



Union - Discipline - Travail



**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE FELIX HOUPHOUET-BOIGNY
UFR BIOSCIENCES
CENTRE NATIONAL DE FLORISTIQUE**

RAPPORT D'ATELIER

**FORMATION DES FOURNISSEURS DE DONNEES BIOLOGIQUES AUX BONNES
PRATIQUES DE COLLECTE ET DE MOBILISATION DES DONNEES
BIOLOGIQUES ET DE DIFFUSION DE L'INFORMATION SCIENTIFIQUE A
TRAVERS LA REDACTION DES DATA PAPERS**

18-22 Juillet 2022, salle MORFALL, CAMPC

Université Félix HOUPHOUËT-BOIGNY, Cocody, Abidjan (Côte d'Ivoire)

1-Introduction

Sur invitation du Centre National de Floristique (CNF), en partenariat avec des Institutions nationales et internationales comme les UFR des Sciences Biologiques des Universités Félix HOUPHOUËT-BOIGNY (UFHB), NANGUI Abrogoua (UNA) , Jean Lorougnon GUEDE (UJLoG), San-Pedro (USP), Peleforo GON COULIBALY (UPGC), le point focal CDB du Ministère de l'Environnement et du développement Durable (MINEDD), l'Institut Pasteur de Côte d'Ivoire (IPCI) et le GBIF France s'est tenu du 18 au 22 Juillet 2022, à Abidjan, un atelier sur la « **formation des fournisseurs de données biologiques aux bonnes pratiques de collecte et de mobilisation des données biologiques et de diffusion de l'information scientifique à travers la rédaction des DATA PAPERS**».

Cette formation est en faveur de l'exécution du projet financé par le GBIF (Biodiversity Information for Development (BID)-BID-AF2020-128-NAC). Ce projet a pour objectif d'une part de mobiliser des données et informations biologiques concernant les organismes vivants (animaux, insectes, champignons, algues, virus, bactéries, etc.) et d'autres part de publier ces informations sous forme de jeux de données biologiques sur le portail du GBIF, ainsi que sous le format des *Data Papers*. La mobilisation des données auprès de 14 fournisseurs de données biologiques animales et végétales, a permis d'obtenir plus de 24000 données d'occurrences. C'est donc dans le but de former les fournisseurs de données biologiques aux bonnes pratiques de collecte, de mobilisation et de diffusion à travers la rédaction des *Data Papers* que cet atelier de formation a été initié.

2- Objectifs de l'atelier

L'objectif général de cet atelier est de former les fournisseurs de données biologiques aux bonnes pratiques de collecte, de mobilisation et de diffusion de l'information sur les données biologiques à travers la rédaction des *Data Papers*. De façon spécifique, il s'est agi de :

1. présenter les outils informatiques utilisables en biodiversité informatique pour la saisie et le nettoyage des données biologiques,
2. publier des données biologiques sur la plateforme du GBIF,
3. rédiger les *Data Papers*.

3-Participants

Ont participé à l'atelier de formation le Vice-président de l'Université Félix Houphouët Boigny (UFHB), en Charge de la Recherche et de l'Innovation Scientifique, l'Université Jean Lorougnon Guédé (UJLOG), les points focaux du Ministère de l'Environnement et du développement Durable (MINEDD) ainsi que le Directeur du Laboratoire des Milieux Naturels et Conservation de la Biodiversité (LMNCB) de l'UFR Biosciences de l'Université Félix Houphouët Boigny.

Aussi, ont-ils répondu présents à l'invitation du Centre National de Floristique, les représentants du Centre National de Floristique, de l'Université Jean Lorougnon Guédé, de l'Université Peleforo GON COULIBALY, de l'Université Nanguy Abrogoua (UNA) et de l'Institut Pasteur de Côte d'Ivoire (IPCI). La liste des participants est jointe en annexe.

