



**GBIF Work Programme 2022**  
***Annual Update to Implementation Plan 2017-2022***

Version 1.0, 2021-10-26

# Table of Contents

Colophon	1
Suggested citation	1
Persistent URI	1
Document control	1
Overview	1
GBIF Annual Work Programme 2022	2
Financial situation	2
Planned expenditure	4
Secretariat staffing	5
Overview of 2022 Work Items	5
2022 Implementation Plan Update	11
Priority 1: Empower Global Network	11
Activity 1a: Focus on people	11
Activity 1b: Strengthen skills	14
Activity 1c: Equip Participant nodes	17
Activity 1d: Equip data publishers	21
Activity 1e: Expand national participation	24
Activity 1f: Plan implementation	28
Activity 1g: Coordinate resources	30
Priority 2: Enhance Biodiversity Information Infrastructure	34
Activity 2a: Modernize data standards	34
Activity 2b: Deliver names infrastructure	38
Activity 2c: Catalogue collections	41
Priority 3: Fill Data Gaps	44
Activity 3a: Identify priority gaps	44
Activity 3b: Expand data streams	46
Activity 3c: Engage data holders	49
Activity 3d: Rescue datasets	52
Activity 3e: Liaise with journals	53
Priority 4: Improve Data Quality	55
Activity 4a: Ensure data persistence	55
Activity 4b: Assess data quality	57
Activity 4c: Enable data curation	59
Priority 5: Deliver Relevant Data	61
Activity 5a: Engage academia	61
Activity 5b: Document needs	64
Activity 5c: Support biodiversity assessment	67
Activity 5d: Assess impact	70

# Colophon

## Suggested citation

GBIF Work Programme 2022: Annual Update to Implementation Plan 2017–2022: GBIF Secretariat.  
<https://doi.org/10.35035/doc-jjrz-b144>

## Persistent URI

<https://doi.org/10.35035/doc-jjrz-b144>

## Document control

Approved version

## Overview

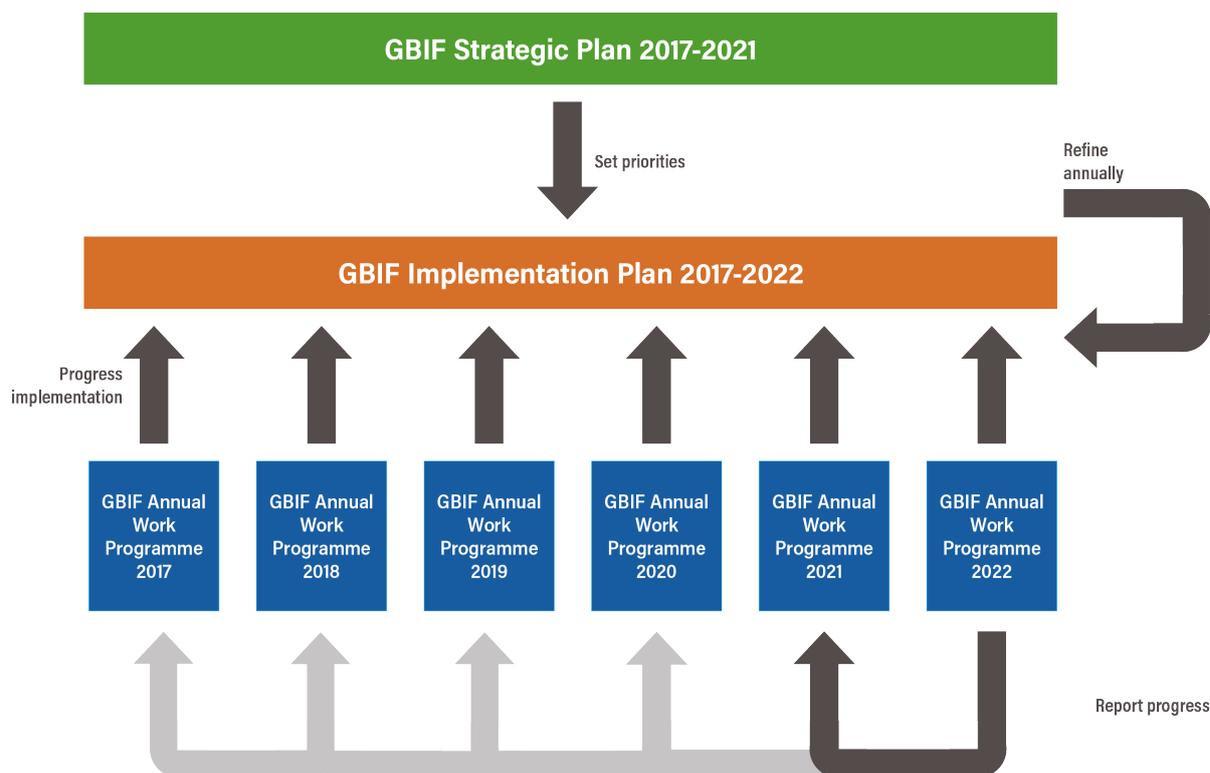
This document serves both as the 2022 revision of the GBIF Implementation Plan for the period 2017-2022 and also as the GBIF Annual Work Programme 2022. Both elements are based upon the [GBIF Strategic Plan for 2017-2021](#), which was extended for one year due to the challenges created by the global pandemic.

The GBIF Annual Work Programme 2022 represents the final programme based on the current strategic plan. The new strategic planning framework for the period 2023-2027 was approved at the 28th GBIF Governing Board meeting (GB28) in October 2021, and it will form the basis of a new implementation plan to be developed during 2022.

The Strategic Plan 2017-2021, agreed at the GB22 meeting in 2015, established five priorities for GBIF's work. The Implementation Plan identified 22 activities which GBIF must undertake to address these priorities during the six-year period.

Each year, GBIF has reviewed progress against these activities and available resources (including Secretariat staff and other uses of core and supplementary funding, as well as resources committed by GBIF Participants). Using this information, an Annual Work Programme is defined for the following year. Each Annual Work Programme identifies a prioritized set of work items during the year. Annual Work Programmes in the second and subsequent years have also reported on progress against the plan.

The relationship between these documents is represented by the following graphic.



The Work Programme is presented here with updates reflecting progress during 2021 and planned activities for 2022.

## GBIF Annual Work Programme 2022

Information on activities planned for 2022 is included under each of the Activity sections included in the 2022 implementation plan update below. For convenience, this information has also been collated as an [Overview of 2022 Work Items](#) listing the items for which approval is sought from the Governing Board at GB28.

