



## **Mainstreaming Freshwater Invertebrates Biodiversity Data into Government Decision-Making**

**(BID-AF2020-169-USE)**

### **Freshwater Insects, Mollusks and Crustaceans (FIMC)**

#### **Data mobilization plan**

**September, 2021**

## S1. Introduction

The project aims at collating existing freshwater Insects, Mollusks, and Crustaceans (FIMC) species data of northeastern Tanzania from various institutions and database/archive sources, to create a useable and accessible FIMC (checklist and occurrence) dataset, distribution maps and database that can be adopted by policymakers and be integrated into decision-making processes by different stakeholders in different sectors.

So far, logistic-wise project's achievements includes: (i) formulation of inter-basin teams of northeastern Tanzania's data holders and users from across different institutions in Tanzania, (ii) signing of Memorandum of Understanding (MoUs) between TanBIF/COSTECH and data holders for material/ data transfer agreement, (iii) two physical meetings (one kick-off meeting as a familiarization and laying out of project logistical plan and one on training on data mobilization and use) took place between July and September 2021. The two meetings have generated among other data mobilization plan (Tables 1 and 2). Official launch of data mobilization between trained data mobilization team and data holders (with high level commitment of attendees from different institutions) was done.

To achieve FIMC data mobilization as a next step, mobilization will include planning for resources, (people and institutions and their roles, assigning tasks, defining expected outputs and setting timeline), validating mobilized FIMC datasets, cleaning and publishing in TanBIF/BGIF portal as detailed in Tables 1 and 2. It is to our expectations that the next steps will be successfully implemented that will enable realization of the achievement of the project outputs and impacts.

## S2: Planning resourcing

Tables 1 and 2 below has divided the project tasks based on the project implementation phases that were put in place during the proposal stage and subsequently appearing in grant letter.

**Table 1:** People and institutions and their roles

People and Institutions	Affiliations	Stakeholder	Roles
One fresh water invertebrates expert, One Community ecologist and One Conservation Biologist will be needed	1. The Nelson Mandela African Institution of Science and Technology (NM-AIST)	1. Dr Grite Nelson 2. Dr Issakwisa Ngondya 3. Dr Linus Munishi	1. To oversee the implementation of all project activities as planned. Also, to participate in project activities that include; ✓ To check data quality, ✓ To prepare FIMC checklist datasets, ✓ To prepare FIMC occurrence data sets, ✓ To select species of interest for developing species

			<ul style="list-style-type: none"> <li>✓ distribution maps, To develop (spatial-temporal) FIMC species distribution maps, to mobilize FIMC data</li> </ul>
One data mobilization expert affiliated with TaNBIF/ GBIF	1. Tanzania Commission for Science and Technology (CoSTECH)	1. Dr Hulda Gideon	<ul style="list-style-type: none"> <li>1. To use FIMC data</li> <li>2. To train data holders, mobilizers and partners</li> <li>3. To publish FIMC checklists and occurrence datasets</li> </ul>
Database developers	<ul style="list-style-type: none"> <li>1. Nugsoft Technologies</li> <li>2. Open University of Tanzania (OUT)</li> </ul>	<ul style="list-style-type: none"> <li>1. Mr Vincent Matsiko</li> <li>2. Dr Khamis Kalegele</li> </ul>	<ul style="list-style-type: none"> <li>1. To develop FIMC database</li> <li>2. Link existing FIMC data base with TanBIF</li> </ul>
Five FIMC experts from basin authorities and research institutes	<ul style="list-style-type: none"> <li>1. Pangani Basin Authority</li> <li>2. Lake Victoria Basin Authority</li> <li>3. Tanzania Fisheries Research Institute (TAFIRI)</li> <li>4. Tanzania Wildlife Research Institute (TAWIRI)</li> <li>5. University of Dar-es-salaam (UDSM)</li> </ul>	<ul style="list-style-type: none"> <li>1. Ms Arafa Maggidi</li> <li>2. Ms Rosemary Masikini</li> <li>3. Mr Hillary Mroso</li> <li>4. Ms Neema Kilimba</li> <li>5. Dr Lulu Kaaya</li> </ul>	<ul style="list-style-type: none"> <li>1. To provide FIMC data</li> <li>2. To mobilize FIMC data</li> <li>3. To use the mobilized FIMC data for policy and decision making</li> <li>4. To use the mobilized FIMC data, to develop (spatial temporal) FIMC species distribution maps</li> </ul>
Six data mobilization (junior scientists - Volunteers) who are graduate in the field of aquatic and terrestrial ecology	<ul style="list-style-type: none"> <li>1. University of Dar es Salaam (UDSM)</li> <li>2. The Nelson Mandela African Institution of Science and Technology (NM-AIST)</li> <li>3. College of African Wildlife Management (CAWM-MWEKA)</li> <li>4. Sokoine University of Agriculture (SUA)</li> </ul>	<ul style="list-style-type: none"> <li>1. Ms Clean Chrisant</li> <li>2. Mr Ally Ally</li> <li>3. Mr Arnold Shoko</li> <li>4. Ms Scholastica Mbinile</li> <li>5. Mr Stephen Maro</li> <li>6. Ms Salome Milola</li> </ul>	<ul style="list-style-type: none"> <li>1. To participate in the mobilization and validation of FIMC data</li> </ul>

**Table 2:** Tasks, expected outputs and timeline

Stage	Task	Responsible person	Expected output	Timeline
Initiating	<ul style="list-style-type: none"> <li>1. Team member familiarization</li> <li>2. Defining user needs</li> <li>3. Fine tune types of targeted species</li> <li>4. Identify data gaps</li> </ul>	1. All team members	1. User needs defined (species of conservation potential, indicators of habitat and ecosystem quality, species of human-health concern, and food	1. 5-11/7/ 2021

