

National Portals Workshop: taxonomic databases, species data information and visualization

Programme: CESP

Project ID: CESP2020-006

Project page: <https://www.gbif.org/project/4YJIFEvYJi5kfuUVzNcfYH/from-shared-experiences-to-shared-knowledge-and-common-solutions-collaborating-to-improve-web-based-tools-in-latin-america-and-the-caribbean>

Project lead organization: Brazilian National Research and Education Network

GBIF Participant Represented: Brazil

Project implementation period: 1/8/2020 - 20/12/2022

Report approved: 2/3/2023

Narrative Final report

Executive Summary

The project was a challenge from the beginning. The project was approved before the pandemic changing the main objective, which was a face-to-face meeting of the nodes in the LAC region. In this way and trying to keep the event, we started to hold online meetings to get to know the group, expectations and needs. After a year, even without the possibility of a meeting, we readjusted the objectives, divided by modules and readjusted the budget, changing everything to salaries. The module 1 that focused on national taxonomic databases was able to advance and finalize the update of the Brazilian taxonomic database. But in addition to updating, we managed to improve the bank's project, we improved the API to create and the code and tutorial videos will be made available so that any partner country can adapt the code to the needs. For module 2 of Plinian Core, the project supported the progress of the development and implementation of Plinian Core in the Living Atlas and we organized sample datasets for testing in addition to presenting the results in the TDWG event. Finally, we are improving the Living Atlas infrastructure and sharing the results with the ALA community. Finally, it should be noted that although it was a lengthy project, it exceeded expectations with high quality products delivered and a starting point for new challenges as a region. More meetings are needed now to discuss new steps both as a region, which will now be in Guatemala as well as in Living Atlas as each has similar needs that need to be shared and discussed. Finally, the videos will not be released yet, given their relevance in Brazil, it will be necessary to review the project and the APIs and organize an inter-ministerial launch event for all institutions that work with biodiversity.

Report on Activities

Activity progress summary

All activities as well as deliverables are described and attached. In the case of the videos, the Portuguese and English versions are not the final versions, lacking minor adjustments. The rest can be checked on the Project website <https://www.sibbr.gov.br/page/cesp/cesp-2020.html> and on the GitHub of CRBio in Costa Rica <https://github.com/AtlasBiodiversidadCostaRica/doc-bie-pliic>.

Completed activities

Activity name: Module 1. Videos. Storyboard

Description: Video script in Portuguese, Spanish and English

Start Date - End Date: 1/9/2022 - 30/9/2022

Verification Sources: Attached.

Activity name: Module 1. Videos. Slides

Description: Slides and video content in Portuguese, English and Spanish

Start Date - End Date: 1/9/2022 - 30/9/2022

Verification Sources: Attached

Activity name: Module 1. Videos edit

Description: Videos editing

Start Date - End Date: 1/10/2022 - 31/12/2022

Verification Sources: Attached in spanish

Activity name: Module 2. Plinian Core. TDWG

Description: Participation in TDWG.

Start Date - End Date: 1/10/2022 - 31/10/2022

Verification Sources: <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Activity name: Module 2. Species dataset

Description: Organization of a dataset with species description information

Start Date - End Date: 1/9/2022 - 30/9/2022

Verification Sources: Attached

<http://ipt-test.crbio.cr:8080/ipt/>

Report on Deliverables

Deliverables progress summary

All deliverables are available on SiBBR's Youtube channel (<https://www.youtube.com/user/SiBBR>) and SiBBR website <https://www.sibbr.gov.br/page/cesp/cesp-2020.html> and CRBio github available at <https://github.com/AtlasBiodiversidadCostaRica/doc-bie-plic>.

Module1

Once it was decided that video tutorials would be made with step-by-step instructions for building a national taxonomic database based on the example of Brazil, the company that was helping to update the Brazilian taxonomic database was asked to make part of the documentation in videos for facilitate understanding for people who are not experts in data engineering projects. Five videos were defined with the most important step by step from the choice of tools, the consumption of original sources, and the development of two APIs, the first that creates a raw taxonomic database per kingdom and the API that creates an integrated taxonomic database by joining all kingdoms in a single database. The last video is for generating views and inserting other data such as endemism and threat lists. The SiBBR team, with the preliminary videos, the structure and content of the videos and the most relevant information, prepared an script and a storytelling and some slides and delivered it to a company hired with the project's money to prepare the final videos. Once approved, the script was translated into Portuguese and English and subtitles were inserted. The videos and the code will be released by YouTube in 2023 once the taxonomic database is launched in Brazil. Unfortunately, the video company took longer than expected to deliver the videos and insert the subtitles, making it necessary to revise everything several times. In addition, during the database update process, several inconsistencies were found in the data that required validation and revision, with some adjustments and validations still necessary. It should be noted that the videos are for understanding the main step-by-step and where all the information is available so that each partner can later develop their own taxonomic database construction project.