### Financial situation

As of September 2021, current and imminent voting participant contributions are slightly above expected contributions approved in the GB27 budget. This targeted total, however, takes into consideration expected inability of some countries to pay contributions. Therefore, the total financial contribution provided by countries remains below the goal set forth in the financial model approved at GB22.

The 2021 budget approved by GB27 included funds for travel and in-person meetings in 2021 (including regional meetings and GB28), but very little of this was used, and these budget items were downscaled in a mid-year revision. Partly due to these savings, €200K budgeted within the 'COVID-19 reserve' was deemed not to be needed, and following mid-year approval by the Executive Committee the Secretariat removed this level of reserve from the budget.

The 2021 Work Programme budget items were adjusted with the overall budget level similar to the budget adopted in GB27. Changes approved by the Budget Committee and Executive Committee for

each Strategic Plan Priority are listed below. Information on each new or amended item is included in the 2021 progress reported for all activities.

## 2021 expenditure on external contracts and workshops

In addition to salary and operating costs (including support for GBIF Governance mechanisms), funding was allocated from core funds to the 2021 Work Programme. These projects and amounts were updated throughout 2021 with EC approval. Estimates for total 2021 spending for the following activities are:

### Priority 1:

- Communications to promote value proposition responses to the 20 review (€30K).
- Support for Living Atlas Activity (€20K).
- *alliance for biodiversity knowledge* (€10K).
- Capacity Enhancement Support Programme (€80K).

### Priority 2:

- Consultancy to begin development of improved data models (€30K).
- Contracted Web Development (€20K).
- To support contractor and costs to maintain and update processes for constructing the GBIF taxonomic backbone in conjunction with COL+ (€108K).

### Priority 3:

- Human disease vector data mobilization campaign (€25K)
- Soil biodiversity data mobilization campaign (€25K)

### Priority 5:

- Up to €80,000 for a quantitative analysis of the value and impact of the GBIF network

## Supplementary income

The following additional expenditure was enabled by supplementary funds received or budgeted during 2021:

- JRS Biodiversity Foundation additional contribution to African Biodiversity Information for Development (BID) projects: €411K
- Biodiversity Community Integrated Knowledge Library (BiCIKL) project (EU Horizon 2020): Developer costs €66K
- Support for Asia regional support contracts through SYNTHESYS+ (European Union): €109K
- Salary subsidies (BID, SYNTHESYS+, DiSSCo Prepare and BiCIKL): increased to €152K

As a result of the past year's staffing and Work Programme changes, GBIF is expected to close 2021 with net capital of around €398,181, with an additional €461,000 in the COVID-19 Reserve.

The 2022 proposed budget includes a smaller allocation (€400K) for contracted activities in the Work Programme expenditure (i.e. non-salary costs allocated to calls, workshops, contracts, etc.) than currently budgeted for 2021. The proposed 2022 budget uses just over half (€261K) of the COVID-19 Reserve. The planned expenditure for 2022 is €188,155 greater than the income projected in the budget. This is considered to represent a responsible balance between advancing GBIF's mission and reserving some funds against future risks (€471,026, net capital forecast for the end of 2022).

GBIF is a maturing infrastructure and network of people. This maturation has enabled infrastructure improvement, inclusion of new data types and stronger global engagement. It also comes at an increased cost to maintain the infrastructure and engagement at current levels. The current core budget is based on €3.5 million from Voting Participants, but the amount actually received is typically less than this. Staff growth in past years means that staff salaries and basic operational costs (including insurance) are now similar to or exceed Voting Participants' contributions. This leaves little room for increased feature development and engagement using core funds.

To meet the development needs, the Secretariat has initiated a resource mobilization strategy to meet these needs with supplementary funding. The Secretariat is confident that the value presented by BID and other supplementary programmes, present infrastructure and data quality improvements will be translated into a strong external funding stream.

## Planned expenditure

In addition to salary and operating costs (including support for GBIF Governance mechanisms), core funding is reserved in the Work Programme 2022 for the following activities:

### Priority Area 1

- Commission risk assessment study for sensitive species data (€15K)
- Launch a new call for proposals under the Capacity Enhancement Support Programme (€80K)
- Support further development of the Living Atlases community (€15K)
- Support for 2022 Regional Nodes Meetings (€32K) - note that some meetings will be supported from supplementary funds or funds reserved from previous budgets
- Support for the *alliance for biodiversity knowledge* (€20K)

### Priority Area 2

- Update GBIF data models (€40K)
- Maintain and update processes for constructing the GBIF taxonomic backbone. €108K has been allocated in the budget to support GBIF contractor costs
- Web Developer contract (€30K)

## Priority Area 3

- Support for thematic data mobilization (€40K)

## Priority Area 5

- Support for phylogenetic integration with occurrence data (€20K)

## Other Funds

- Governance funds: Launch 2022 calls for the Ebbe Nielsen Challenge and Young Researchers Awards (€25K)

## Secretariat staffing

The Secretariat is in the process of hiring two new staff members. In response to GB value proposition discussions and the recommendations of the CODATA review of GBIF's first 20 years, the Secretariat has hired a new communications officer to begin in October 2021. This new staff member is part of the Participation and Engagement staff. We currently are recruiting an administrative assistant to supplement our capacity in our administrative staff. These new hires will bring our staffing level to 30 staff, three of whom work part-time. In addition to the Director and Deputy Director, the Secretariat is structured as four teams, as follows:

- Participation and Engagement (nine staff)
- Data Products (four staff)
- Informatics (eight staff)
- Administration (seven staff)

All staff work towards accomplishing the goals of the five Priority Areas in the 2017-2022 Implementation Plan, with most staff working towards multiple areas. The Secretariat internally reports salary to the Work Programme Priority Areas and Secretariat management based on an estimate of staff time spent on the areas.

## Overview of 2022 Work Items

The following is a summary of all 2022 Work Items proposed in the 2022 update of the Implementation Plan below. It is intended to summarize the work items for which approval is sought from the Governing Board at GB28. Accordingly, it does not include the additional commitments made and reported by Participants in the detailed activity sections.

