13.12.2019

# BALTIC GENDER DATA MANAGEMENT PLAN

Version 3

**Project Name:** Baltic Consortium on Promoting Gender Equality in Marine Research Organisations
**Project Identifier:** Baltic Gender
**Grant Title:** 710363
**Principal Investigator / Researcher:** Katja Matthes
**Project Data Contact:** Hela Mehrtens, hmehrtens@geomar.de
**Description:** My plan (Horizon 2020 DMP) - Baltic Gender DMP
**Funder:** European Commission (Horizon 2020)

## 1. DATA SUMMARY

*Provide a summary of the data addressing the following issues:*
*\* State the purpose of the data collection/generation*
*\* Explain the relation to the objectives of the project*
*\* Specify the types and formats of data generated/collected*
*\* Specify if existing data is being re-used (if any)*
*\* Specify the origin of the data*
*\* State the expected size of the data (if known)*
*\* Outline the data utility: to whom will it be useful*

**1.1 Aggregated and anonymous gender segregated data** are collected from partner institutes (as part of the work undertaken in WPs 1-4 towards realizing a database of indicators, D6.1). The respective human resources departments in the partner organizations deliver the data, which are completely anonymized. The objectives are manifold: (i) to assess the status of gender equality in the partner institutions; (ii) to develop gender-sensitive indicators as long-term monitoring tools; (iii) to produce possible targets and innovative strategies toward Gender Equality Plans (GEPs). The data are collected in spreadsheets. Existing statistics of gender studies are re-

used, e.g. from She figures. The indicators, which are calculated from the collected data, comprise:

- 7 quantitative indicators: scissor diagram for academic staff, glass ceiling index, gender pay gap, part-time employment, sex of the chief scientist on scientific cruises, women's representation in committees, percentage of women in the recruitment process
- 6 qualitative indicators: flexible working arrangements, child care service, incorporation of gender analysis in research design and management, incorporation of GEPs in research project plans and implementation, recommendations on gender sensitive language, recommendations on gender sensitive didactics

The size of the data is less than one megabyte.

**1.2** Data are collected in the form of **interviews**, **questionnaires, focus groups, case studies and online surveys** (hereafter named as studies) from project partners. The purpose of the studies is to collect information that is otherwise not available in a written document at the departmental / institutional level. The studies help to: (i) gather information about practices, processes, procedures relevant for achieving gender-equality in the partner institutions; (ii) evaluate the value of a practice or scheme by gathering personal experiences; (iii) showcase personal experiences of staff on career advancement and work-family balance in the form of a blog. The data collected comprise the following:

- Interviews are carried out with female scientists/marine engineers from partner organisations. Resultant stories are published as part of a blog on "portrays of women in marine science and technology" (WP1, Task 1.1, D1.1).
- Interviews are carried out with female or male staff members from partner organisations. Resultant stories are published as part of a blog on "experience of families sharing child-care" (WP2, Task 2.1.2, D2.2).
- A survey is carried out with staff members, who had career breaks (WP2; Task 2.2). The survey results are used to produce a check-list (D2.3) for best practices within the consortium regarding maintaining contact with individuals taking family breaks or returning to work as well as best measures for the persons-on-leave to keep up with colleagues and scientific developments in their field of research.

- Focus group discussions, consisting of approximately five persons from the staff of one partner institution in a single session, are carried out on practices, gaps and challenges with respect to actions taken to catalyse structural change (WP3; Task 3.1 lead by CAU). The results are used towards compiling a report on best practice examples for catalysing structural changes (D3.1).

- Case studies focusing on specific research project or task are carried out to test the draft method protocol for incorporation of gender aspects in marine research (WP4, Task 4.1).

- An online questionnaire is sent out to each participant before and after the training sessions for staff (WP5; Task 5.1 lead by IOW) to understand the value of the trainings. A report (D5.1, month 39) on the value of the trainings is prepared based on these surveys.

- The outcome of the mentoring scheme (WP5; Task 5.2) is evaluated by all involved parties based on questionnaires and mentee's reports, which are used to deliver a synthesis report (D5.2).

- Interviews and questionnaires are used for the formative and summative assessment of the GEP implementation (WP8; Task 8.4; D8.3&D8.4).