Module 2

Part of the resource was passed on to the team of Costa Rica, which was responsible for coordinating Plinian Core activities from ratification and presentation of results in TDWG to the development of the indexer for these files to BIE-INDEX for publication of species information on the Living Atlas platform. With the transfer of the amount to Costa Rica, the technical staff got support to speed up the process by hiring support to achieve the project's results. The entire process is available on CRBio's github and the documentation will soon be available for consultation on GitHub as well.

Module3

Module 3, which was focused on Living Atlas, is a continuous and parallel activity being carried out by SiBBR's technical team. The team is automating the execution of several routine processes on SiBBR's Living Atlas platform. These processes were summarized in a presentation made by SiBBR called "Self-service Living Atlas in containers, Kubernetes and Rancher". With only using co-funding resource and in continuous contact within the support of GBIF Spain programmer and living atlas focal point, Vicente J. Ruiz, we are updating and disseminating how the installation and infrastructure that houses the Living Atlas platform was conceived and designed from SiBBR. All these advances are available on the ALA - Brasil GitHub. The presentation is available for download on the SiBBR website. Finally, it is from this presentation held on December 16th during the closing of the project that the possibility of holding a face-to-face meeting was discussed to discuss priorities, news and data entry with the countries that implement the Living Atlas.

Completed deliverables

Title: Building a national Taxonomic Database . Video 1

Type: Documentation

Description: Tools, infrastructure and programming languages to build a national Taxonomic Database This video will present the recommended tools and infrastructure for building a taxonomy database from several previously consolidated and validated species lists.

Sources of verification: Attached in documents. Soon available on the SiBBR youtube channel

Title: Building a national Taxonomic Database . Video 2

Type: Documentation

Description: Taxonomic Database of raw data Original data sources. In this video tutorial, we present the different original data sources and national reference lists that were used to consolidate the taxonomy database of Brazil and make it available for other systems.

Sources of verification: Attached in documents. Soon available on the SiBBR youtube channel

Title: Building a national Taxonomic Database . Video 3

Type: Documentation

Description: Taxonomic Database of raw data API - Taxonomy – Bruto Flora complementary. This video tutorial describes how the raw database is created from the API-Taxonomy application. In addition, the script used for Flora's complementarity will be detailed.

Sources of verification: Attached in documents. Soon available on the SiBBR youtube channel

Title: Building a national Taxonomic Database . Video 4

Type: Documentation

Description: Integrated Taxonomic Database API - Taxonomy - Integration. In this video we will talk a little more about updating the taxonomy database of Brazil, when it comes to the integrated database, that is, the integration of all the reference sources that were mentioned in the previous videos

Sources of verification: Attached in documents. Soon available on the SiBBR youtube channel

Title: Building a national Taxonomic Database . Video 5

Type: Documentation

Description: Taxonomic Integrated Database Generating views. Adding new data and checklists This video is intended to show the complete integrated database and the export of the database in Darwin Core for publication at IPT or for consumption on platforms such as Living Atlas.

Sources of verification: Attached in documents. Soon available on the SiBBR youtube channel

Title: Plinian Core

Type: Documentation

Description: All presentations, webinars, TDWG abstract and code development in github

Sources of verification: Module 2 in <https://www.sibbr.gov.br/page/cesp/cesp-2020.html> and CRBio GitHub <https://github.com/AtlasBiodiversidadCostaRica/doc-bie-plic>

Title: Module 3. Living Atlas of SiBBR. Infrastructure Kubernetes

Type: Documentation

Description: Presentation of how is Living Atlas installed and maintained.

