

# European Marine Omics Biodiversity Network (EMO BON)

## *EMO BON metadata*

The collection of rich and rigorous metadata to accompany the data is high priority for EMO BON. Metadata are the key that places the data in a broader context and ensures the correct data FAIRification. To collect the metadata, we have prepared detailed metadata log-sheets that should be completed by the contact people of each EMO BON station. The log-sheets are prepared as online spreadsheet files, in the EMBRC google account. Each file corresponds to one EMO BON observatory. In addition, there are separate files for the water column and the soft substrates samplings for each observatory.

You will receive an invitation by email to access the file(s) that corresponds to your observatory station. We ask that you complete the metadata spreadsheets as soon as possible after the sampling event. This will mark that the sampling took place, that the samples are collected and will create the list of samples that belong to your observatory. The metadata spreadsheet file includes 4 sheets:

### **1) READ\_THIS\_FIRST**

This sheet includes brief instructions on what the other sheets contain. It also asks that the structure of the file is not changed when completing the metadata. Keeping the structure of the file as is, is of great importance. These files are going to be used for uploading the metadata to public databases and if their structure changes, the upload will not be possible. This means that it is not acceptable to rename the columns and to add or remove columns.

### **2) observatory**

EMO BON metadata are separated in categories as described in the EMO BON Handbook. This sheet contains the Observatory Metadata that are all information regarding the observatory station, the sampling sites, their characteristics, and the contact people. This sheet should be completed by the stations after registration to EMO BON. Each row is a different ARMS unit ("ARMS\_units\_id") and each column represents a different metadata field.

### **3) sampling**

EMO BON metadata are separated in categories as described in the EMO BON Handbook. This sheet contains the Sampling Metadata which include all the information related to the specific sampling event and the fate of the samples. For example, information on the date and time of sampling, sample handling details and deviations from protocols are included here. This sheet should be completed by the stations after every sampling event. Each column represents a different metadata field. Each row corresponds to one sample. The field "source\_mat\_id" is where the sample label should be added. This is the unique identification for each sample collected and will follow the sample to the data generation.



#### 4) definitions

This sheet contains the definition of each metadata field included in the sheet's "observatory" and "sampling". The metadata fields are listed in alphabetical order. For each metadata field, it is given: its full name, a detailed definition, the value and units that should be used, an example, which metadata category it belongs to, if it is mandatory or optional and if it is mandatory in other checklists.

Common mistakes: There are certain metadata fields which we have noticed are commonly filled in incorrectly, so please read the following to make sure you don't make the same mistakes:

- All numbers should be entered with a dot and no commas: 10002.345 for example
- It is vital that the sample IDs are correct, as these are the key identifiers that tie all EMO BON data together. For example, the identifier EMOBON\_VH1\_Ha\_210608\_220608\_SF40\_1 is the first replicate for the sample SF40 of that event, EMOBON\_VH1\_Ha\_210608\_220608\_blank is the blank sample for that event, etc. Please do not forget the \_ characters.
- The column "images\_in\_PlutoF" should be filled with a Y when you have uploaded your images to PlutoF together with the image description spreadsheet, N if you have not yet done that but will, and a None if the images will never be uploaded to PlutoF.
- Remember that colours and comments are not understood when the logsheet is harvested into GitHub and added to the EMO BON dataset. If any vital information needs to be part of the logsheet, they need to be added as normal characters in the cell (e.g. if you will not provide a mandatory measurement, use N/A rather than colouring the cell in red).
- Dates are YYYY-MM-DD; you do not also need to provide the time.
- Do not use the comma in comments, as this character is used to divide cell values from each other when the logsheet is downloaded as CSV. Use ";" instead.

The contact people of the observatories may contact the EMO BON Science Officer any time for questions and/or assistance when filling in the metadata. Samples for which metadata are incomplete or improperly documented will be rejected and will not be further processed.

**Ioulia Santi**

**European Marine Biological  
Resource Centre**

Paris Headquarters

4 Place Jussieu,

Tour 46/00, 1er étage, bureau 101

75252 Paris Cedex 05 (FR)

[www.embrc.eu](http://www.embrc.eu)

Email: [ioulia.santi@embrc.eu](mailto:ioulia.santi@embrc.eu)

Phone: +30 2 81 03 37 727

Mobile: +30 6 94 51 78 545