Biodiversity data mobilization in natural history state collections of the Dominican Republic

Programme: BID
Project ID: BID-CA2020-031-NAC
Project lead organization: Instituto de Investigaciones Botánicas y Zoológicas, Universidad Autónoma de Santo Domingo
Project implementation period: 1/8/2021 - 31/5/2023
Report approved: 7/7/2022

Narrative Midterm report

Executive Summary

Achievements to date:
1. Integration of the teams of both institutions in a satisfactory manner.
2. Publication of two datasets, one for each institution.
3. Obtained two certifications from two members of the project, Oniel Álvarez and Gabriel de los Santos.
4. Presentation of the project to the community at the CEICYT RD 2021, the JIC 2021 and the XVII CIC 2022. (see attachment)
5. Digitization of an important part of the records committed to both institutions.
6. All the profiles required to work with the datasets have been generated in the Symbiota portals. For example:
• Fishes: https://csvcoll.org/portal/collections/misc/collprofiles.php?collid=1066
• Insects: https://serv.biokic.asu.edu/ecdysis/collections/misc/collprofiles.php?collid=37
• Arachnida: https://serv.biokic.asu.edu/ecdysis/collections/misc/collprofiles.php?collid=67

How you have been monitoring and evaluating your project's progress:

A coordinator has been designated for the team at the Museum (G. De los Santos), who maintains continuous communication on the progress of the work with the other colleagues. At IIBZ, together with O. Álvarez, we follow up on the progress of the work teams that we have formed according to the datasets committed: a botanical team with the students that have been hired and a group of five volunteers that have been trained, a malacology team, and one for entomology.

Any changes that you wish to make to ensure the successful implementation of your project in the remaining project period:

1. We want to change the name of some datasets to make them more precise. For example, the dataset called "Ants" we would like to call it "Ants of two urban parks in Santo Domingo".

2. We want to merge the two workshops into one and have it at the end of the project, in April.

Progress against milestones

Has your project published at least one dataset through GBIF.org?: Yes

Dataset published:

Dataset	DOI
Colección Entomológica del Instituto de Investigaciones Botánicas y Zoológicas / Ants	https://doi.org/10.15468/h6453y
Museo Nacional de Historia Natural “Prof. Eugenio de Jesús Marcano” - Colección de Arácnidos	https://doi.org/10.15468/v9pt63

Has at least one member of your project team received certification following the BID capacity enhancement workshop?: Yes

Name of the workshop participant: Biodiversity Data Mobilization

Certification obtained: Advanced Badge

Report on Activities

Activity progress summary

Activity 3 Determination and organization of specimens

Specimens determination is being done at the lowest possible taxonomic level. All specimens to be digitized have already been separated for processing. The datasets of savannah insects, aquatic macroinvertebrates and insects from the IIBZ entomological collection are starting to be processed. All the Museum's datasets already have this part completed.

Activity 5 Digitization of specimens

The digitization of the specimens has been progressing satisfactorily. The Museo has succeeded in cleaning a good part of the data. Based on the programming of the work plan, we estimate that this activity will not have any setbacks in its execution.

Activity 8 Input of the collection data from the Museum into the Museum database

The activity presented difficulties due to technical issues and currently no records are being entered in the SQL database. The solution has been to continue adding the records directly in Excel sheet; this is the case for Amphibians, Reptiles, Bats and the type collection. The Orthoptera collection will be entered directly into the Ecdysis portal.

Activity 9 Migration of the museum collections data to GBIF.

This activity was scheduled to start in January 2022, but has not been possible due to problems with the SQL database. Understanding the advantages of working with Symbiota, the Museum collaborated with the Arizona State University - Biodiversity Knowledge Integration Center to upload the data first in the respective Symbiota platform portals and then mobilize them to GBIF. In this order, the records of arachnids (not spiders) have already been entered in the Ecdysis portal and published in GBIF. The CSVColl portal will host the vertebrate records. The Museum's type catalog will be published through IPT.

Activity 11 Workshop for data users

We have considered that given the little developed knowledge on digitization and mobilization of biodiversity data in the Dominican Republic, it would be convenient to have a much more complete workshop than the one we had initially thought.