### Activity 1a: Focus on people

- Improve representation of individuals on GBIF.org, with more coherent and integrated handling of the roles and profiles of Secretariat staff and network members, including biographies, activities and other related information

## Activity 1b: Strengthen skills

- Develop new training module on data quality management, targeted at GBIF nodes
- Monitor impact of GBIF training programmes, based on uptake of existing training resources, skills certification through digital badges, replication of courses throughout the network, and outputs of data and other products linked to participation in training
- Further develop joint guidance with OBIS on publication of marine datasets through both networks

## Activity 1c: Equip Participant nodes

- Mature the technical capabilities and community processes to expand the hosted portal service, focusing on the needs of GBIF participant nodes. This activity assumes positive feedback from participants of the pilot program. The expanded programme will determine the level of service that will be available to each node (including number of portals), and possible additional services that would be available through extra financial contributions
- Continue support for the Living Atlases collaboration through extended contract for the technical coordinator (€15K)
- Support in-person regional nodes meetings for 2022, if permitted by the travel situation (€32K from core funds)
- Actively explore supplementary funding for additional regional support contractors in Latin America, the Pacific, and E Europe/Caucasus/Central Asia
- Refresh existing guidance materials for GBIF nodes and associated support, including node-level mentoring, capacity self-assessment and training for national-level project teams
- Develop guidance on best practices for working with indigenous and local knowledge, including application of the CARE principles, following recommendations of the working group established during 2021

## Activity 1d: Equip data publishers

- Building on the 2021 work of the sensitive species working group, commission an analysis of the current scale of sensitive species data published through GBIF, evaluation of the evidence that such data presents risks of harm, and recommendations for applying a risk assessment approach (€15K)
- Based on work carried out in 2021, develop plans for the future of the IPT
- Implement new technical documentation website

## Activity 1e: Expand national participation

- Review the current status and models of participation in GBIF both at national and organization level, with a view to developing recommendations for the new strategic implementation period at GB29
- Take a more proactive role with HoDs through global science collaboration activities such as the Group of Senior Officials

- Use the delayed agreement of the post-2020 Global Biodiversity Framework to further emphasize GBIF's value to countries and promote expanded participation

### **Activity 1f: Plan implementation**

- Organize the 3rd Global Biodiversity Informatics Conference (GBIC3), hosted by the LifeWatch European Research Infrastructure Consortium (ERIC), to address the ongoing operation and governance of the *alliance for biodiversity knowledge* (€10K)
- Organize and run the 2nd SYNTHESYS+ workshop relating to citation of collections

### **Activity 1g: Coordinate resources**

- Launch a call for proposals under the 2022 Capacity Enhancement Support Programme (€80K)
- Following approval of the resource mobilization strategy at GB28, develop implementation road map and begin implementation, with a view to targeting additional supplementary funds as well as consolidating core contributions from Voting Participants
- Subject to an improved travel situation in 2022, organize an in-person engagement meeting in Asia, using supplementary funds from BIFA (€20K) and the SYNTHESYS+ project (€19K), co-located with the 2022 Asia regional nodes meeting

### **Activity 2a: Modernize data standards**

- Continue to participate in the ongoing Open Digital Specimen and Extended Specimen Network standards development and strive towards a common solution
- Together with the Atlas of Living Australia (ALA), support the Living Atlas community, in adoption of the shared data pipelines
- Following open review of the revised data model, prototype new infrastructure that allows the demonstration of data mapping, publishing and indexing into new catalogues, developed in an open manner. This will involve development in the IPT, new data pipelines, APIs and user interfaces to explore content and communities of data publishers willing to shape data for the pilot. It is envisaged one or more developer workshops will be arranged to encourage more open development and collaboration
- Ensure all vocabularies are included in the GBIF Registry and maintained in a collaborative manner

### **Activity 2b: Deliver names infrastructure**

- Complete outstanding tasks to integrate the extended Catalogue of Life in GBIF, partly within the framework of the BiCIKL project
- Maintain and update processes for constructing the GBIF Backbone Taxonomy, including monitoring the content and helping to prioritize editorial effort. €108K has been allocated in the budget to support GBIF costs. This work is in collaboration with Catalogue of Life
- Explore feasibility of supporting national taxonomies for exploring GBIF occurrence data to better enable national-level reporting (carried over from 2020 and likely to be incomplete in 2021)

- Continue and expand the taxonomy curation pilot projects in collaboration with COL and the taxonomic community

### **Activity 2c: Catalogue collections**

- Complete outstanding tasks to deploy an enriched GRSciColl providing search and access of collections, specimens and people
- Focus on content of GRSciColl: cleanup of existing entries and registration of new ones by promoting use and giving training and support to editors, and promoting consistent use of codes within data shared
- Seek to identify links between journal articles and collections based on the collection codes, within the framework of the EU-funded BiCIKL project
- Support user interface translations for GRSciColl
- Carried over from 2020: Explore synchronization of content with the Consortium of European Taxonomic Facilities (CETAF) Registry (under development)

### **Activity 3a: Identify priority gaps**

- Publish updated best practices on prioritizing data mobilization based on data gap analysis, for incorporation into guidance for nodes and data publishers
- Explore the use of data gap analysis to prioritize support for data mobilization through funded programmes coordinated by GBIF, and to document the value added by existing projects
- Work with partners to address the information needs of the post-2020 Global Biodiversity Framework through filling priority data gaps, including for those countries not yet participating in GBIF

### **Activity 3b: Expand data streams**

- Building on the framework developed in 2021, establish data mobilization campaigns for additional thematic data streams (possibly freshwater, DNA-derived data and tropical biodiversity) (€40K)
- Seek supplementary funding to support scaling up of support for thematic data mobilization throughout the GBIF nodes network (see also Activity 3c)
- Continue working with partners to increase mobilization and improve visibility of biodiversity monitoring datasets in GBIF

### **Activity 3c: Engage data holders**

- Building on preparatory work and partnerships developed in 2021, scale up mobilization of primary biodiversity data from the private sector, through the Data4Nature initiative and establishment of help desk services
- Engage additional sectoral associations, financial and other initiatives to integrate biodiversity data-sharing commitments into best practices for Environmental, Social and Governance (ESG) or regulatory requirements

### **Activity 3d: Rescue datasets**

- Continue to implement workflow for prioritizing and drawing upon potential data sources reported through the 'suggest a dataset' tool, including involvement of nodes, mentors and crowdsourced solutions (carried over from 2021)
- Based on the outcome of the re-evaluation, work with interested parties to roll out a workflow for processing data mobilization requests arising from the 'suggest a dataset' tool

### **Activity 3e: Liaise with journals**

- In consultation with journal publishers, explore solutions to enable academic authors to submit and publish datasets to GBIF alongside research articles
- Further develop the model for sponsored data papers, in connection with mobilization campaigns from thematic data streams
- Continue BiCIKL collaboration to improve integration of data extracted from research literature

### **Activity 4a: Ensure data persistence**

- Support adopters of cloud technologies for accessing GBIF-mediated data, and revise procedures and formats as necessary.
- Explore opportunities to strengthen integration of the GBIF network within research data infrastructures