	5. Identification of data holders/ providers and users		<p>and farming potential)</p> <p>2. Types of targeted species fine tuned</p> <ul style="list-style-type: none"> <li>✓ Proposed 12 spp shrimps only 2 were in GBIF database without GPS coordinates</li> <li>✓ Proposed 1 spp of crabs, only 1 was found in GBIF database with GPS coordinates</li> <li>✓ Proposed 41 taxa of insects, only 18 was found to have records in GBIF database</li> <li>✓ Nine (9) genus of mollusks were in GBIF with records but without coordinates</li> </ul> <p>3. Data gap identified</p> <ul style="list-style-type: none"> <li>✓ All crustaceans and crabs to be Mobilized</li> <li>✓ Over 23 taxa of insects to be mobilized</li> <li>✓ Nine (9) genus of mollusks needs to be mobilized</li> </ul> <p>4. Seven data holders/ providers identified</p> <ul style="list-style-type: none"> <li>✓ Pangani water Basin Authority</li> <li>✓ Lake Victoria Water Basin Authority</li> <li>✓ Tanzania Fisheries Research Institute</li> <li>✓ Tanzania Wildlife Research Institute</li> <li>✓ Dr. Grite Nelson</li> <li>✓ Dr. Lulu Kaaya</li> <li>✓ Mr. Gordian Mataba</li> </ul>	
Planning	<ol style="list-style-type: none"> <li>1. Virtual training of FIMC data mobilization (August 2021)</li> <li>2. Physical training of</li> </ol>	<ol style="list-style-type: none"> <li>1. Six graduate volunteers and three research members</li> <li>2. Six graduate volunteers, six data providers/</li> </ol>	<ol style="list-style-type: none"> <li>1. Six graduate volunteers virtually trained on FIMC data mobilization</li> <li>2. Six graduate volunteers, project members, and six data holders/users physically trained on FIMC</li> </ol>	<ol style="list-style-type: none"> <li>1. 1 &amp; 8/9/2021</li> <li>2. 12-18/9/2021</li> </ol>

	FIMC data mobilizers and holders/ providers (September 2021)	holders (stakeholders) and nine research members	data mobilization	
Monitoring	1. Pre FIMC mobilization virtual (zoom)meeting with FIMC data mobilizers and holders/ providers (training follow up)	1. Nine team members 2. Six data holders/ providers 3. Six graduate volunteers	1. Improve gained FIMC data mobilization skills 2. Increase social interactions between data mobilizers (trained graduate volunteers) and data holders	1. 04/10/2021 and 11/10/2021
Executing	1. Mobilization of FIMC data (checklist and occurrence ) from: <ul style="list-style-type: none"> <li>✓ GBIF database</li> <li>✓ Pangani Basin Authority</li> <li>✓ Lake Victoria Authority</li> <li>✓ Individual researchers (Dr. Grite Nelson, Dr. Lulu Kaaya, Mr. Gordian Mataba)</li> </ul> 2. Publication of checklist and occurrence FIMC data to TaNBIF/GBIF 3. Develop (spatial-temporal) FIMC species distribution maps 4. Develop Northern Tanzanian FIMC species database	1. Four team members (Dr's Grite Nelson, Linus Munishi, Hulda Gideon and Issakwisa Ngondya) 2. Six data holders/ providers 3. Six graduate volunteers	1. Approximately > 10000 records of FIMC species occurrence mobilized and published in TaNBIF/GBIF 5. Spatial-temporal FIMC species distribution maps developed 2. A database for northern Tanzanian FIMC species developed	3. 15/10/2021 to 30/10/2022
Closing	1. Organize project's dissemination workshop	1. All project members 2. All data providers/ users 3. Local government authorities 4. Relevant Non-Governmental Organizations (NGO's)	1. Project's findings communicated to users	31/1/2023

### **S3: FIMC Data Quality and Standardization** (*Data validation check: data cleaning*)

Mobilized FIMC data will be validated and cleaned using GBIF validation and data cleaning tools (<https://www.gbif.org/tools/data-validator>). Data cleaning procedures will include checking all columns conformity to Darwin Core Standard (<https://dwc.tdwg.org/terms/>). For both checklist and occurrence datasets all required columns (taxonID, scientificName, taxonRank and occurrenceID, basisOfRecord, scientificName, eventDate respectively) and recommended columns will be filled in and cleaned. Sources of errors such as taxon match fuzzy (a match with different spelling), taxon match higher rank (no match found at the same taxonomic rank but only at higher rank), taxon match none (no match found), blanks, data format, data type and incompleteness will be corrected. Species names (nomenclature) will be confirmed using species matching tool and or catalog of life tool (<https://www.gbif.org/tools/species-lookup> or <https://www.catalogueoflife.org/> respectively). For occurrence records that will be missing global positioning system (gps) coordinates, geo-locate tool (<https://www.geo-locate.org/web/default.html>) will be used to source the gps coordinates.

### **S4: FIMC Dataset publication**

Following a thorough dataset quality check and standardization procedures (see S.3 above), mobilized FIMC datasets will be published in TaNBIF/ GBIF portal. All licensing requirements will be followed whereby recommended and required license compliances including those based on agreements with data holders/sources will apply. GBIF Integrated Publishing Toolkit (IPT) tool will be used to publish the datasets (<https://www.gbif.org/ipt>). The procedures for publication will include (i) selection of a dataset/datasets for publication (ii) transform the dataset in accordance to Darwin Core Standard (iii) upload datasets to IPT (iv) map datasets to taxon (v) fill in metadata (vi) publish dataset to TaNBIF/ GBIF database (vii) register datasets.