4-Cérémonie d'ouverture

Après la présentation du programme du déroulement de l'activité par le Maître de cérémonie, Dr Cissé Abdoulaye, la cérémonie a débuté par le Mot de bienvenue du Professeur. TIEBRE Marie-Solange Epouse Ransonnet, Professeur Titulaire et Directrice Adjointe du Centre National de Floristique (CNF), à toutes les délégations venues participés à l'atelier. Tout en remerciant les organisateurs de l'atelier notamment Dr KOUA Béranger, Coordonnateur du projet, Professeur TIEBRE à clos ses propos en souhaitant une bonne formation à tous les participants.

Suite au mot de bienvenue de la Présidente du comité d'organisation de l'atelier, le Vice-président de l'Université en Charge de la Recherche et de l'Innovation Scientifique, Prof. Sangaré Abdoulaye, a adressé au nom du Président de l'UFHB son mot de bienvenue à l'atelier de formation. Il a tenu à féliciter le comité d'organisation pour la réalisation de l'atelier et a souligner l'importance de cet atelier de formation ainsi que la mise en avant du service à la communauté au travers des solutions proposées à la recherche. Après avoir remercié les formatrices Mesdames Sophie PAMERLON et Alice AINSA de l'équipe du GBIF-France, le Vice-Président a déclaré ouvert l'atelier sur la « **Formation des fournisseurs de données biologiques aux bonnes pratiques de collecte et de mobilisation des données biologiques et de diffusion de l'information scientifique à travers la rédaction des *Data Papers*** ».

5-Déroulement des activités

➤ Jour 1 des activités : 18-07-2022

5-1-Présentation du réseau GBIF par Madame Sophie PAMERLON -Equipe GBIF-France

Les activités de la première journée ont débuté par la présentation du portail GBIF par Sophie PAMERLON de l'équipe du GBIF-France. Dans son intervention, elle a présenté le GBIF en tant que réseau intergouvernemental et infrastructures de données qui fournit un accès libre aux données d'occurrences sur la biodiversité. Elle a aussi indiqué que grâce aux collectes de données à travers les pays membres et fournisseurs de données, plus de 2,2 milliards de données sont disponibles sur le portail du GBIF. Selon ses propos, le partage des données permet d'ouvrir une fenêtre sur où, quand et par qui a été observée une espèce. Dans la dynamique d'évolution du portail GBIF, d'autre type de données telles que les analyses d'ADN sont en cours d'intégration. Les données du GBIF, permettent de collecter des données brutes et de les convertir en des modèles de simulations pour le suivi et la protection de la biodiversité. Il sert donc de base des décisions prises au niveau mondial pour la conservation et évaluations de la biodiversité au niveau international. En ce qui concerne les données, elles sont issues de diverses institutions scientifiques et de recherches sur la biodiversité, les agences gouvernementales, les ONGs, les réseaux de sciences citoyennes, les réseaux thématiques (données marines, agro-biodiversité, séquences d'ADN), les Entreprises et consultants, *etc.*

Pendant la présentation des exemples de pages GBIF d'espèces, d'éditeurs de données et du processus (le principe FAIR pour la collecte et la mise à disposition des données) de publication des données sur le portail ont été présentés.

Madame Sophie PARMERLON a terminé sa présentation en mettant en évidence les différents domaines d'application des données du GBIF. Il s'agit des domaines tels que les politiques de conservation, la sécurité alimentaire, le changement climatique et la santé humaine.

5-2--Présentation du portail GBIF (Alice AINSA/GBIF-France)

Madame Alice AINSA, dans son intervention, a présenté la prise en main du portail GBIF accessible à partir du lien www.gbif.org. Sur le site plusieurs informations relatives à l'espèce, à la taxonomie, au lieu de collecte, etc. peuvent être consultées. De même certaines

statistiques (tableaux, figures) relatives à l'espèce ainsi que ses occurrences et le statut de conservation UICN sont fournis.

A la fin des deux présentations, un échange entre les participants, les organisateurs de l'atelier et les représentants du GBIF a eu lieu.

5-3-Présentation des standards et des types de données du GBIF (Madame Sophie PAMERLON -Equipe GBIF-France)

Les différents types de données disponibles sur le portail du GBIF ont été listé par Madame Sophie PAMERLON. Ces types sont les suivants :

-Référentiels et listes taxonomiques : liste d'espèces inventoriées ; liste d'espèces valide leurs hiérarchie et synonymes ; liste rouge des espèces avec leur statut de vulnérabilité.