### **Activity 4b: Assess data quality**

- Implement improvements to data validation reports identified in community consultation
- Promote integration of the data validator in tools connected to GBIF, including the IPT, through use of the validator API

### **Activity 4c: Enable data curation**

- Explore what can be enabled to offer annotation services in GBIF.org following ideas offered during the consultation topic during 2021 (see Activities 1f and 2a).
- Continue to explore the use of the GBIF data index to support stable persistent resolvable identifiers for all specimens and occurrence records (see Activity 2a)

### **Activity 5a: Engage academia**

- Consolidate training resources aimed at academic users and publishers of GBIF-mediated data, for use in funded programmes and guidance for nodes
- Support implementation of the BioDATA and BioDATA Advanced projects through partners in Norway, Ukraine, Armenia, the Russian Federation and South Africa
- Promote GBIF at relevant fora including the 9th International Barcode of Life Conference, scheduled for August 2022, and the International Congress of Entomology, scheduled for July 2022

- Begin pilot project to integrate phylogeny with occurrence data (€20K)

### **Activity 5b: Document needs**

- Coordinate work of the expert group on mobilization and use of data on zoonotic diseases based on outputs from the systematic review commissioned in 2021
- Continue programme for mobilization of soil biodiversity data through partnerships and work streams initiated in 2021
- Continue partnership with the IUCN Invasive Species Specialist Group and the Convention on Biological Diversity Secretariat to improve content and curation of data relating to invasive alien species

### **Activity 5c: Support biodiversity assessment**

- Through the *alliance for biodiversity knowledge*, hold a workshop on the outcomes of the analysis of primary data needs for the post-2020 Global Biodiversity Framework, involving partners in GEO BON, the Biodiversity Indicators Partnership, CBD Secretariat and others
- Implement a process enabling key checklists to be used in filtering occurrence data, such as Red Listed species and invasive alien species (carried over from 2020 and likely to be incomplete in 2021)
- Work with the new secretariat of GEO BON to improve connections between Biodiversity Observation Networks (BONs) and nodes at national, regional, thematic and global levels
- Continue collaboration with IUCN to implement the Memorandum of Cooperation, including exploration of publication of reference datasets associated with Red List assessments, and development of training material for use of data in the Key Biodiversity Area (KBA) process
- Continue to engage with the IPBES Data and Knowledge Task Force to support use of GBIF-mediated data in ongoing and upcoming assessments

### **Activity 5d: Assess impact**

- Publish quantitative analysis of the value and impact of the GBIF network, and promote in engagement with governments, funders and other partners
- Launch Ebbe Nielsen Challenge and Young Researchers Awards competitions for 2022 (€25K), and generate communications connecting past awardees with subsequent work and impact
- Explore options for a 'Nodes Award Scheme' to recognize progress, performance and innovation across the community
- Produce communication materials including short videos demonstrating the impact of GBIF's funded programmes such as BID and BIFA
- Complete and promote communication materials demonstrating GBIF's impact across thematic communities, thereby supporting the case for increasing the number of Voting Participant countries

# 2022 Implementation Plan Update

Note that the 'Rationale' section in each activity is presented with the original wording from the Implementation Plan in 2017, to show the original intention of the activity. Updates on Participant contributions and Participant plans provided by the nodes are included in this final version for review alongside the Work Programme.

## Priority 1: Empower Global Network

Ensure that governments, researchers and users are equipped and supported to share, improve and use data through the GBIF network, regardless of geography, language or institutional affiliation.

### Activity 1a: Focus on people

#### Rationale (for the period 2017-2022)

GBIF is the result of work by thousands of people in agencies and institutions worldwide. This network's long-term sustainability depends on demonstrating the value of such contributions and justifying continued investment of effort.

The GBIF Secretariat can enhance efforts to develop capacity within the network and build an effective distributed help desk system that acknowledges and showcases relevant skills and experience that people across the network possess. Users of GBIF data products would also benefit from showcasing the network's indispensable human resources and their impact on assessing and improving biodiversity data.

For these reasons, the next round of improvements to GBIF.org should enhance the network's capacity to serve as a hub for the GBIF community by identifying its active contributors, integrating information on their relevant GBIF activities, and supporting broader knowledge exchange and skill development.

#### 2021 Progress

Building on previous efforts during the current Implementation Plan, GBIF has continued to make use of the growing cadre of volunteers— **Biodiversity Open Data Ambassadors**, **mentors** and **translators**—who contribute their time and energy toward fulfilling the network's global mission. While the opportunities for representing the GBIF community in relevant fora were limited due to the COVID-19 pandemic, the GBIF volunteer community has remained active in numerous virtual events this year, including mentoring and training workshops, e-conferences and symposia.

Apart from maintaining lists of volunteers, progress on this activity since 2017 has mainly focused on establishing the ability to connect individual ORCID IDs with various activities on GBIF.org and developing the **GBIF Community Forum** as an effective space for both general discussions and targeted consultations. More ambitious plans to represent and increase connections between the

individual contributions across GBIF.org have not been implemented, due in part to challenges of effectively managing and protecting personal data. The final year of the Implementation Plan presents an opportunity to review and consolidate this work in order to make progress toward the original objectives of this activity.

## 2022 Work items

- Improve representation of individuals on GBIF.org, with more coherent and integrated handling of the roles and profiles of Secretariat staff and network members, including biographies, activities and other related information

## 2021 Participant contributions

- **Andorra:** We are working on purely national activities (migration of the portal to an AmazonWS server, script correction, etc.).
- **Argentina:** Once the pandemic is resolved (we have an estimate of the possibility of vaccination only in April 2021) the plan is to resume the pending activities of 2021. All our node performs its tasks with the home office mode and it has been impossible to continue with many activities that depend on of the Ministry. Therefore, being able to resume activities normally would be our priority one.
- **Belgium:** Many mentors/trainers/translators activities.
- **Brazil:** Citation tracking possibilities were evaluated, but not implemented. Was implemented a tool that allows the user to know the number of downloads made for each collection, thus valuing the datasets of a given publisher.
- **Chinese Taipei:** TaiBIF staff members facilitated the BIFA workshop as mentors and helped the evaluation of BIFA and CESP projects. TaiBIF node manager is serving as Deputy Representative for Asia.
- **France:** We are involved as mentors and translators in different online trainings, we organized a webinar on data mobilization and we are engaged in Living Atlases symposiums. We are promoting the use of ORCID identifiers among our community of data providers.
- **Korea, Republic of:** The new two institutes has participated in Korean Biodiversity Consortium in 2021. Field surveys on biodiversity data with the Korean Biodiversity Consortium (approx. 70 people) has been conducted in June and September 2021 to collect biological samples and occurrence data from the principal classification. KBIF collected the data from the institutions participated in the field investigations, and incorporated the data into GBIF data portal.
- **Mauritania:** Improvement of knowledge on the biodiversity of three wetland areas in Mauritania.
- **Spain:** We did some dissemination on the matter and got back some interest ... and GBIF ambassadors. We collaborated in the proposal for a BID project led by Cuba, that was finally awarded. We look forward to start working in it soon. We led a CESP proposal in partnership with Argentina and Colombia that was turned down. Besides we participate, learn and contribute in initiatives such as LifeWatch, DiSSCo, and EOSC.
- **Sweden:** GBIF-Sweden staff have contributed to several evaluation processes organized by GBIF (BID, Hosted Portals) and by our neighbouring colleagues in Norway (data mobilization projects

proposals).