In addition, it seems necessary that we design some reference manuals for the Biology students of the Universidad Autónoma de Santo Domingo, who in the present and near future will begin to change the way in which we capture biodiversity data in the country.

It will also be necessary to strengthen the relationship with other institutions in order to progress more quickly in this area. Therefore, we propose that the workshop should be held in April 2023, instead of the initially scheduled date.

Completed activities

Activity: BID Capacity Enhancement workshop – Data mobilization and publication

Description: A mandatory workshop for representant of
the project on data mobilization
Start Date - End Date: 3/8/2022 - 6/8/2022
Verification Sources: (1) Biodiversity Data Mobilization, Basic Badge - Oniel Álvarez
(2) Biodiversity Data Mobilization, Advanced Badge - Gabriel de los Santos

Activity: Biodiversity data mobilization plan

Description: Mandatory milestone attached to the Early progress report. A data mobilization plan will be designed for both dataholder institutions
Start Date - End Date: 14/9/2021 - 29/10/2021
Verification Sources: See the file referent to work meeting: Reunión proyecto BID 2022-03-04

Activity: The workshop was postponed as a result of the management of the COVID-19 situation. The staff of each institution was trained by instructors Oniel Alvarez (IIBZ) and Gabriel De los Santos (Museum). We had a face-to-face meeting of both teams. Then we acce

Description: The workshop was postponed as a result of the management of the COVID-19 situation. The staff of each institution was trained by instructors Oniel Alvarez (IIBZ) and Gabriel De los Santos (Museum). We had a face-to-face meeting of both teams. Then we accessed the workshop "Introduction to the Digitization of Biological Collections with the Symbiota platform" (October 28, 2021) at the XXIV Mesoamerican Society for Biology and Conservation Congress-El Salvador (online).
We also participated in the course Publication of biodiversity data with IPT (17-18 March 2022), as part of the project Extending knowledge on biodiversity data quality and publication in the Spanish-speaking community, Capacity Enhancement Support Programme, GBIF. Regular meetings were scheduled with Samanta Orellana (Arizona State University - Biodiversity Knowledge Integration Center) to train all staff in the use of the Symbiota portals and IIBZ collection profiles.
Start Date - End Date: 28/10/2021 - 18/3/2022
Verification Sources: Work Meeting BID-CA2020-031-NAC List.pdf; photographs: Bastardo, R. Publicación con IPT; Work meeting Symbiota-IIBZ; Symbiota workshop El Salvador 1, 2, 3 y 4.

Activity: Gain certification at BID Capacity Enhancement Workshop

Description: The BID Capacity Enhancement Workshop was successfully completed with the award of two certifications. We requested permission to present two candidates, one for each institution: Gabriel De los Santos (Museo-Advanced Badge) and Oniel Alvarez Abreu (IIBZ-Basic Badge).
Start Date - End Date: 3/6/2021 - 10/3/2022
Verification Sources: 1. Biodiversity Data Mobilization - Gabriel de los Santos,2. Biodiversity Data Mobilization – Oniel Álvarez Abreu

Activity: Publication of at least one dataset through GBIF

Description: Two datasets were published ahead of schedule, one for each institution (IIBZ and Museum). For the IIBZ dataset (Ants) we managed to digitize more records than expected.
Start Date - End Date: 21/4/2022 - 31/5/2022
Verification Sources: Ants: https://doi.org/10.15468/h6453y
Colección de Arácnidos: https://doi.org/10.15468/v9pt63

Report on Deliverables

Deliverables progress summary

We committed 20 datasets of occurrences, checklists and sampling events. They are distributed among the participating institutions of the project, the IIBZ and the Museum. Both institutions have already published a dataset.

Datasets from the Instituto de Investigaciones Botánicas y Zoológicas are:

• Ants. 400. Published
• Fishes from the Lago Enriquillo. All 100 records digitized, pending of publishing.
• Molluscs: 130 out of 400 committed records have been digitized for the dataset.
• Herbarium USD: 18 % digitized. 178 records from 1000 digitized in the database.
• Insects from savannah: The digitizing of this dataset will start in July 2022, and it is expected to be published in January 2023.
• Insect collection IIBZ: 2000 (Orthoptera. 400, Butterflies. 700, Chrysomelidae. 800, Blattodea) 100. Digitization will start in July 2022.
• Padre Fuertes collection at Herbarium USD: 500. The digitization of this dataset will start in August 2022.
• Aquatic macroinvertebrates: 0 % complete. 500. The digitization of this dataset will start tin August 2022. This need some advices from Leonardo Buitrago, because it is the only one event sample dataset.
• Odonata. 2000 records. Currently digitizing specimens. Publishing expected by April 2023.
• Fishes from Lago Enriquillo. The digitization of this dataset will start in July 2022.
• Padre Fuertes Collection. This dataset is expected to conclude by March 2023.

