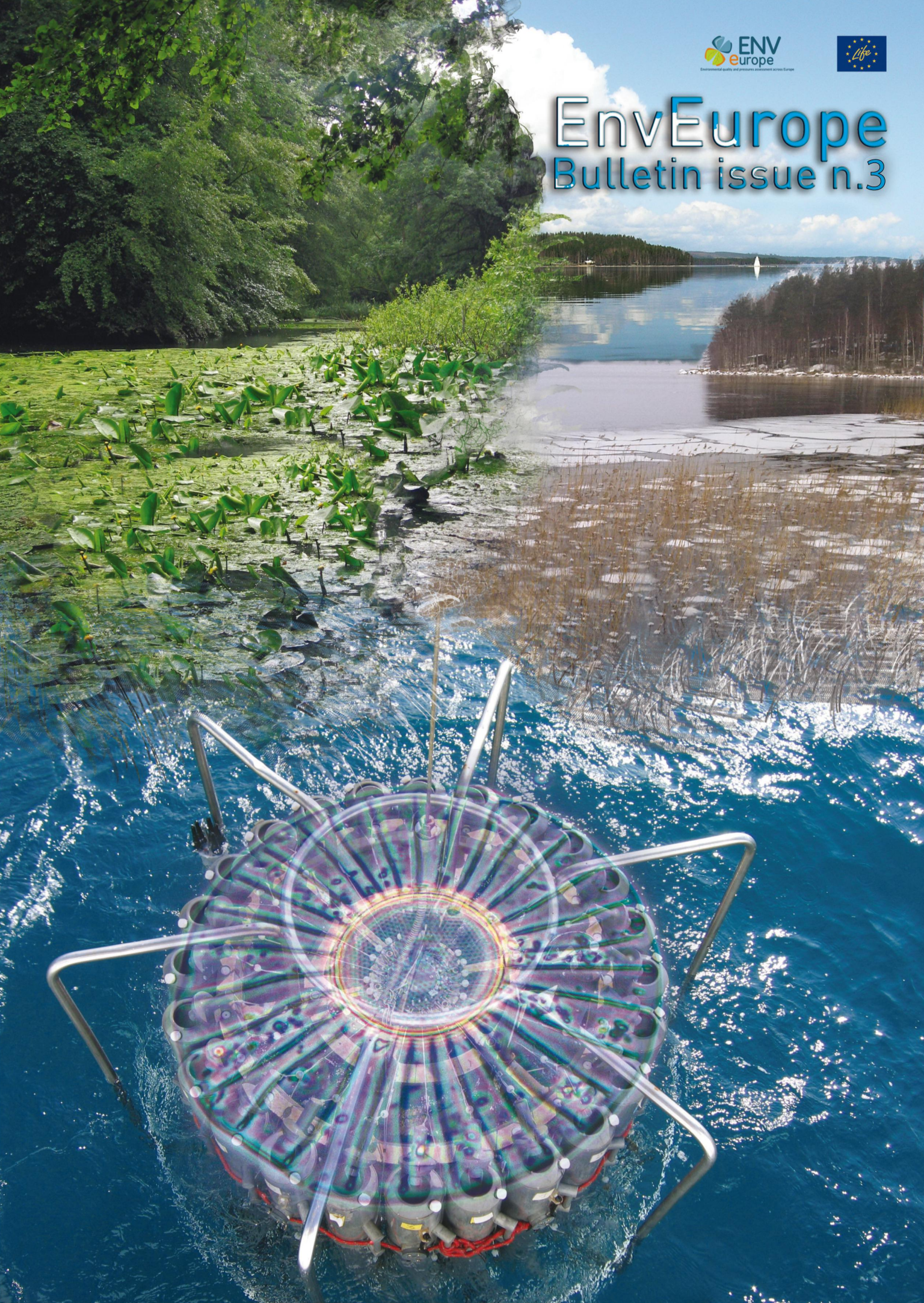
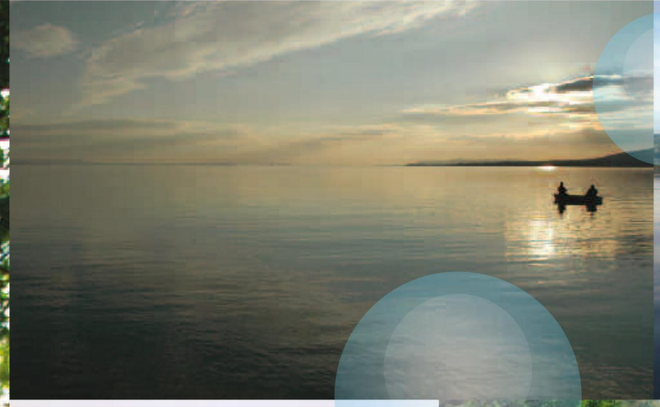
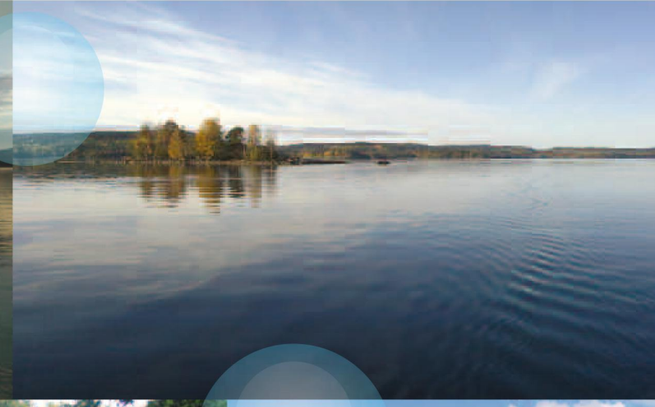
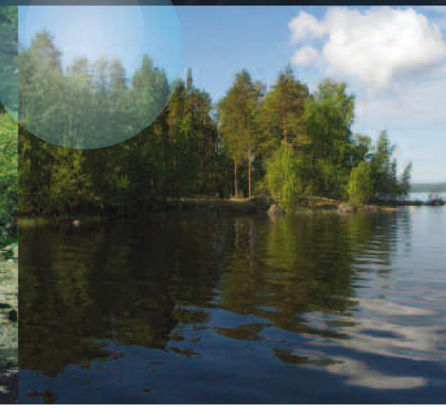