- **Uganda:** In 2021, the Uganda node supported the response to the Biodiversity for Development programme call for concept notes. A total of 5 concept notes were endorsed.

## 2022 Participant plans

- **Andorra:** Significantly increase the number of data on the national portal, with an important data reorganization. Then make the GBIF portal more widely available to biodiversity managers in Andorra. All these actions must have to permit a better network's long-term sustainability in Andorra.
- **Argentina:** Unfortunately, we (partially) follow the restrictions due to COVID19 and our situation is not different from the previous year. It remains our priority to reactivate face-to-face activities as soon as possible, despite everything, we have transferred the activities of data publication and support to publishers to virtual media, as well as training.
- **Belgium:** Organize EBRII conference to gather all Biodiversity Informatics people of our country. Continue our mentors/trainers/translators activities.
- **Brazil:** Continue the assessment of citation tracking possibilities; communicate the benefits to publishers.
- **Chinese Taipei:** TaiBIF will continue to work closely with the Secretariat on these.
- **France:** We will continue to support and organize workshops, webinars, symposia, and other events aimed at increasing digitization, mobilization, and research use of biodiversity data. Continue to update the metadata of the organizations and disseminate the benefits of using ORCID.
- **Korea, Republic of:** The field surveys of 2022 on biodiversity data with the Korean Biodiversity Consortium will be conducted in June and September 2022. And new two institutes will be participated in Korean Biodiversity Consortium in 2022.
- **Spain:** Support of data providers, data users, and interested parties in GBIF and biodiversity data is provided through a range of core operations: annual training program, e learning platform. Data publication and quality control services (inc. IPT), helpdesk, software development, communication services, Citizen Science services.
- **Sweden:** To the extent possible and asked for we will continue to provide such support as described above. We encourage being included in the public catalogue of potential trainers, mentors, technical support persons etc.
- **Taxonomic Databases Working Group:** At TDWG, our 2021 Conference Theme: Connecting the world of biodiversity data: standards uniting people, processes, and tools highlights that we understand, value, and note the need to focus on the fact that people create much biodiversity data (or create the tools that collect automated data), and people use, curate, develop policies and tools, and publish using this information. Our plans include developing a new strategic plan for TDWG, and continuing work to build a community understanding of a) how to contribute to standards development, b) how to implement them, and c) help the community to adopt any changes in the standards. This work will include continuing to strategize around what the Technical Architecture Group of the future looks like for TDWG. Additionally, TDWG notes the worldwide strain on all for constant "virtual" meetings (due to Covid). While in-person meetings require a lot of logistics, so too do virtual or hybrid meetings. The costs (both for humans, and

software to do this well) are high. We have very actively sought support this year in order to enhance what our small group can manage. We are already doing similar for next year.

- **Uganda:** The successful proposals will be helped to publish the datasets in the GBIF data portal through participating in the activities preceding data publishing.

## Activity 1b: Strengthen skills

### Rationale (for the period 2017-2022)

The strengthening of personal skills through international collaboration has been one of the great successes of the GBIF global network. During this implementation period, GBIF must build on this past experience (including the support mechanisms developed for the BID programme) to reinforce efficient training and capacity enhancement across the network.

Central to this is the development of a collaborative help desk capability and the alignment of relevant aspects of national training initiatives with a global curriculum to facilitate direct reuse of resources.

### 2021 Progress

The current implementation period has seen very substantial progress in skills development across the GBIF network, including the consolidation of training resources in multiple languages, mainly developed through BID but re-used across a range of projects and programmes. Increasingly these resources are available both for rollout in Secretariat-led workshops (virtual as well as in-person) and for self-instruction.

A new [training and e-learning page on GBIF.org](#) now provides open access to a core curriculum including courses on a general introduction to GBIF, node management, data mobilization, and data use for decision making. These courses use a flexible format enabling easy updates, self-guided instruction and embedding of videos. They complement the [e-learning platform hosted by GBIF Spain](#) which is still used to organize training workshops for participants in programmes such as BID and BIFA.

The following progress can be reported on the specific work items listed in the 2021 Work Programme:

- During 2021, collaboration between nodes and secretariats from GBIF and the [Ocean Biodiversity Information System \(OBIS\)](#) resulted in [a series of videos prepared for the virtual Global Nodes Meeting](#), outlining the overlaps and differences involved in publishing marine data for the two networks. This collaboration is the first step towards aligning guidance and training materials between OBIS and GBIF, with further opportunities provided by the data mobilization workshops in the second phase of BID, especially in the Caribbean where several funded projects have a marine focus.
- Following the release of manuals updating guidance on [georeferencing](#), [data cleaning](#) and [generalizing data on sensitive species](#), GBIF's core documentation programme proceeded with identification of further guidance to be commissioned in 2021. These include updated guidance on digitization and publication of data from camera traps, and new guidelines on species

migration and telemetry data, targeted filling of data gaps and data quality.

- Work on developing an e-Learning strategy to support GBIF's training needs has begun and will be completed in the final quarter of 2021.
- While opportunities to engage with external projects remained limited due to the pandemic, GBIF continued to support and develop training materials for the BioDATA programme, and a graduate programme for the UK Natural Environment Resource Council (see Activity 5a).
- Work was scheduled for the final quarter of 2021 to develop the curriculum for a data quality management course aimed at nodes, with rollout expected to carry over to 2022.
- The Secretariat launched a new sign-up form for volunteer mentors, and further steps for engaging and monitoring the work of mentors were planned for the final quarter of 2021.
- Improvements to the guidance materials for project writing, including better organization of FAQs and subtitles for guidance videos, are under way and expected to be substantially completed by the end of 2021.

