University of Abomey-Calavi (UAC)
****

Faculty of Agronomic Sciences (FSA)
****

Workshop report

Capacity building for national partners on biodiversity data mobilization, digitization, cleaning, and publication

Abomey-Calavi, Sainte Anouarite reception house
November 30th and December 1st , 2021

December 2021

Contents

List of photos

Introduction

A capacity building workshop for national partners took place at Sainte Anouarite reception center (South Benin) on November 30th and December 1st, 2021, to train participants. This workshop aimed to build participants' capacities on data mobilization, provide them the necessary technical support for the publication of data, and help them become more visible by registering on GBIF website and publishing biodiversity data. Participants came from various institutions working in the field of biodiversity. Their list is in appendix 01. The agenda of the workshop (see appendix 02) was structured in five parts.

- Opening ceremony ;
- Digitization of biodiversity data;
- Data cleaning and formatting;
- Publication of data;
- Closing ceremony.

1. Opening ceremony

The installation of the participants started at 8:30 am and the opening of the workshop started at 9:00 am. The opening ceremony (Photo 01) was marked by the opening address of the Dean of the Faculty of Agronomic Sciences of the University of Abomey-Calavi, Professor AHOHUENDO Bonaventure. He first welcomed the participants, then, presented GBIF through its objectives, missions, member countries, etc. He also covered the history of GBIF. He also reviewed the main achievements of GBIF Benin. He presented the objectives of the workshop and reiterate the commitment of the Faculty to support GBIF-Benin in all its activities. Following this, he particularly thanked GBIF and the European Union for their financial support to GBIF-Benin and officially launched the workshop.

Photo 01: Officials during the opening ceremony

After the opening ceremony, the participants introduced themselves. The moderator, Dr. AOUDJI Augustin, Associate Professor then, presented the workshop program.

2. Digitization of biodiversity data

2.1. Presentation of GBIF

Dr KOURA Kourouma presented GBIF (Photo 02) through its history, the GBIF network with its 102 member states and organizations, the GBIF mission, the GBIF data portal with its more than 1,902,001,522 data of occurrences as of 30th November 2021. According to her, GBIF's mission is to facilitate open access to primary biodiversity data to support scientific research, natural resource conservation and sustainable development. Through this network, biodiversity data is mobilized and made available free of charge to users around the world.

Photo 02: Dr KOURA Kourouma during her presentation of GBIF

2.2.   Presentation of GBIF Benin

This presentation was made by Professor Jean Cossi GANGLO (Photo 03). In his presentation, he highlighted the importance of biodiversity, the threats to biodiversity, and the main achievements of GBIF-Benin. It should be remembered from his presentation that biodiversity provides eco-systemic services indispensable to the survival of humanity. There are four of them: provisioning services, regulating services, cultural services, and support services. Despite the importance biodiversity it is threatened by: the destruction of habitats

and the modification of environments; overexploitation of resources; climatic extremes; the introduction of invasive alien species; pollution and diseases. These threats create a significant loss of forest area and other losses of biodiversity. In the face of threats, the Global Biodiversity Information Facility (GBIF) was needed to define and facilitate biodiversity conservation strategies and actions. This system, an initiative of the Organization for Economic Co-operation and Development (OECD), is a global network which was created in 2001. Benin has been a member of this network since December 2004 and acquired the status of voting member in October 2011. The objectives of GBIF-Benin are to mobilize data on biodiversity; to provide technical and scientific training for national and regional partners in the field of biodiversity informatics; to use the data mobilized and make the research results available to inform decision-making on biodiversity conservation and, to promote national partners not only by strengthening their capacity in biodiversity informatics but also by fostering their visibility through data publication on GBIF site. Professor GANGLO presented the achievements of GBIF-Benin as follows:

- The publication, as of 30th November 2021, of nearly 900,000 published occurrence biodiversity data of Benin out of which at least 85 % (nearly 800,000) are by GBIF-Benin (http://www.gbif.org/country/BJ/summary);
- The organization each year of at least one capacity building workshop to train partners in data mobilization and data use;
- The development of the GBIF Benin internet portal (https://gbif-benin.org);
- The establishment of a master and Ph. D program in biodiversity informatics at the Faculty of Agronomic Sciences (FSA).

Photo 03: Professor Jean Cossi GANGLO during his presentation on GBIF-Benin

2.3. Types of biodiversity data published on GBIF site

Ph. D students Gorgias AIKPON and Justine KOTIN presented the types of biodiversity data taken into account in the framework of GBIF's activities. Three types of data are published through the GBIF portal:

- Occurrence data;
- Sampling protocols;
- Checklists

According to them, occurrence data or primary biodiversity data have 04 essential attributes: the full updated scientific name; the complete description or geographic coordinates of the place of occurrence (observation, collection); the complete date of collection (day, month, and year) and the nature of data collected (Basis of record). The possible basis of records are: plant specimens in herbaria, animal collections, data based on survey protocols, observational data, and multimedia data. According to them, metadata are indispensable for data publication; they are structured descriptions of datasets; they provide essential details such as the geographic and taxonomic scope of the data, methods of collection or observation, contacts and citation requirements. Checklists are lists of species belonging to given categories (eg : taxonomic, geographic, trait-based, red list). Some of the errors to be avoided are duplicate data. These duplicate data are data relating to the same species, observed at the same date and at the same location.