ENV
europe
Environmental quality and pressures assessment across Europe
Life
EnvEurope
Bulletin issue n.3

Contents







Introduction

Analysing trends of environmental changes, identifying drivers and providing options for mitigation is one of the key challenges of long term ecological monitoring and research in Europe. Access to reliable data and harmonised access rights are, therefore, a crucial aspect for this kind of challenges. Nevertheless, we are facing a heterogeneous situation with regard to data management standards and access ranging from data portals to file based data exchange. Different data models, code lists, data policies or data access frameworks, to name just some of the obstacles, are still hampering easy data integration and cross domain and cross site analysis.

The action on management for long term ecological data within the Life+ project EnvEurope (LIFE08 ENV/IT/000399) is aiming to test tools and technologies in order to enhance the accessibility of data from different LTER sites and provide recommendations for site and network level data management. The results in this respect also should provide further inputs for the establishment of a Shared Environmental Information System (SEIS) for Europe.

From Metadata to Data

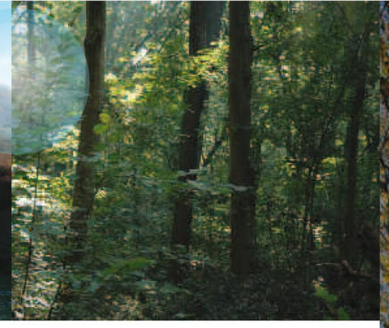
Providing sufficient information about observations is one of the key aspects with regard to usability of data and successful data sharing. Easy ways to discover and interpret data are important not only on the level of networks but also on the institutional level. A catalogue of existing data as well as easy access to it is, therefore, an important task for integrated ecological data management. In this respect not only metadata about datasets are concerned, but also about sites and persons involved. Based on existing standards, user requirements and related legislation community metadata profiles were developed. EnvEurope intends to provide standards and guidelines for the entire LTER-Europe network. Even though, the EnvEurope project is not including the entire about 420 LTER sites it is representative for the whole network.

