

1ST Data Management Workshop- LifeWatchGreece RI

Period: June 30th - July 3rd 2014 **Location:** Crete, Heraklion, Gournes, Hellenic Centre for Marine Research (HCMR)

SUMMARY REPORT

Monday June 30th

The first LifeWatchGreece Research Infrastructure (RI) Data Management Workshop was organized between June 30th and July 3rd at the Hellenic Centre of Marine Research in Crete. The meeting was attended by 29 participants who represent several Universities and Institutes from all over Greece. The main aims of the meeting were to invite all Data Managers for a first contact, to describe the kind of data that are available and to understand the kind of problems we have to deal with.

A welcome and introductory speech was held by Dr Christos Arvanitidis who presented **LifewatchGreece** RI, its major aims, objectives and related activities. The participants introduced themselves, presented their data collections and described the way they could contribute their data in LifewatchGreece RI. They also detailed all the problems they have faced from the very first step of data collection up to the last step of data publishing. One of the issues emerged concerned the storage of photos and Quick Response Codes and videos for several species such as bats, ants and dolphins in the Infrastructure. A Google doc was prepared where all participants filled in their problems and questions concerning LifewatchGreece RI, and these were discussed during the last day of the workshop.

Next item on the agenda was the **Data sharing and policy agreement** presented by Dr Sarah Faulwetter. A very fruitful discussion followed, leading to many clarifications and additions. This document, which will be available by the end of July, will provide the data managers, the data providers and the LWG RI users with arguments on why to support the emerging policy of data sharing. The participants supported the idea of producing a brochure pinpointing the major aspects and benefits of data sharing on LifewatchGreece RI. As an outcome of the discussions it was decided that each data manager will keep a diary regarding their work, which will be uploaded at the data collection forum- LifewatchGreece RI portal. This diary will sum up the discussions with the potential data providers, as well as any related problems and suggestions, in order to have a clear picture on what actions may be needed to be followed in the future.

The last presentation of the day by Niki Keklikoglou was about the Biodiversity Data Journal (BDJ), which is a community peer-reviewed, open-access, comprehensive online platform, designed to accelerate publishing, dissemination and sharing of biodiversity-related data of any kind.

Tuesday July 1st

The first presentation of the second day by Dr Sarah Faulwetter concerned **Taxonomy and Nomenclature**. Rules and common problems concerning synonyms, homonyms, authority and ranking of taxonomy groups were discussed.

Dr Vaggelis Pafillis presented the subject of **genetic metadata**. He referred to the Ocean Sampling Day and the protocol used during this event for data capturing. He also described the "Environments" project, an application capable of identifying environment descriptive terms, and noted that the Encyclopedia of Life can act as an enhancement to "Environments" aiming to process the EOL Taxon pages to extract descriptions of their environmental context.

The next presentation was based on **spatial data** and was presented by Mrs Matina Nikolopoulou. Specifically the presentation concerned spatial data management and display, differences between geographical and projected coordinate systems, accuracy and precision on longitude and latitude measurements, common mistakes and solutions on spatial data management. Mrs Nikolopoulou continued on geoweb services and applications, which are used in order to control, check, store and publish spatial data.

Wednesday July 2nd

Mrs Dimitra Mavraki covered the topic of Data Management and quality control of raw data and metadata. She described all the web services that can be used in order to clean and check data, mainly focused on marine data. Data quality control (QC) procedures are important for detecting missing mandatory information, errors made during the transfer or reformatting of data, duplicates and remaining outliers (out of scale data, vertical instabilities etc.). In addition she distributed a Quality Control Manual for Marine Data and asked the participants to prepare a similar manual for their own data that might differentiate from marine data in many aspects (e.g. museum collections).

A training session followed, during which a number of exercises were distributed. Each participant was able to try on his own laptop how to fill in the metadata document, in order to better understand the main problems on cleaning raw data, as well as to get more familiar with using web applications for quality control.

The next item on the agenda was a presentation by Dr Sara Faulwetter concerning **data structuring**, during which she described the ways and formats that data managers can use to collect and store data for sufficient interpretation and analysis.

Thursday July 3rd

On the last day, Mrs Dimitra Mavraki referred to The **Darwin Core standards**. These include a glossary of terms (that might be called properties, elements, fields, columns, attributes, or concepts) intending to facilitate the sharing of information about biological diversity by providing reference definitions, examples, and commentaries.

The last session of the workshop included a general discussion on the issues and solutions that the participants had filled in the Google doc concerning LifewatchGreece RI. All PowerPoint presentations will be made available for forum members at the LifeWatchGreece-Data collection workspace: <http://www.lifewatchgreece.eu/?q=forums/data-collection-dc>.