**TRAINING SESSION IN THE FRAMEWORKF OF THE**

**EMODNET BIOLOGY PHASE IV**

 **ON THE FEASIBILITY STUDY FOR RECOGNITION OF SPECIFIC ECOLOGICAL TRAITS AND SAMPLING DEVICES/METHODOLOGIES IN TEXT**

**Intoduction**

The European Marine Observation and Data Network (EMODnet) is a network of organisations supported by the EU’s integrated maritime policy. These organisations work together to observe the sea, process the data according to international standards and make that information freely available as interoperable data layers and data products. EMODnet provides access to European marine data across seven discipline-based themes: [Bathymetry](https://emodnet.ec.europa.eu/bathymetry), **Biology**, Chemistry, Geology, Human Activities, Physics and Seabed Habitats.

**EMODnet Biology** is the EU service for European in-situ marine biodiversity data. It was established in 2009, following the publication of the Maritime Policy Blue Book in 2007. Its main objectives include the curation, quality control and publication of marine biodiversity data in European marine regions and the Caribbean seas where European countries have overseas territories or by European scientists and/or institutions outside of the European area. EMODnet Biology focuses not only on recent but also historical data, i.e. data that have been collected over 50 years ago. Part of the work is also dedicated to creating interoperable and reproducible products based on open data and engaging with the regional sea conventions, EU and international reporting mechanisms.

**Overview**

The training session is organised in the framework of EMODnet Biology and in particular on the ongoing Feasibility Study on Ecological Traits and sampling devices/methodologies identification with text mining. This training session focuses on the interdisciplinary field of text mining and its application on marine ecological traits. During the session, participants will receive an overview of state-of-the-art text mining technologies and bioinformatic tools. A balanced approach of lectures and practical sessions will provide a general background and the latest advancements in the field as well as its relationship with other disciplines.

The session will start with training the participants about literature curation workflows, standards and resources. Participants will prepare their own literature and texts and follow the workflow to extract information on traits and species names. Using machine learning algorithms and tools they will evaluate the results and provide feedback on user interface, usability and feature requests. Thus, the participants have a unique opportunity to work on the complete workflow from literature preparation to data analysis. This session will also address the annotation of texts to improve the current models. Programming is a key part of the session, during which an introduction on how to process data from raw text to the application of algorithms for text mining, will take place. Trainers will provide short lectures with follow-up hands-on exercises in order to have the participants learning and assessing their knowledge.

**Audience**

This session is aimed at marine ecologists, curators, data managers, students who are doing or planning to start research on curation and information extraction of marine ecological literature. Prior knowledge on programming is recommended.

**Agenda**

**9:00 – 9:15 Welcome. Introduction to EMODnet Biology.**

(Mrs Dimitra MAVRAKI, Data Manager, Hellenic Centre for Marine Research, Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC))

**9:15 - 10:00 Overview of pre-course work, introduction round.**

(Mr Savvas PARAGKAMIAN, PhD Candidate, Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC))

**10:00 - 10:30 Introduction to Text Mining for Ecology.** Definitions, applications, workflow, available tools GUI and CLI, ecological entities, towards trait definitions

(Ms Lydia KALAITZAKI, Operations Manager, Advance Services)

**10:30 - 11:00 A presentation of the examined tools and challenges faced.**

(Ms Lydia KALAITZAKI, Operations Manager, Advance Services)

**11:00 – 12:00 A demonstration of the performance of the desirable text mining models/ Hands-on Training**

(Ms Melina LOULAKAKI, Software Engineer, Advance Services & Christos Papacharalampous, Data Engineer, Advance Services)

**12.00 - 12.30 Advice and further development plan (answers and questions session).**

(Ms Lydia KALAITZAKI, Operations Manager, Advance Services)

**Practical Information**

**Date: 20 March 2023**

**Time Zone:** **Central European Time**(CET)

**Location: online training** (the link will be sent to participants after their express of interest at *info@advancesvs.com*)

**The training session is free of charge for participants.**