The dataset level metadata model is based on requirements defined by the target stakeholder groups. The user requirements for the content were harmonized with the INSPIRE metadata regulation and implemented using Ecological Metadata Language (EML) metadata specifications.

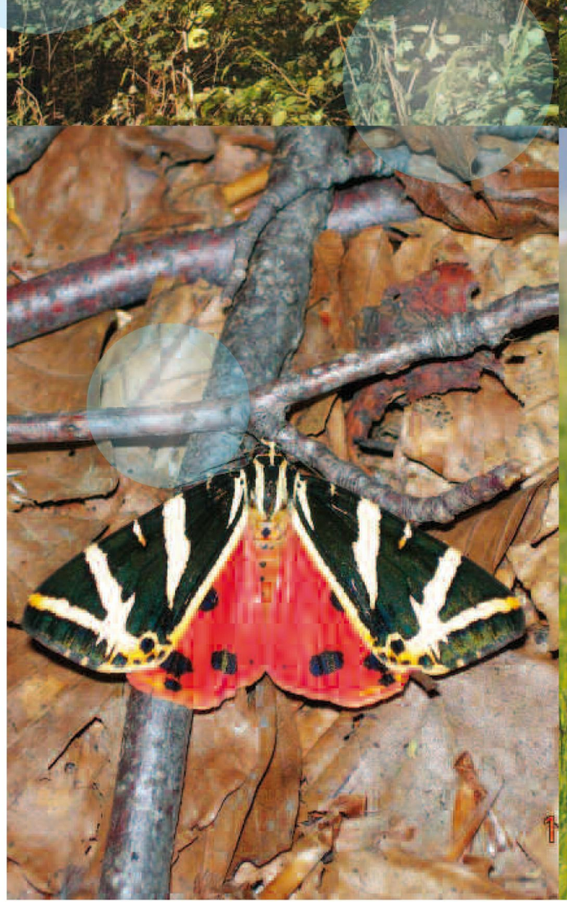
The site level metadata model is based on LTER InfoBase model, developed in the frame of ALTER-Net. New requirements needed for the evaluation of the site network arising within the EnvEurope project were included to provide a consistent and comprehensive description of the site characteristics (see http://www.lter-europe.net/info_manage/infobase-field-description).

The person level metadata model is based on the US LTER model and updated with the user requirements. In order to facilitate the user navigation between related information were all three metadata models automatically interlinked.

A DRUPAL based metadata catalogue (DEIMS) for managing and discovering collected metadata was developed using existing work from the US LTER community. DEIMS (http://data.lter-europe.net/deims/) is one of the core components of the data management within EnvEurope. The metadata specifications aim to ensure feasible interoperability level also with other levels of the international LTER network. Therefore, EML compliant EnvEurope metadata records will be harvested into the LTER-Europe MetaCat, which will be metadata node within the ILTER network. In addition, the LTER-Europe MetaCat server could be a DataOne node (see https://www.dataone.org/) when DataOne becomes operational. INSPIRE discovery service providing INSPIRE and ISO compliant metadata is planned to be implemented in the forthcoming months.




Speaking a common language

Speaking a common language is crucial for exchanging metadata and data. Therefore, within the Life+ project EnvEurope the establishment of a common Controlled Vocabulary was decided. It shall serve as the basis for semantic integration of data resulting from long term ecological research and monitoring in Europe. Based on existing thesauri and other controlled vocabularies, it is extended according to the needs of LTER Europe. EnvThes is still in development but has already been adopted by the ILTER community.

The aims of EnvThes (see http://vocabs.lter-europe.net/EnvThes.html) are manifold. Most important is to provide thematic terms for annotating metadata and data made available via EnvEurope services as well as to support discovery and interpretation of metadata and the related data. Nevertheless, the long term goal is to foster not only structural but semantic interoperability using a common set of concepts describing environmental observations.

EnvThes is based on SKOS/RDF and consists of significant terms and definitions used for long term ecological monitoring, research and experiments. It is based on existing controlled vocabularies using the US LTER Controlled Vocabulary (http://im.lternet.edu/VocabTOR) as backbone. Missing concepts are added using a collaborative approach to provide a comprehensive list of concepts and their definitions. The needs arising from the addressed scientific domains are analysed and existing controlled vocabularies from different domains are searched to avoid re-definition.