Sources of verification: Module 3 in <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Events

Module1. Webinar 1. National Taxonomic databases. Countries using Living Atlas platform

Dates: 13/5/2021 - 13/5/2021

Organizing institution: SiBBR, CRBio (Costa Rica), CONICET (Argentina)

Country: Brazil

Number of participants: 40

Comments: Webinar 1, held on May 13, 2021 and moderated by Brazil, had the participation of the

Events

Module 1. Webinar 2. National Taxonomic databases. Countries with National Portals

Dates: 27/5/2021 - 27/5/2021

Organizing institution: SiBBr, SibColombia, MMA Chile, CONABIO (mexico)

Country: Brazil

Number of participants: 40

Comments: The second webinar, held on May 27, 2021, featured presentations from Conabio from Mexico, SiB Colombia and representatives from the Ministry of the Environment from Chile.

Website or sources of verification: <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Events

Module 1. Webinar 3. National Taxonomic Databases. International Taxonomic Databases

Dates: 28/9/2021 - 28/9/2021

Organizing institution: SiBBr

Country: Brazil

Number of participants: 70

Comments: For webinar 3, we invited three experts to present how they organize taxonomic data on other international platforms.

Website or sources of verification: <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Events

Module 2. Plinian Core

Dates: 14/9/2021 -

Organizing institution: SiBBr, SibColombia, Costa Rica

Country: Brazil

Number of participants: 70

Comments: With the participation of speakers from Costa Rica and Colombia, the second webinar was held in Spanish on May 14, 2021. It was moderated by Brazil and had the support of the Atlas of Living Costa Rica (CRBio) and the Information System on Biodiversity in Colombia (SIB Colombia).

The teams of these systems follow the development, update and rectification of the Plinian Core standard in the Taxonomic Databases Working Group (TDWG), responsible for the development and update of the Biodiversity standards for publication and data sharing. The presentations and discussions held at the Webinar enabled a better understanding of the Plinian Core standard, the status of the approval and implementation of the standard and the respective extensions in the IPT (Integrated Publishing Toolkit) - GBIF's tool for publishing data. Furthermore, the CRBio team is in the process of improving a module, aimed at consuming and viewing information about species, which will use the Plinian Core standard on the Living Atlas platform.

Website or sources of verification: <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Events

Final Webinar

Dates: 16/12/2022 - 16/12/2022

Organizing institution: SiBBr, CRBio

Country: Brazil

Number of participants: 30

Comments: Final webinar with all results from the three modules.

Website or sources of verification: <https://www.sibbr.gov.br/page/cesp/cesp-2020.html>

Communications and visibility

All information is published in <https://www.sibbr.gov.br/page/cesp/cesp-2020.html#s>. On the SiBBr page you will find all the information, presentations and access to the webinars directly on the SiBBr Youtube channel. As it was a project that, due to the pandemic, lasted much longer than planned, it allowed us to strengthen partnerships and more contact with partners in Latin America. During the last webinar and during the presentations of the results of the different modules, it became clear the individual potential of each country as well as the potential as a region. We need more opportunities to meet and exchange ideas and knowledge both with the countries that implement the Living Atlas and with the rest of the countries. We need more meetings as a group and we believe that the next face-to-face meeting in Guatemala is an excellent opportunity to continue producing content. Examples of taxonomic databases, official lists and sharing of information about species with Plinian Core show the

progress of different Latin American countries.

Monitoring and evaluation

Monitoring, evaluation and lessons learned

Final Evaluation:

The project brought great results and exceeded the expectations of the different nodes. Instead of a face-to-face meeting in São Paulo, to discuss about national portals and sharing information about taxonomic data and the species descriptions, it resulted in a project where all the resources were destined to produce material for the organization and publication of data on species, whether taxonomic or descriptive. All countries, although they are at different stages when it comes to organizing species, will be able to use the project results as subsidies for organizing information on national biodiversity. The project, in addition to exchanging information on the challenges and results of each country in the region, brought several presentations by international specialists on how to organize taxonomic databases. The advantage of dividing the project into three major modules allowed to subdivide deliveries and divide responsibilities, with Brazil responsible for modules 1 and 3 and Costa Rica and Colombia responsible for module 2. In this way, it was possible to organize deliveries and produce better results. Regarding module 1, the videos are subtitled in Spanish and English and can reach a large number of audiences around the world and although it is not necessary to follow exactly the example of the Brazilian national taxonomic database, some of the tools and codes disclosed may be used or adopted to the needs of different countries. The resource destined to module 2 on Plinian Core accelerated the development and integration of the Plinian Core standard to the BIE module of the ALA, as well as allowing tests to be carried out with spreadsheets on descriptions of species from Costa Rica and Brazil being published in the IPT test with the Plinian extension implemented. Regarding module 3, Brazil presented the infrastructure used to maintain and operate the Living Atlas platform. The project strengthened partnerships, produced a lot of very relevant material that will make it possible to advance along several lines, such as creating an official list of species for Latin America, sharing information on common species on the continent, and made it clear that countries that use Living Atlas such as Brazil and Costa Rica has much to offer the Living Atlas community and a meeting between these countries is essential for the years to come.