## 2022 Work items

- Develop new training module on data quality management, targeted at GBIF nodes
- Monitor impact of GBIF training programmes, based on uptake of existing training resources, skills certification through digital badges, replication of courses throughout the network, and outputs of data and other products linked to participation in training
- Further develop joint guidance with OBIS on publication of marine datasets through both networks

## 2021 Participant contributions

- **Argentina:** Participation as mentor and trainer (national, regional and global), (start to) publish the National Biodiversity Inventory of the Environment Secretariat on the ALA Portal of Argentina . Continue with the advice of our publishers to facilitate and promote the publication of data sets.
- **Belgium:** Help developing an online data cleaning course, to support guidance on expected data quality services.
- **Brazil:** Tutorials were made available on the national platform, and an alignment with OBIS Brazil was carried out.
- **Chinese Taipei:** TaiBIF developed a training module on publishing data paper and is developing a training module on the use of GBIF-mediated data.
- **Finland:** FinBIF has actively been training FinBIF staff and the Finnish biodiversity informatics community in various skills in biodiversity informatics: georeferencing, data cleaning and data management by providing workshops and trainings.
- **France:** GBIF France supplied trainers and mentors for online workshops organized by GBIF during 2021 under BID and BIFA for capacity enhancement in data publishing. We cooperated with Cameroon and Benin to prepare a mentoring CESP proposal which has been accepted in 2021.
- **iDigBio:** iDigBio hosts a large number of workshops, webinars, symposia and other event on digitization and data related topics. iDigBio gave a presentation about several of these

opportunities at the April 2021 GBIF Community Webinar.

- **Mauritania:** Training of students; raising public awareness of endangered species and sources of pollution in continental wetlands.
- **Netherlands:** The Dutch node has supported and guided the development of a successful BIFA application from Indonesia. NLBIF as DiSSCo-NL national node representative has provided a workshop on harmonizing and updating GRSciColl records for the entire DiSSCo consortium.
- **South Africa:** The annual Forum will continue in the form of online virtual engagements with the community.
- **Spain:** Capacity building for the different communities involved in GBIF is a pillar of our ongoing strategy. Our e-learning platform and the use that GBIF and its communities make of it so proves it. We plan to continue and expand this area of activity and collaboration. Activities directed to enhance the network's capacity are well under way. These may be summarized as follows: Workshops (this year all held remotely on the [GBIF Spain eLearning Platform](#))
  - **Data management and data publishing in Polar research**
  - Data management and data publishing in LTER studies
  - Data visualization and analysis with R (scheduled for September)
  - Getting the most of GBIF Data portals (Scheduled for November) Webinars: This year focused on Collections. **One carried out, two scheduled**
  - Workshop with Spanish Environmental administrations on sharing and disseminating Biodiversity Information (scheduled for November)
- **Sweden:** In giving several courses directly or indirectly related to biodiversity informatics (basic and advanced training) in part using tools developed by, and in connection to GBIF - incl. other nodes - GBIF-Sweden has contributed to strengthening the national and regional community of bio- and biodiversity informaticians. We have actively contributed to the development of tools for genomic data provisioning and use.
- **Uganda:** The Ugandan Node together with the French node provided technical help for the data publishing organizations in Uganda notably National Fisheries Research Resources Institute.

## 2022 Participant plans

- **Argentina:** Participation as mentor and trainer (national, regional and global). Continue with the advice of our publishers to facilitate and promote the publication of data sets. Continue with the implementation of the Regional Portal for LAC, as well as the activities proposed by the CESP in which Argentina is the leader/participant. Strengthen ties with other nodes/organizations through thematic networks. Coordinate and improve the data of the institutions/collections/staff of Argentina in the GRSciColl.
- **Brazil:** Alignment and technical guidance with OBIS Brazil will be continued.
- **Chinese Taipei:** TaiBIF will continue to expand the topics in the national training workshops and will develop virtual training workshops.
- **France:** GBIF France will continue support for capacity enhancement activities and workshops. We plan to reuse the BID data mobilization training and some other available materials at national level and we will spread the joint guidance with OBIS on publication of marine datasets

as soon as it will be available.

- **iDigBio:** iDigBio will continue offering its training and workforce development activities via workshops, webinars, symposia, and other events.
- **International Centre for Integrated Mountain Development:** Based on the biodiversity data mobilization survey we are doing in 2021, ICIMOD will try to explore opportunity for capacity strengthening for data mobilization in collaboration with CAS Node. We will also be seeking help from GBIF Secretariat in terms of technical expertise.
- **Mauritania:** As part of the response to climate change, we plan to work on the seagrass beds of the National Park, Banc d'Arguin, and the restoration of coastal ecosystems in this protected area, which is an international connectivity point for the green turtles.
- **Netherlands:** Motivate and support potential GBIF partners to develop and submit applications to The BID and BIFA calls.
- **South Africa:** SANBI-GBIF will continue to conduct training workshops and develop communities of practice in identified biodiversity informatics areas.
- **Spain:** Besides continuing and expanding this area of activity and collaboration, we plan to upgrade our eLearning Platform to incorporate new functionalities, to make easier for everyone to re-use the material hosted there, and to facilitate finding them.
- **Sweden:** Through integration into the new Swedish Biodiversity Data Infrastructure (SBDI) context, comprising all major active units within the biodiversity informatics field in Sweden within one organization, we will increase our capacity and activity to strengthen and encourage data mobilization, enhance data flow and and support data use.
- **Taxonomic Databases Working Group:** Our plans now, and going forward include our TDWG Outreach and Communications Committee, currently led by Mareike Petersen. With this group's efforts, we are planning events like webinars (e.g. Darwin Core Hour) that help us bridge and expand participation in the world between those who develop standards, those who build tools to implement them well, those who use them (e.g. in their CMS, on export from their CMS), those who work to make our data more discoverable (e.g. schema.org, bioschemas.org), those who use these data for regional / national / global resource tracking and planning, and those who publish using these data.
- **Uganda:** Train more stakeholders on the use of the Integrated Publishing Toolkit for data paper publishing.

## Activity 1c: Equip Participant nodes

### Rationale (for the period 2017-2022)

By coordinating national, regional and thematic networks, Participant nodes play an essential role in helping GBIF engage the broadest possible community of institutions, initiatives and individuals engaged in biodiversity informatics. GBIF must provide learning materials and tools to support nodes efficiently as they work to mobilize biodiversity data, promote the reuse of available data and support users by improving data management and quality. Preferred approaches enable any one node to invest in developing tools and capacity that others can easily leverage for the benefit of the whole community. The skills and experience of the node managers and other team members are recognized as uniquely valuable in helping new Participants establish their nodes and allowing the

community to develop together.