Links to controlled vocabularies like GEMET (http://www.eionet. europa.eu/gemet), EARth (http://uta.iia.cnr.it/earth_eng.htm), AgroVoc (http://aims.fao.org/ standards/agrovoc/about), EuroVoc (http://eurovoc. europa.eu/drupal/), EUNIS habitat list (http://eunis.eea.europa.eu/habitats-code-browser.jsp) as well as Wikipedia (http://en.wikipedia.org/wiki/Main_Page) are established and the degree of concept matches (e.g. exact match, close match) is defined, thus creating a web of linked thesauri. Multilingualism is supported by providing translations of the concepts into native languages occurring in the project context.

The direct use of the EnvThes within the project frame is related to the discovery of data and services. Unambiguous and well defined keywords are needed to annotate metadata in order to be able to search consistently for data. In addition to that, the use of common concepts e.g. within data generation enhances semantic interoperability and their usability. EnvThes is therefore an important step towards an integrated network of information resulting from the LTER Europe network as well as the EnvEurope project.

From Data to Analysis

Web service based data access

In order to support distributed data access and sharing reference models and implementation strategies from biodiversity relevant ICT projects on the European level (e.g. LifeWatch, SANY, etc.) and other biodiversity related projects (e.g. EBONE, TERENO, etc.) were analysed for the relevance to the project. The most promising being XML based OGC services as well as the emerging RDF based Linked Data technologies.

For the EnvEurope project a linked data service using D2RQ was established providing online access to data. With a SPARQL plugin for R statistical software the data from distributed resources can directly be used in the analysis or displayed. In addition, SOS server and SWE client provided by 52°North are used to evaluate XML based data services. These services can also be directly consumed by R statistical software or other clients.

By focusing on web service based data provision the basis for distributed data management and access is laid. Metadata are provided in a consistent form using DEIMS.

Data reporting format and centralised database

Based on existing data reporting formats which are used in the domain of environmental monitoring (e.g. ICP Forest, ICP Integrated Monitoring) a generic common reporting format was developed which allows reporting data in a simple way. Flexibility and extensibility together with basic metadata, describing the observation background (e.g. methods), were the main guiding principles in the definition. Data are reported using simple ftp-file storage and being described with its metadata using DEIMS. Reported data are quality checked using simple check routines and imported to central relational data storage. This central database is also used in a first run for testing the different data services.

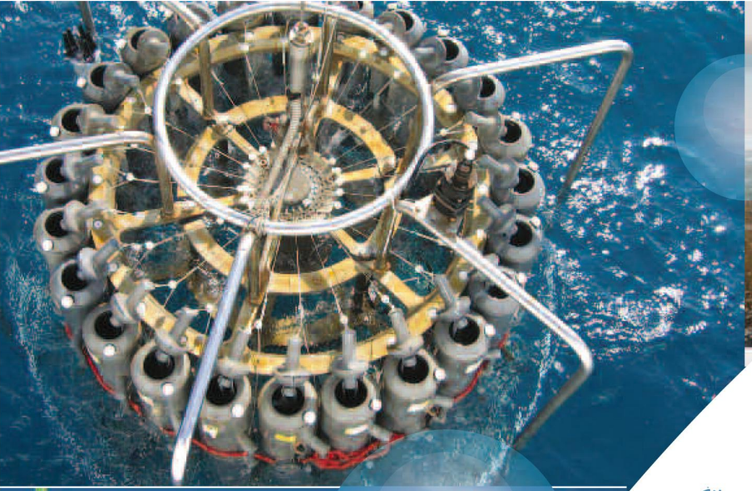
The long term ecological monitoring and research network in Europe is characterised by its heterogeneity with regard to data management. Analysis of the data management practises within the EnvEurope community showed a great variation from file based data storage to well-developed data portals. Nevertheless, the majority of data is stored in simple formats and offline data access is the most common means. EnvEurope in this respect resembles a characteristic subset of the LTER Europe network and provides a good test environment both for data exchange and technical evaluation. In order to provide recommendations for the LTER Europe network a two-fold strategy was applied. First defining a common reporting format in order to facilitate existing data in a structural harmonised way and second the evaluation of data service technologies in order to provide a framework for distributed data access.