Photo 04: Ph. D students Justine KOTIN and Gorgias AIKPON during their presentation on the types of biodiversity data

2.4. Importance of biodiversity information and benefits of data publication

Professor GANGLO presented the importance and benefits of publishing biodiversity data. In his presentation, he, first, enlightened the audience on the notion of data and the notion of publication of data. He, then, expanded on data sources before addressing the importance of biodiversity data and the interest linked to the publication of data. Finally, he addressed property rights in order to reassure data providers about the accessibility of the data they publish.

2.5. Procedure for registering institutions on the GBIF website

Professor Jean GANGLO made it clear to the participants that the registration of institutions on GBIF website gives them international visibility. It is also essential for the publication of biodiversity data if the institution wishes to maintain its authorship on the data published. Among the institutions registered on the GBIF website, we noted 32 Beninese institutions. Prof. GANGLO, then, presented the institution registration process on the GBIF website. The form to be completed is at the following link: https://www.gbif.org/fr/become-a-publisher.

2.6. Digitizing biodiversity data (practical session)

The GBIF data entry template was presented by Ph. D students GBENOU Tania, APELETE Eben-ezer, and Sunday Berlioz KAKPO (Photo 05). They explained to the participants, each field of the canvas and how to fill them out.

To make the training practical, participants received data from a floristic inventory and practiced filling out the GBIF data entry spreadsheet.

Photo 05: Ph. D students GBENOU Tania, APELETE Eben-ezer, and KAKPO Sunday Berlioz during their presentation

3. Data cleaning and formatting

3.1. Basic data cleaning concepts and data management tools

The trainers and Ph. D students Miss GBENOU Tania, Mr. APELETE Eben-Ezer, and Mr. Sunday Berlioz KAKPO supervised the participants on the basic concepts of data cleaning and the tools of data management. In their first presentations, they notably addressed aspects of data digitization, inventory data entry and probable errors that one might encounter in databases after submission for publication. They were able to make participants understand that data cleaning is a process used to improve data quality by correcting detected errors. It encompasses improving the quality of data to make them "fit for use". The cleaning steps are:

- Define and determine the types of errors (look for and identify the errors of occurrences);
- Correct the errors;
- Document the cases and types of errors;
- Modify our data entry process to reduce future errors.

In the rest of their presentation, they warned participants about the different types of probable errors that one might encounter in databases:

- Errors related to nomenclature (Taxonomic data);
- Temporal data errors;
- Spatial data errors.

The second presentation focused much more on the main tools for visualization and geo-referencing of biodiversity data. As tools, we have:

- Excel spreadsheet (TXT, CSV) for converting geographic coordinates, converting date (format) and splitting cells;
- OpenRefine for data cleaning;
- Canadensys for contact details conversion;
- Global name resolver for taxonomy;
- Geolocate, USGS, for georeferencing;
- QGIS, ARCGIS software for transformation, analysis, visualization, cleaning.

3.2. Using OpenRefine in Data Cleansing

Doctor Kourouma KOURA communicated on the use of OpenRefine to clean data formatted according to the GBIF spreadsheet. Thus, each participant installed the software on his computer and applied the instructions of the presenter in order to detect and correct errors in the dataset submitted to them. In addition, participants learned to use other OpenRefine resources such as column creation and deletion tools, tools to generate different taxa using the genus of a species.

3.3. Data publication

The last presentation of the workshop was made by Professor GANGLO. He briefly returned to the basic concepts of Biodiversity Informatics (BI) and highlighted the particularities of biodiversity informatics. He explained the possibilities and potentialities offered by biodiversity informatics for an effective conservation of biodiversity in Benin and the world in general. To be concrete, he presented the results of BI research from the Forest Sciences Laboratory (LSF). Finally, he noted the obstacles linked to the publication of data and possible solutions in order to dispel any doubt linked to the publication of biodiversity data.

4. Closing ceremony

The closing ceremony of the workshop took place as follows:

- The speech of the representative of the General Director of Water, Forests and Hunt, Captain Doctor OROU MATILO Augustin;
- The speech by the Node Manager of GBIF Benin, Professor GANGLO Jean Cossi.

4.1. Speech of the representative of the General Director of Water, Forests and Hunt

Captain Doctor OROU MATILO Augustin, in his speech, recognized that this workshop is of great importance as it constitutes major technical support for national partners in the field of mobilization, publication and use of biodiversity data. It is therefore important to continue to perpetuate this capacity building. He thanked the Faculty of Agronomic Sciences for these efforts in the field of biodiversity conservation and GBIF Benin for the continued support and importance it places on the publication of biodiversity data.

He, then, launched an appeal to the institutions represented to encourage them to register on GBIF website and publish their data. This will not only enhance the visibility and recognition of their institutions but also raise the image of Benin at international level.

4.2. Speech of the Node Manager of GBIF Benin

The Node Manager of GBIF-Benin, Professor Jean Cossi GANGLO, thanked the various heads of institutions, in particular the Dean of the Faculty of Agronomic Sciences of the University of Abomey-Calavi, the Minister of Living Environment and Sustainable Development, and the General Director of Water, Forests and Hunt. He also mentioned that GBIF-Benin needs more support from them and also from other institutions involved in biodiversity management to meet the challenges of publishing data and training in Biodiversity Informatics. To end his speech, Professor Jean C. GANGLO thanked GBIF and the European Union for their financial supports and all the trainers for their active involvement in the various activities of GBIF-Benin.