-Fichier modèle GBIF pour les données taxonomiques.

-Specimens et échantillons issus des collections d'Histoire naturelle : planche d'herbier ; specimens conservés dans du liquide de conservation ; fossiles et autres matériels

-Données issues de publications scientifiques

-Notes et enregistrements de terrain : enquêtes, études d'impacts ; carnets de terrain, fiches ; données de sciences participatives

-Fichier modèle GBIF pour les données d'échantillonnage : pour les données d'inventaires

-Autres sources de données : données obtenues par GPS, radar, satellite ; données de cartes papier, impressions d'images ; photos ou enregistrements.

Les activités de l'après-midi ont concerné des exercices pratiques de prise en main et de nettoyage des données. Pendant ces exercices, Il s'est agi d'effectuer des vérifications dans la base de données mise à notre disposition d'éventuelles erreurs qui s'y sont insérées et de procéder aux différentes corrections qui y découlent.

➤ Jour 2 de l'atelier : 19-Juillet-2022

5-4-Suite des exercices et présentation de OpenRefine

La deuxième journée de l'atelier a débuté par des exercices d'application entamé le jour précédant. Cet exercice a porté sur l'utilisation sur le nettoyage d'une base de données avec des outils de réconciliation et de taxonomie (Taxonomic Name Resolution Service : TNRS ; Canadensys : correction des coordonnées géographiques ; etc.)

Par la suite, Madame Sophie PAMERLON a fait une présentation du logiciel OpenRefine. Sa présentation a porté sur l'Utilisation de OpenRefine dans le nettoyage et la standardisation des données. C'est un outil puissant pour travailler sur les données à nettoyer avant la publication sur le portail du GBIF. Elle a aussi indiqué le processus de nettoyage qui se fait en trois étapes qui sont le Facettage-la Fusion-la Réconciliation. A la différence des autres logiciels, OpenRefine permet la modification de plusieurs cellules en simultané ; l'exploration et la transformation facile des données et la visualisation interactive. Par ailleurs, c'est un logiciel gratuit d'application qui fonctionne hors-ligne et dont la Fonction "annuler" est infinie. Il permet l'utilisation d'API et d'importer et exporter différents types de données.

Succédant à la présentation du logiciel OpenRefine, des exercices ont permis aux participants d'effectuer le nettoyage d'un jeu de données mis à leur disposition.

5-5- Présentation du processus de publication des données des données sur le GBIF

Les partenaires du GBIF fournissant les données sont entre autres MNHN, Symbiota, des IPT nationales, etc.

Le dispositif de publication se fait selon le processus suivant :

1. Nettoyage des données et standardisation
2. Publication des données sur le GBIF via l'IPT
3. Indexation et affichage des données sur GBIF.org
4. Visualisation, filtrage et téléchargement
5. Vérification des données téléchargées

Les données publiées sont généralement associées à des métadonnées fournissant toutes les informations relatives aux jeux de données fournies au GBIF. Le but des métadonnées étant de décrire les données de manière précise pour les autres utilisateurs. Par ailleurs le choix de licence est un élément important dans la publication des données sur le GBIF. Concernant la publication des données sur le GBIF, l'IPT (Integrated Publishing Toolkit) représente l'outil principal. Un exercice d'application a permis aux participants d'effectuer une prise en main de l'IPT et de créer une ressource.

➤ **Jour 3 de l'atelier : 20-Juillet-2022**

La journée a débuté par la présentation du concept du DATA PAPERS. Dans son intervention, Madame PARLMELAN a défini le DATA PAPER comme une publication scientifique, dont le but principal est de décrire un ensemble de données ou un groupe d'ensembles de données, plutôt que de rendre compte d'analyses et de résultats de recherche. Son but est de mettre en valeur les jeux de données, ainsi que le temps passé par les gestionnaires de données à les structurer, nettoyer et mettre en ligne. L'intérêt des Data Papers s'est de décrire les données sous forme structurée et lisible par un humain. Il permet aussi de promouvoir et faire connaître les données. Les DATA PAPERS ont aussi valeurs d'articles scientifiques qui sont publiés dans des revues scientifiques avec des preuves d'indexation. A ce niveau, la rédaction du DATA PAPERS pour la publication se fait généralement en anglais. Par ailleurs, l'ajout d'analyses statistiques ou de représentations graphiques est possible.