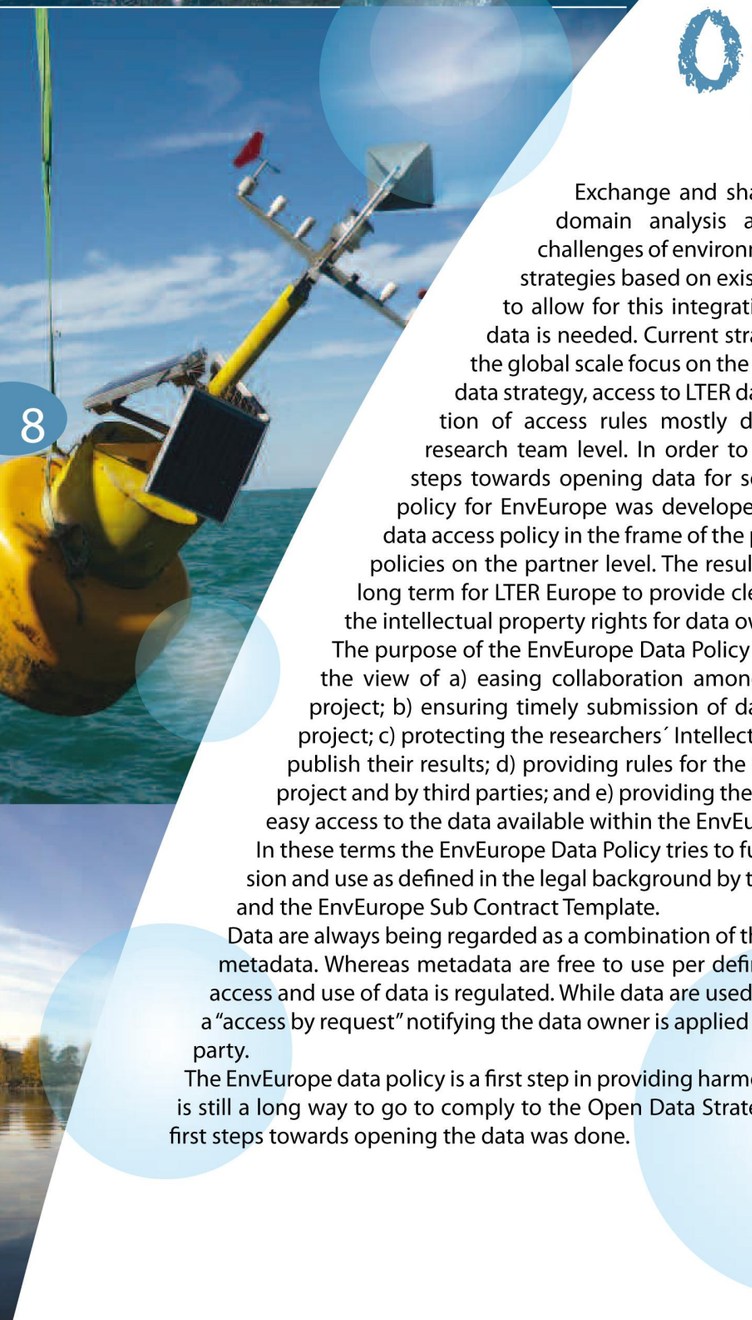
From Analysis to Knowledge

Analysing trends of environmental changes, identifying drivers and providing options for mitigation is one of the key challenges of long term ecological monitoring and research in Europe. Within the EnvEurope project, action 3 – Cause-effect analysis and scientific evaluation – is focused on the analysis of historical and new ecosystem monitoring data coming from a series of LTER sites. It aims at evaluating the status, the trends and the cause-effect relationships at different spatial and temporal scales. The evaluation is taking regional variability and differences between freshwater, marine and terrestrial habitats into account. Bringing together existing long term datasets will be beneficial not only to the network but to each particular site/partner, because it will give a new dimension to the future analysis – the network dimension - and it will allow cross-site comparison. New types of analyses will be started, such as to extend from the site to the regional scale. It will improve the visibility of the network and the LTER sites, and will stimulate new projects directly at the sites.