• Types of Herbarium USD. March 2023.
• Ants from urban areas of Santo Domingo. 20 %. 2022-12-30

It is expected to publish all datasets from Museo Nacional de Historia Natural “Prof. Eugenio de Jesús Marcano” between July 2022 and April 2023. Datasets from the Museo are:

• Arachnids Collection MNHNSD: 27 % Published, but incomplete.
• Bats Collection MNHNSD. Currently 249 of the 400 records digitized.
• Herpetological Collection MNHNSD: 28 % of 2000 records digitized.
• Birds Collection MNHNSD. 400 from 400 digitized.
• Collection of Types of MNHNSD: 439 of 500 records digitized.
• Orthoptera Collection MNHNSD: 81 of 500 records digitized.
• Checklist of types collection MNHNSD. This dataset will start in July 2022.

Progress towards deliverables

Dataset deliverables

Herbarium USD
Dataset type: Occurrences
Dataset scope: Currently digitizing specimens
Number of records: 1,000
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Neotropical Flora portal
% complete: 13%
Status update: Currently digitizing specimens.
DOI:
Expected date of publication: 2023-02-28

Padre Fuertes collection at Herbarium USD
Dataset type: Occurrences
Dataset scope: Nearly 500 specimens collected by Padre Miguel Fuertes deposited at Herbarium USD.
Number of records: 500
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Neotropical Flora portal
% complete: 0%
Status update: The digitization of this dataset will start in August 2022.
DOI:
Expected date of publication: 2023-04-30

Insect collection IIBZ
Dataset type: Occurrences
Dataset scope: Hispaniola: Dominican Republic. Partially old insect collection of Eugenio Marcano and recentlly collection from several entomologists. Some small subsets: Blattodea, Chrysomelidae, butterflies, Orthoptera.
Number of records: 2,000
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Ecdysis portal
% complete: 0%
Status update: Currently selecting specimens to be digitized. Digitization will start in July 2022.
DOI:
Expected date of publication: 2023-04-30

Odonata
Dataset type: Occurrences
Dataset scope: Hispaniola: Dominican Republic. Actual specimens of adult Odonata from the IIBZ Entomology Collection included in the paper http://www.scielo.org.co/pdf/acbi/v41n111/0304-3584-acbi-41-111-72.pdf
Number of records: 2,000
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Ecdysis portal
% complete: 10%
Status update: Currently digitizing specimens.
DOI:
Expected date of publication: 2023-04-30

Ants
Dataset type: Occurrences
Dataset scope: Dominican Republic. Actual (2017) ants collected using various sampling techniques

in urban parks in Santo Domingo city (FONDOCyT research project on ant as a model in biodiversity assessment in urban areas).
Number of records: 400
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Ecdysis portal
% complete: 100%
Status update: Published more occurrences than what was compromised. Many more occurrences are being added to this dataset.
DOI: https://doi.org/10.15468/h6453y
Expected date of publication:

Insects from savannah
Dataset type: Occurrences
Dataset scope: Dominican Republic. Insects associated to high elevation savannah of Danthonia domingensis in four protected areas, 2014-2015. (FONDOCyT research project).
Number of records: 500
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Ecdysis portal
% complete: 0%
Status update: The dititizing of this dataset will start in July 2022
DOI:
Expected date of publication: 2023-01-31

Molluscs
Dataset type: Occurrences
Dataset scope: Hispaniola: Dominican Republic. Land and freshwater snails and bivalves from historical and actual malacological collection of IIBZ (1958-2000).
Number of records: 400
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, InvertEBase portal
% complete: 0%
Status update: Half of the records are being digitized and under their final cleaning process, so they can be databased in its Symbiota portal.
DOI:
Expected date of publication: 2022-11-30