## 2021 Progress

The following progress can be reported on the work items in the 2021 work programme:

- Due to the continuing COVID-19 travel restrictions, plans for in-person regional nodes meetings during 2021 were impossible to implement. Instead, and through consultation with the Nodes Steering Group, the Secretariat decided to organize a **virtual Global Nodes Meeting in June/July 2021**, returning to the usual cycle of holding such meetings every second year. As part of the series of sessions held over five days, regional breakout sessions were organized for all six of the GBIF regions. To maximum engagement during the meeting, a series of videos produced by Secretariat staff, nodes and other partners were made available ahead of time, with the live sessions focused on panel discussion and interventions from participants. The virtual format attracted record attendance for a Global Nodes Meeting, with 168 participants attending at least one live session, including node managers, node staff and observers from all regions.
- The involvement of OBIS in the virtual Global Nodes Meeting laid the ground for improved connections between OBIS and GBIF nodes, and this is evident from the progress reported by several nodes. Further work is planned for the remainder of 2021 to map the overlap and gaps between the two networks and determine where further connections could be made, including at regional level.
- The Secretariat established a working group involving several nodes to share experiences and make recommendations on working with indigenous and local knowledge (ILK), especially with regard to application of the **CARE Principles for Indigenous Data Governance**.
- The Secretariat began development of guidelines to provide more consistent visual branding of products and platforms associated with GBIF nodes, regions and publishers, as well as externally-managed platforms with a dependency on GBIF services. This was in connection with the hosted portal programme (see below), and outcomes of the work are planned for rollout in the final quarter of 2021.
- Following a **call for expressions of interest from nodes to take part in the pilot phase of the GBIF hosted portal programme**, the Secretariat accepted 19 hosted portal projects, covering institutional, national, regional and thematic examples. The first version of the technical framework was complete as of July 2021, and participants were in the process of preparing their content and styling. Terms were drawn up through legal consultation enabling the Secretariat to provide this hosted service through agreements with participating nodes. The first portals were expected to be made public on the internet in August 2021.
- Consideration of possible pricing structures for additional services was deferred until the completion of the hosted portal pilot, and will be taken up in the final quarter of 2021 and continued into 2022.
- Following discussion of the concept of regional offices at GB27 a stepwise approach to regional support was adopted in 2021, and the selection of regional project support contractors for Africa, the Caribbean and Asia (see Activity 1g) may be seen as the first step in this process.
- The Living Atlas community has continued to collaborate virtually throughout the Covid disruption with activity from 16 partners and virtual technical sessions with five. The **toolkit** that configures and manages an installation has been reworked to simplify the processes, and

installations are now able to include DOI citation and Sensitive Data Service modules. The community of users are increasingly contributing code to the core ALA software, particularly from France, Sweden and Spain, and the Technical Coordinator has contributed to the **GBIF / ALA data pipeline work** in preparation for rolling out across the wider Living Atlas community. A new **website** is near completion, and revisions to the community MoU, branding and documentation have been ongoing.

## 2022 Work items

- Mature the technical capabilities and community processes to expand the hosted portal service, focusing on the needs of GBIF participant nodes. This activity assumes positive feedback from participants of the pilot program. The expanded programme will determine the level of service that will be available to each node (including number of portals), and possible additional services that would be available through extra financial contributions.
- Continue support for the Living Atlases collaboration through extended contract for the technical coordinator (€15K)
- Support in-person regional nodes meetings for 2022, if permitted by the travel situation. The Asia regional nodes meeting will co-located with the Asia regional engagement meeting and supported by supplementary funds (see Activity 1g); €28,000 is carried over from 2020 for hosting of the Africa regional meeting in Togo; €32,000 will be available from the 2022 Work Programme budget to support additional meetings.
- Actively explore supplementary funding for additional regional support contractors in Latin America, the Pacific, and E Europe/Caucasus/Central Asia
- Refresh existing guidance materials for GBIF nodes and associated support, including node-level mentoring, capacity self-assessment and training for national-level project teams
- Develop guidance on best practices for working with indigenous and local knowledge, including application of the CARE principles, following recommendations of the working group established during 2021

## 2021 Participant contributions

- **Argentina:** Mentor to new nodes. Support and organization of regional nodes meeting.
- **Belgium:** CKAN Data portal, a light weighted solution for national nodes. See Belgian Biodiversity Data portal.
- **Brazil:** Only one assessment for the private sector approach was carried out. Regarding the infrastructure of the national platform, the Living Atlas tool was updated, in addition to important updates and improvements to the website.
- **Chinese Taipei:** TaiBIF staff members form a reading group on the topics related to biodiversity informatics.
- **Finland:** FinBIF has shared best practices in developing national node for European colleagues by having several meetings. There has also been relevant cooperation with the Nordic countries in e.g. taxonomy.
- **France:** We keep on to contribute and support the Living Atlases community while implementing the French national portal "openobs": sharing code on GitHub, support other nodes of the

community, contribution to workshop and symposium.

- **iDigBio:** iDigBio regularly participates in the GBIF community webinars, Global Nodes Meetings, Governing Board meetings, and North America regional meetings. iDigBio continues to host regular meetings of the Collections Data Infrastructures Working Group [CDIWG](#).
- **International Centre for Integrated Mountain Development:** Explore possibilities to jointly organize HKH level regional workshop to define data gap in collaboration with CAS and other GBIF Asian Node. This work did not materialize, however the biodiversity data mobilization capacity survey we are doing this year will allow us to understand the need on part of the our HKH member countries, and guide our future actions to support the data mobilization process as per their need. Partnership and support with other institutions will also be formally explored in 2022. The possibility of joint work between ICIMOD and CAS was outlined by the HKHBIF node presentation during the Asian Regional Node meeting. Prof. Ma Keping, Head of Delegation of CAS China Node acknowledged the possibility of collaboration with ICIMOD in future. Formal discussion between Head of Delegations would be necessary.
- **Korea, Republic of:** New IT staff has been employed to operating and managing the KBIF IPT server in May 2021, and constantly working on updating the data to increase usability in use in research, exhibition, and education.
- **South Africa:** SANBI-GBIF has supported the establishment of the ACM, as part of the broad concept to establish regional offices (bullet 5). The Node has developed a draft resource diversification strategy for consideration. SANBI-GBIF will also be presenting the ACM efforts at a Digitization Conference organized by GBIF/iDigBio and The U.S. National Museum of Natural History – Smithsonian (NMNH).
- **Spain:** We (in conjunction with GBIF-Portugal and technological partners - EGI, Ibergrid) have submitted a proposal to the EC for a project aiming fort the developing of a pilot biodiversity data portal for Europe. All Living Atlases activities continue. Progress: The mentioned EC project was granted and initial works started. [More details here](#). Currently prototyping some ALA spatial module capabilities.
- **Sweden:** GBIF-Sweden staff participated in workshops e.g. customizing ALA portals (Tanzania) ans on sensitive species, and in consultations on digital extended specimens (Alliance for Biodiversity Knowledge and internationally. We have continued to develop and provide tools for biodiversity data management under the SYNTHESYS+ and DiSSCo headings.
- **Uganda:** The Uganda node participated in the virtual events in 2020 and 2021.