One of the analysis projects conducted in EnvEurope is focusing on nitrogen deposition and vegetation change in LTER sites. Chronic nitrogen deposition poses a threat to biodiversity as a result of an eutrophication of sensitive ecosystems. Long-term monitoring data from 28 forest sites from the LTER network (including sites from the UNECE ICP Integrated Monitoring and ICP Forests Programme) were used to analyse temporal trends in species cover and diversity. The result showed that the cover of plant species that prefer nutrient-poor soils (oligotrophic species) significantly decreased at sites where the measured N deposition exceeded the empirical critical load for eutrophication effects. Forest plant species that prefer nutrient-rich soils (eutrophic species) showed an opposite trend but only with a marginal significance. Contrary to changes in species cover, species diversity did not correlate with nitrogen critical load exceedance. While the cover of oligotrophic plant species have decreased in European forest ecosystems, diversity is still not affected by airborne N deposition.

This project used data not only from the EnvEurope context but tried to integrate data from other relevant networks using the common data reporting procedures and formats. While being a test case for cross site analysis it also evaluated the achievements for data management.





Opening the Data?

Exchange and sharing of data for cross site and cross domain analysis are important to address the key challenges of environmental change and provide mitigation strategies based on existing data and analysis results. In order to allow for this integrating analysis a clear defined access to data is needed. Current strategies on the European as well as on the global scale focus on the open access of data. Despite the open data strategy, access to LTER data is still hampered by a diverse situation of access rules mostly defined on organisation or even on research team level. In order to overcome this situation and to take steps towards opening data for scientific research and analysis a data policy for EnvEurope was developed. The aim was to define a common data access policy in the frame of the project and to harmonise the different policies on the partner level. The resulting data policy could be used in the long term for LTER Europe to provide clear rules for data users and guarantee the intellectual property rights for data owners.

The purpose of the EnvEurope Data Policy is to set up fundamental principles in the view of a) easing collaboration among the participants of the EnvEurope project; b) ensuring timely submission of data for the use within the EnvEurope project; c) protecting the researchers´ Intellectual Property Rights (IPR) and rights to publish their results; d) providing rules for the use of the data within the EnvEurope project and by third parties; and e) providing the broader scientific community with an easy access to the data available within the EnvEurope project.

In these terms the EnvEurope Data Policy tries to further detail the aspects of data provision and use as defined in the legal background by the EnvEurope Partnership Agreement and the EnvEurope Sub Contract Template.

Data are always being regarded as a combination of the observation data and its describing metadata. Whereas metadata are free to use per definition without any further limitations access and use of data is regulated. While data are used openly within the EnvEurope context a "access by request" notifying the data owner is applied in case the data will be used by a third party.

The EnvEurope data policy is a first step in providing harmonised access rules to LTER data. There is still a long way to go to comply to the Open Data Strategy but within the EnvEurope project first steps towards opening the data was done.

Future steps

Action 1 within the EnvEurope project aims to provide tools and recommendations for ecological data management. In the previous period the foundation was led and need to be completed in 2013. This focuses mainly on the following topics:

Data collection and metadata description

A common data basis is one of the strength of EnvEurope and the LTER Europe network. Historic data for selected parameter groups had already been collected. An update and completion with data collected within action 5 during the field season 2012 will be done. Metadata describing these datasets will be created using DEIMS. In addition metadata update describing persons and research sites involved in the EnvEurope project will be continued. Thus, a good level of information, to be used for further analysis, is created and can be used beyond the runtime of the project.

Tool development and evaluation

The completion of the technical tasks in relation to the metadata framework, data services and data presentation will be done. One of the major tasks is to provide discoverable metadata for the related European (SEIS, INSPIRE) or Global (GEOSS) information infrastructure initiatives with a standardized catalogue service.