Aquatic macroinvertebrates
Dataset type: Sampling Event
Dataset scope: Hispaniola: Dominican Republic. Actual data from the project intended to adapt the BMWP and IBF biotic index for monitoring water quality in Dominican rivers.
Number of records: 500
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Instituto de Investigaciones Botánicas y Zoológicas
% complete: 0%
Status update: The digitization of this dataset will start tin August 2022.
DOI:
Expected date of publication: 2023-02-28

Arachnids Collection MNHNSD
Dataset type: Occurrences
Dataset scope: Occurrence data of Arachnids (non-Acari) collected in Dominican Republic from 1970s to actual time.
Number of records: 3,000
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Symbiota, Ecdysis portal
% complete: 27%
Status update: Published, but incompleted.
DOI: https://doi.org/10.15468/v9pt63
Expected date of publication: 2022-12-15

Herpetological Collection MNHNSD
Dataset type: Occurrences
Dataset scope: Occurrence data of Amphibians and Reptiles collected in Dominican Republic between 1970s to date.
Number of records: 2,000
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
% complete: 28%
Status update: About 28 % of the occurrences have been digitized.

DOI:
Expected date of publication: 2023-04-30

Orthoptera Collection MNHNSD
Dataset type: Occurrences
Dataset scope: Occurrence data of Orthopterans collected in Dominican Republic from 1970s to date.
Number of records: 500
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Symbiota, Ecdysis portal
% complete: 16%
Status update: Currently being digitized.
DOI:
Expected date of publication: 2023-02-28

Types of Herbarium USD
Dataset type: Checklist
Dataset scope: Checklist of the type specimens of the Herbarium USD.
Number of records: 93
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Instituto de Investigaciones Botánicas y Zoológicas
% complete: 5%
Status update: Certain progress have been made in the checklist. It is expected to conclude by March 2023.
DOI:
Expected date of publication: 2022-12-30

Ants from urban areas of Santo Domingo
Dataset type: Checklist
Dataset scope: Checklist of species from two urban areas in Santo Domingo.
Number of records: 33
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Instituto de Investigaciones Botánicas y Zoológicas
% complete: 20%
Status update: Certain progress have been made in the checklist. It is expected to conclude by December 2022.
DOI:
Expected date of publication: 2022-12-30

Fishes from Lago Enriquillo
Dataset type: Checklist
Dataset scope: Checklist of fish species from the Lago Enriquillo.
Number of records: 15
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Instituto de Investigaciones Botánicas y Zoológicas
% complete: 0%
Status update: The digitization of this dataset will start in July 2022.
DOI:
Expected date of publication: 2022-12-30

Padre Fuertes Collection
Dataset type: Checklist
Dataset scope: Checklist of plants of Padre Miguel Fuertes deposited at Herbarium USD
Number of records: 524
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Instituto de Investigaciones Botánicas y Zoológicas
% complete: 15%
Status update: This dataset is expected to conclude by March 2023.
DOI:
Expected date of publication: 2023-03-30

Types collection MNHNSD
Dataset type: Occurrences
Dataset scope: Specimens of the type collection of MNHNSD, Dominican Republic.
Number of records: 500
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
% complete: 88%
Status update: This dataset will be published by August 2022.
DOI:

Expected date of publication: 2022-08-31

Fishes from Lago Enriquillo
Dataset type: Occurrences
Dataset scope: Hispaniola: Dominican Republic. Fishes collected during Lago Enriquillo flooded in 2014.
Number of records: 100
Data holder: Instituto de Investigaciones Botánicas y Zoológicas
Data host institution: Symbiota, Consortium of Small Vertebrate Collections portal
% complete: 100%
Status update: Currently 100 % digitized.
DOI:
Expected date of publication: 2022-12-30

Types collection MNHNSD
Dataset type: Checklist
Dataset scope: Checklist of the type collection MNHNSD, Dominican Republic.
Number of records: 148
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
% complete: 0%
Status update: The digitization of this dataset will start in July 2022.
DOI:
Expected date of publication: 2023-03-30

Birds Collection MNHNSD
Dataset type: Occurrences
Dataset scope: Occurrence data of Birds collected in Dominican Republic between 1970s-2000.
Number of records: 400
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Symbiota, Consortium of Small Vertebrate Collections portal.
% complete: 100%
Status update: The validation of this dataset needs to be done.
DOI:
Expected date of publication: 2023-02-28