- An online survey is circulated to scientific staff in the partner institutions to understand how Marine Science research operates (WP 3, Task 3.2; WP 6, D6.1).

## 2. FAIR DATA

**2.1 Making data findable, including provisions for metadata:**
*\* Outline the discoverability of data (metadata provision)*
*\* Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?*
*\* Outline naming conventions used*
*\* Outline the approach towards search keyword*
*\* Outline the approach for clear versioning*
*\* Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how*

The deliverables are discoverable from the institutional repository of GEOMAR, "OceanRep", where metadata are OpenAire compatible and are harvested by several distributors, e.g. BASE. DOIs are given to reports and the final indicator dataset. Reports and data spreadsheets have names including the title, author's name and the publication date. Search keywords are provided by the authors. The repository offers a versioning system. The metadata system used is Dublin Core.

The deliverables are uploaded on the GenPORT, a community sourced internet portal for sharing knowledge on gender and science. The contributions uploaded on GenPORT include a link to OceanRep and keywords.

**2.2 Making data openly accessible:**
*\* Specify which data will be made openly available? If some data is kept closed provide rationale for doing so*
*\* Specify how the data will be made available*
*\* Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?*
*\* Specify where the data and associated metadata, documentation and code are deposited*
*\* Specify how access will be provided in case there are any restrictions*

Aggregated and anonymous gender segregated data (**Section 1.1**) received from the human resources departments are only available to the project partners and are subject to data-screening by the Coordinator and respective work-package leaders. Only the chosen indicators from this data set are made publicly available.

The data collected from offline and online studies (**Section 1.2**) are not publicly available as they may contain sensitive information. However, the deliverables generated from these studies are made publicly available. The deliverables generated from the studies are in the form of user-friendly outputs such as a blog, list of best-practices, recommendations, check-list, reports.

The deliverables (including the indicator database) are openly accessible from the repository OceanRep, the website of the Baltic Gender Project (http://www.baltic-

gender.eu) and the GenPORT portal (https://www.genderportal.eu). The deliverables are also electronically distributed to the relevant divisions of staff as well as regional networks and alliances of partner institutions. The project outputs are presented to all disciplines of natural sciences through conferences, workshops and other events in Europe (General Assembly of European Geosciences Union, European Maritime Day, Helmholtz Sustainability Summit and many more).

**Table 2.1. Deliverables of Baltic Gender**

| |
|---|
| Baltic Gender Action brochure |
| Blog on portrays of women in science |
| Blog for experiences on families sharing child-care |
| Guideline for GEPs for marine research proposals/projects |
| Newsletter for the mentoring scheme and information on women's scientific networks |
| Report on the status of the existing GEPs from partner organisations |
| Project outputs for "Insight into women in Marine Science and Technology" |
| Practices and recommendations on family friendly strategies |
| Data base for indicators |
| Review of national legislations on gender equality in S&T |
| Best-practices /recommendations for decision making- processes, recruitment procedures and allocation of resources |
| Report on promotion project |
| Check-list for maintaining contact with individuals taking career breaks |
| Report on surveys for staff training on gender competence |
| Best-practices in career-building measures |
| Report on the activities of the established grass-root networks |
| Brochure& training material on gender sensitive teaching methods / content |
| A synthesis report for the mentoring programme |
| Report on "Marine Gendered Innovations" |
| Final report, including a systematic synthesis of the gained lessons and a summative evaluation of the status of the GEPs |

**2.3 Making data interoperable:**

*\* Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.*

*\* Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?*

The data and metadata are taken in a common format by the project partners to allow comparability in the project. Where possible, internationally accepted standards/formats are used for the gender-sensitive indicators. We adopt to metadata and terms from earlier studies in this field (e.g., She Figures), to make data interoperable.

**2.4 Increase data re-use (through clarifying licenses):**
*\* Specify how the data will be licenced to permit the widest reuse possible*
*\* Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed*
*\* Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why*
*\* Describe data quality assurance processes*
*\* Specify the length of time for which the data will remain re-usable*

Deliverables are licensed as CC-BY to allow reuse of data by citing the reference. As the data and reports are made accessible through an institutional repository and international data centers, the length of time for which the data will remain re-usable is not limited to the existence of the Baltic Gender webpages.