EnvThes – the semantic backbone for LTER

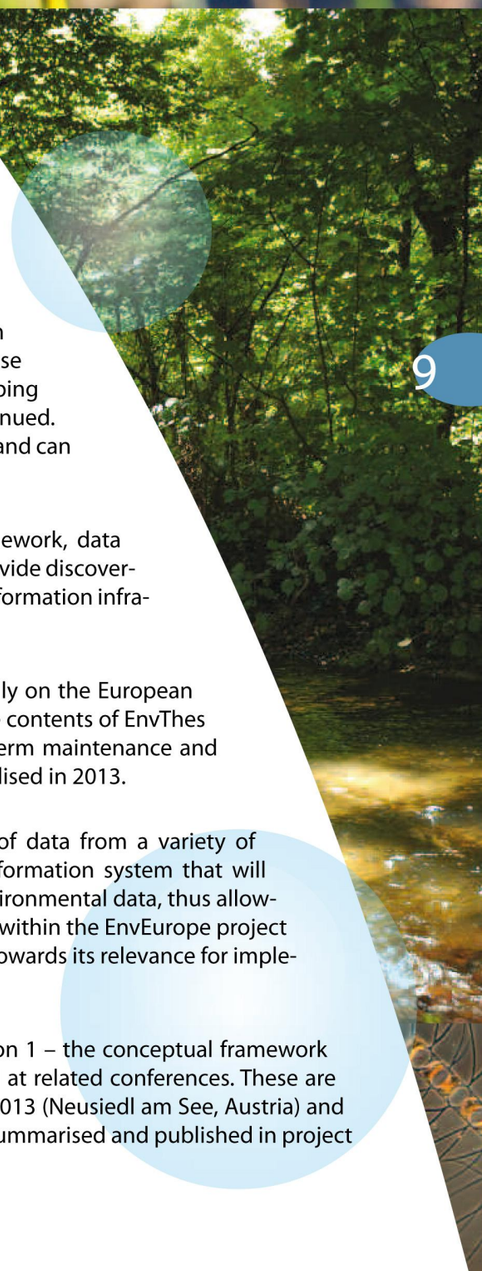
EnvThes will provide the semantic backbone for the LTER network not only on the European but also on the global level. Joint efforts in the further development of the contents of EnvThes together with ILTER and ExpeER will continued. A concept for the long term maintenance and development of EnvThes in the frame of ILTER and LTER Europe will be finalised in 2013.

Shared Environmental Information System (SEIS)

Tackling today's environmental challenges depend on the assessment of data from a variety of sectors and sources. This is why the European Union decided on an information system that will provide decision-makers at all levels (local to European) with real-time environmental data, thus allowing them to make immediate and life-saving decisions. The developments within the EnvEurope project in the field of ecological data management and sharing will be evaluated towards its relevance for implementing a Shared Environmental Information System (SEIS) in Europe.

Dissemination

Dissemination of the results gained from the activities performed in Action 1 – the conceptual framework description and the implementation practice outcomes will be presented at related conferences. These are inter-alia the European EGU 2013 conference (Vienna, Austria), the ISESS 2013 (Neusiedl am See, Austria) and the EnviroInfo conference (Hamburg, Germany). Moreover, results will be summarised and published in project reports as well as peer reviewed papers.



References:

1. Report on the metadata model for datasets
   (PD_A1.1.2b_Kliment_et_al -MD_Specification - Rev.2_Final)

2. Report on the metadata model for sites
   (WD_A1.1.2a_Peterseil_et_al - SiteLevelMetadataModel - Rev.2)

3. Report on data stored in the database
   (PD_A1_D1_4_4_Peterseil_et_al - DataManagementStatus - Rev0.5)

4. Data model for data reporting
   (WD_A1.3.4b_Peterseil_et_al -EnvEurope_DataReporting_
   FieldDescription - Rev.0.4)

5. Data policy
   (Enveuroope_Draft_Data_Policy_V2_3)

Authors:

Johannes Peterseil [1], Herbert Schentz[1], Alessandro Oggioni [2], Tomas Kliment [3], Thomas Dirnböck [1]

with the contributions of

Cristian Mihai Adamescu [4], David Blankman [5]

1 Environment Agency Austria, Austria

2 National Research Council - Institute of Institute for Electromagnetic Sensing of the Environment, Italy

3 National Research Council - Institute of Marine Sciences, Italy

4 University of Bucharest, Romania

5 ILTER Information Management Committee