Best Practices and Lessons Learned:

This project must have been one of the most affected by the pandemic. The project was completely adapted, rewritten and approved again. Although it was a challenge and the results exceeded expectations, it is important to note that the process was time-consuming and complex and the level of commitment varied among project partners and although it was to be expected some decisions had to be made between smaller groups. A point in favor, with so many nodes from the LAC region, it allowed us to get to know each other better as well as understand the challenges and particularities of each country. For the next meeting of the LAC region, we will have advanced and produced material to continue working as a region. The pandemic has shown the potential of online events, making it possible to bring international guests and experts to contribute and enrich the events. However, someone with experience in using zoom and live streaming on other platforms is always needed. The realization of webinars and the rapid dissemination of materials, presentations and videos allow for greater monitoring of the progress of the project and also the possibility for other people to be able to follow the results. Finally, in this project, where all the resources that were previously defined to be spent on travel were all allocated to salary payments in partnership with Costa Rica, a fact that allowed the hiring of support to accelerate activities that were planned but that would take a while to finish, thus allowing to have tangible and very important results as a region and as sharing data and descriptive taxonomic information giving the deserved relevance for this type of work.

Post Project Activity(ies):

Some activities still need to be finalized. Videos subtitled in Spanish were shared with project partners and are attached to this report. The videos will not yet be released in Brazil given the relevance of the project. The national taxonomic database will be launched in 2023 at an interministerial event where it will be presented to the rest of institutions that work with biodiversity in order to establish a partnership in the use of the database and ensuring the sustainability of this result. The code needs revision and support for minor adjustments before releasing in Brazil. We hope that the launch of this result in Brazil results from a partnership to keep the database updated and working with the Ministry of the Environment and to strengthen the original Brazilian data sources that need support to keep the catalogs of flora and fauna working. Regarding module 2, Plinian Core, the documentation needs to be finalized. Once organized, we will ask the RNP team to implement the code in homologation to carry out other tests and, in parallel, we will be organizing the species information for publication in the IPT test. In addition, the importance of organizing a face-to-face event to talk about the Living Atlas is highlighted. During the project closing event, many questions and doubts arose about how data managers deal with some problems. This exchange of information is essential to better understand how the platform works, facilitating daily work. In this way, although the project ends, a launch event will still be held in Brazil.

Sustainability plans

As mentioned in the post project activities, it will be necessary to launch the videos and taxonomic database in Brazil with the different institutions that deal with biodiversity at a country level to guarantee the use and update of a unique database that is not linked to a platform but to the government. The code will be revised, validated with those responsible for the Brazilian flora and fauna catalogs and the lists of viruses and microorganisms validated with the Brazilian societies of virology and microbiology. The publication of species information will be carried out by groups and prior authorization will be requested from the providers of this information. In addition, regarding module 3, it will be disclosed to the Living Atlas community and we will talk with the Living Atlas focal point to try to organize a face-to-face event later this 2023.

Impact of COVID-19 pandemic on project implementation

The COVID-19 pandemic had a major impact on the implementation of the project since the main activity consisted of organizing and implementing a face-to-face event. We postponed the event by monitoring the data and progress of vaccination. However, the pandemic did not allow a face-to-face meeting or travel. In this way, already in 2022 it was decided together with all the partners that the resources will be invested in tutorials such as videos and documentation and other online workshops. Other activities include support for the team that is working with the implementation of the Plinian Core standard. Finally, Brazil also prepares complementary documentation for those interested in implementing the Living Atlas platform and will organize an online workshop to present.

The Capacity Support Enhancement Programme (CESP) is funded by GBIF