Bats Collection MNHNSD
Dataset type: Occurrences
Dataset scope: Occurrence data of bats collected in Dominican Republic from 1970s to date.
Number of records: 400
Data holder: Museo Nacional de Historia Natural "Prof. Eugenio de Jesús Marcano"
Data host institution: Symbiota, Consortium of Small Vertebrate Collections portal
% complete: 62%
Status update: Currently 62 % of the records digitized.
DOI:
Expected date of publication: 2023-02-28

Other deliverables

Events

Project BID-GBIF (BID-CA2020-031-NAC) Work Meeting

Dates: 2022-03-04 - 2022-03-04
Organizing institution: Instituto de Investigaciones Botánicas y Zoológicas
Country: Dominican Republic
Number of participants: 22
Comments: This meeting was held with both teams of the IIBZ and the Museum. The idea was to formally introduce the project among the members and explain the goals and purposes to be achieved. The participants from each institution received important information from each coordinator about the role of each member of the project. The progress of the project up to that moment was explained.
Website or sources of verification: List of participants and photographs

Communications and visibility

The visibility of the Project and its activities began with several conferences in congresses (see annexes). We also had the public announcement of the uploading of the first datasets through the

Twitter platform (ttps://mobile.twitter.com/coleopterasam/status/1531318766578789624 32; https://mobile.twitter.com/coleopterasam/status/1517298427540885504) via the collaboration of our Symbiota mentor, Samanta Orellana from Arizona State University.

Another important aspect is that the IIBZ is incorporating data quality procedures into its Procedures Manual. All personnel working at the institute must undergo training on the Darwin Core standard, data cleaning and data collection from the field. This will allow future specimens entering the collection to contain curated data.
Likewise, for students of the Biology Bachelor's Degree of the School of Biology of the Universidad Autónoma de Santo Domingo, who use biodiversity data from our collections or are thinking of entering new field data for their thesis, we will be encouraging the use of the DwC standards and GBIF data.
Photographs: Bastardo Primer proyecto de digitalización -IIBZ.pptx; RuthBastardo_SalaB_09junio_presencial XVIICIC

Monitoring and evaluation

Monitoring and evaluation findings

In the monitoring and evaluation plan, we have been providing routine support by the certified personnel to their respective teams. There is a general coordination and regular meetings with both instructors. As part of the GBIF staff, Leonardo Buitrago has been working closely with us and verifying that the concerns that arise are being adequately addressed. On the other hand, Samanta Orellana, from Arizona State University, has been continuously working on the creation and tracking of the collection profiles to upload the datasets via the Symbiota platform. The profiles are hosted in the portals Neotropical Flora, Ecdysis, Consortium of Small Vertebrate Collections, and InvertEBase. This guarantees better protection for the data of both collections. Regarding hired personnel, there were some delays and changes due to the resignation of three members of the IIBZ team. These assistants were replaced, but as a result, we now have funds to hire one more person. We have also encouraged the participation of student volunteers to collaborate while learning about the process.
No delays are expected for the completion of the deliverables on the scheduled dates. Although some datasets, such as the Catálogo de Plantas del Padre Fuertes, we have left to the end because have found they have a more difficult degree to achieve.
One of the most important steps we hope to be able to take is to move towards the publication of datapapers. We would like to have the support of a mentor for this purpose. We are continuing to work with Leonardo Buitrago, so it would be interesting if he could guide us in this direction.

Impact of COVID-19 pandemic on project implementation

Due to the pandemic, the initial meetings of the work teams of both institutions began virtually. We then moved to in-person meetings. This delayed the scheduled workshop (Workshop and training for the digitization of collections) that was scheduled for September 2021. As a result, the team ended up being trained by the certified personnel. Although there has been a slight delay, this does not represent an obstacle to the achievement of the proposed goals. Thus, we plan to use those available funds, upon authorization, to offer a more comprehensive workshop next year for the general public.

GBIF leads the Biodiversity Information for Development (BID), a programme funded by the European Union. The programme provides supplementary support for activities addressing the needs of regional researchers and policymakers through mobilization and use of biodiversity data.