Concernant les métadonnées affiliées aux Data Papers, elles permettent de lister toutes les personnes impliquées dans la publication des données et de fournir plusieurs informations descriptives du jeu de données mis à disposition. Au cours de la présentation, Madame PAMERLON a précisé que Le GBIF collabore avec ses partenaires de l'édition scientifique afin de promouvoir les data papers comme un moyen de d'attribuer citation et reconnaissance académique à toutes les personnes impliquées dans la publication de données, d'alerter la communauté scientifique de l'existence de jeux de données sur la biodiversité et de l'intérêt qu'ils représentent potentiellement pour des projets de recherche spécifiques, et en tant que mécanisme d'évaluation de la qualité des données accessibles via le GBIF et d'autres réseaux.

A la suite de la présentation, des exercices pratiques sur l'IPT et la visualisation des données sur GBIF.org ont été réalisés. Les exercices ont consisté à compléter les métadonnées

en décrivant un jeu de données mis à la disposition des participants afin d'aboutir à la rédaction d'un DATA PAPERS. Les exercices pratiques sur l'IPT, centrée sur les métadonnées permettant de générer un fichier EML nécessaire à la publication d'un Data Paper se sont poursuivis.

➤ **Jour 4 de l'atelier : 21 Juillet 2022**

5-6-Redaction du Data Papers

Les outils pour la rédaction des Data papers ont été présenté aux participants. La présentation a porté sur l'outil de rédaction ARPHA Writing Tool (Pensoft) qui sert à faciliter la mise en page, la soumission, le processus de relecture, la publication, l'hébergement et l'archivage d'articles scientifiques et des Data papers. Par la suite une démonstration de l'utilisation de l'interface <https://arpha.pensoft.net/> pour importer les fichiers EML a été faite.

Dans la session de l'après-midi, Madame Alice, a instruit les participants sur les différentes facettes du portail du GBIF. La démonstration a servi à montrer les différents champs de collecte des données sur le GBIF avec les filtres présents. Ces filtres ont permis de faire des recherches d'occurrences par exemple par fournisseurs de données, par pays, sur une période, etc.

Les activités de la formation ont pris fin par des échanges entre les participants et les formateurs et une évaluation des modules de la formation par les participants.

6-Cérémonie de clôture

La cérémonie de clôture a débuté par le mot de remerciements des participants représenté par Dr Kouamé de l'IPBES, à l'égard des organisateurs et des formatrices. La cérémonie a été également animé par la remise de diplôme de participation à l'atelier de formation des fournisseurs de données de qualité et de rédaction des Data Papers.

Dans son mot de clôture, le Professeur TIEBRE Marie-Solange Epouse RANSONNET, Directrice adjointe du Centre National de Floristique (CNF), a bien voulu remercier les participants, les formateurs ainsi que l'équipe d'organisation de l'atelier notamment Les Docteurs N'GORAN Koua Béranger (Coordonnateur du projet Biodiversity Information for Development (BID)-BID-AF2020-128-NAC), Akossoua Faustine épouse Koffi et la secrétaire du CNF, Mademoiselle KAMAGATE Ami.

Après une brève invitation aux participants à la mise en pratique des bonnes pratiques de saisies des données biologiques et de publication des Data Papers, le Professeur Tiébré en souhaitant aux participants un meilleur retour dans leurs familles respectives, a déclaré clos **l'atelier sur la« Formation des fournisseurs de données biologiques aux bonnes pratiques de collecte et de mobilisation des données biologiques et de diffusion de l'information scientifique a travers la rédaction des DATA PAPERS ».**

Fait à Abidjan le 22 Juillet 2022

Les rapporteurs

Dr YAO Saraka Didier Martial

Dr ASSEH Ebah Estelle