## 3.  ALLOCATION OF RESOURCES

**Explain the allocation of resources, addressing the following issues:**
*\* Estimate the costs for making your data FAIR. Describe how you intend to cover these costs*
*\* Clearly identify responsibilities for data management in your project*
*\* Describe costs and potential value of long-term preservation*

The costs of data management are mainly personal costs, which have been applied for in the project. The projects is allowed to use the infrastructure of GEOMAR for repository and website because the coordination is situated at GEOMAR.

Data curation is carried out by the coordination team in close cooperation with the in-house data management team using an existing portal (https://portal.geomar.de) for data exchange and storage. There is a regular backup. Deliverables are communicated to interested network partners and uploaded to the EU portal and also harvested by Openaire to be found and utilized after the end of the project.

The consortium will raise funds to make sure that the project website with all public deliverables stays accessible and up-to-date (until at least 4 years after the duration of the action).

## 4. DATA SECURITY

***Address data recovery as well as secure storage and transfer of sensitive data***

All **aggregated and anonymous gender segregated data** (**Section 1.1**) received from the respective HR departments are anonymous and do not contain any sensitive information. Secure storage at the internal data storage of the project is provided according to safety rules of the scientific data centre at GEOMAR. The data can only be accessed with personal login.

Any **personal data** (e.g., name, and contact details) collected as part of the studies (**Section 1.2**) are on a separate cover page. These cover pages are stored safely in a locked cabinet at the Baltic Gender project management office at GEOMAR for the duration of the study. The sheet of questions from the survey are linked to the cover page through an identification number, making it possible to relocate and recover data in case of withdrawals.

For the studies, where no personal data is collected, a secret coding system is introduced to relocate the study. The participants create their own code (such as the first three letters of their mother's name and the last four numbers of their cell phone number) on the documents used for the survey.

For the online surveys, the SurveyMonkey platform is employed. The platform provides several survey distribution methods called collectors. Every collector type has different settings or collector options that affect the survey-taking experience and what respondent information is recorded. For Baltic Gender, Anonymous Responses is

turned on in the collector so collector data that makes respondents personally identifiable is not included in the survey results. This also means that IP addresses can not be tracked.

The sheet of questions from the studies are stored in locked cabinets and electronic data are stored in secure servers of the task leader. The task leader is responsible for the analysis of the data and ensures that all data is kept strictly confidential, that it is only used for scientific purposes and that no access is given to third parties or unauthorized persons - all in line with national and institutional data protection guidelines.

All personal data collected by studies will be deleted at least 5 years after the end of the project (unless an earlier withdrawal is requested by the participant).

The data protection procedures are already authorised by the competent Institutional Data Protection Officer or the National Data Protection Authority of each partner. The authorisations are filed at the Baltic Gender project management office at GEOMAR.

## 5. ETHICAL ASPECTS

***To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former***

Any ethical consideration related to the protection of personal data has been dealt with in a previous deliverable (BG-EthReq1-V5; BG-EthReq2-V4). Accordingly, all persons participating in the studies (**Section 1.2**) are asked for their written consent. Participation in the studies is voluntary and the participant is able to withdraw their consent anytime. The data collected from the studies are anonymous in that there is no personal data on the survey sheets. Identification numbers or a secret coding system ensure that the survey results can be located and be deleted should a participant request his/her withdrawal from the study.

For online surveys, where the online platform SurveyMonkey is used to collect individual responses anonymously, IP-addresses are not tracked and responses cannot be linked to individuals. At the end of the survey, the respondent is given an option to withdraw

from the survey. Due to full anonymity, it is not possible to withdraw information once the survey is finalized.

## 6. OTHER

***Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)***

Helmholtz Open Science Policy relevant to GEOMAR:
http://os.helmholtz.de/open-science-in-the-helmholtz-association/

CAU Data Management Guidelines
https://www.praesidium.uni-kiel.de/de/dokumente/leitlinie-zum-umgang-mit-forschungsdaten

'Open Science in Estonia' (principles and recommendations)
http://www.etag.ee/wp-content/uploads/2016/07/Avatud_Teadus_Eestis_1.0.pdf