## 2022 Participant plans

- **Argentina:** Mentor to new nodes. Support and organization of regional nodes meeting.
- **Belgium:** Continue to adapt our CKAN data portal.
- **Brazil:** For the coming year we expect greater involvement with the private sector. Regarding tools and services, training and help desk will be maintained for the IPT and another tools of Living Atlas.
- **France:** We will continue to participate in the activities of the Living Atlases Community.
- **iDigBio:** iDigBio will continue to participate in the GBIF community webinars, Global Nodes Meetings, Governing Board meetings, and North America regional meetings. iDigBio will continue its partnerships, discussions, and collaborations among the major global aggregators

via the Collections Data Infrastructures Working Group (CDIWG).

- **Mauritania:** Equipment for creating databases and for digitizing and sharing information on biodiversity.
- **South Africa:** SANBI-GBIF will engage as a participant in the GBIF-Africa Network.
- **Spain:** As the EGI-ACE will be in full motion we plan to deliver a first prototype of the joint portal, and start evaluating it. It will fully integrated in the EOSC infrastructure and procedures. We plan to increase the capacities of our national data portal to take better advantage of the spatial module. We aim to deploy several hubs in responding a demand by some institutions to present a more coherent view of their publisher data.
- **Sweden:** We intend to maintain, possibly increase, during 2022 our participation in similar activities as in 2021. By remaining an active part in the Living Atlases Community (LAC) we expect to contribute as before to tools and skills development internally and in connection to joint endeavours, and by incorporating colleagues from partners in the Swedish Biodiversity Data Infrastructure into our common capacity pool we will be even more able to assist and support efforts to strengthen node's skills.
- **Taxonomic Databases Working Group:** In general, TDWG activities benefit from Node Member participation and our outputs benefit Nodes. We will continue to align with GBIF efforts and needs as our objectives and reach complement each other.
- **Uganda:** More participation by Ugandans in the events.

## Activity 1d: Equip data publishers

### Rationale (for the period 2017-2022)

Data publishers are an essential component of the GBIF network as they share their content through the common infrastructure. More than 1,700 data publishers actively distribute datasets through GBIF.org, and their ranks increase steadily. Publishers from different parts of the world often face unique challenges, though common themes emerge. These problems range from lack of data publishing experience or skills, lack of equipment, language barriers, difficulties in managing data hosting facilities, and the inability to publish high-quality data or curate data into the future. The Integrated Publishing Toolkit (IPT) requires ongoing improvements and enhancements, including the establishment of hosted instances that reduce the technical burden on data publishers.

### 2021 Progress

The following progress can be reported for work items in the 2021 Work Programme:

- The first bi-annual newsletter to data publishers is scheduled for the final quarter of 2021.
- Enhancements were made to the pages associated with DOI citations of GBIF-mediated data, to enable easier export of citations using standard formats.
- Exploration of push notifications to data publishers for new citations was deferred.
- Further development of regional helpdesk services was redirected into the establishment of the regional project support contractors (see Activity 1g).
- The open curation of vocabularies has been included in the GBIF registry with the LifeStage

vocabulary being the first pilot used by GBIF and the ALA. Work is underway with the TDWG Vocabulary Task Group to migrate all existing vocabularies into the registry. Once complete, communities can get involved in the maintenance of the vocabularies which dictate how data is treated in GBIF.org.

Following publication of the [updated guidance on generalizing data on sensitive species](#) and in response to issues raised during GB27, the Secretariat set up a working group of nodes to document the various approaches across the community in application of national sensitive species lists. This group will make recommendations on a more consistent approach, and options for helping data publishers to identify records that may present risks if published with precise location details.

At the 2020 TDWG Conference, GBIF and the BioCASE team held two virtual workshops to capture ideas and hold discussion on the future of the Integrated Publishing Toolkit (IPT). The [final report](#) captures the outcomes which provided clear direction. Following this feedback GBIF has appointed a new lead developer on the IPT product and released the 2.5.0 version of IPT which addresses many of the key issues.

The framework for a technical documentation website has been created, with the [IPT manual](#) being the first component included. Collating information on architecture, data processing, definitions of formats, tutorials, guides and the API is ongoing but requires more dedicated focus during the remainder of 2021 and 2022.

## 2022 Work items

- Building on the 2021 work of the sensitive species working group, commission an analysis of the current scale of sensitive species data published through GBIF, evaluation of the evidence that such data presents risks of harm, and recommendations for applying a risk assessment approach (€15K)
- Based on work carried out in 2021, develop plans for the future of the Integrated Publishing Toolkit (IPT).
- Implement new technical documentation website

## 2021 Participant contributions

- **Argentina:** Argentina has a collection catalogue (and institutions) since 2003. We keep updating it. Try to reduce the number of IPT to centralize all the data sets in one IPT.
- **Belgium:** Helpdesk IPT and Darwin Core users.
- **Brazil:** Tutorials and training in the use of the IPT were improved and made available on the national platform website. The data flow was discussed with the SiBBR Steering Committee, in order to better structure the role of the national node.
- **Chinese Taipei:** TaiBIF holds regular training workshops on data publishing and mobilization and continues to provide the help desk service to data publishers.
- **Finland:** FinBIF has continued to support the Russian initiatives in sharing data to GBIF through EarthCape Biodiversity Data Platform. FinBIF has supported the sharing of Russian alien species information with GBIF in the project “Collaborative Data and Information Exchange Network for Managing Invasive Alien Species”. FinBIF has also provided support to Vietnam National

University in sharing Bryophyte data to GBIF through FinBIF IPT installation.

- **France:** GBIF France offers data hosting and maintains IPT instances for several national institutions and a dozen of southern countries. Our IT engineer contributed to the development of the IPT by implementing new functionalities and participated in the tests of the future version 2.5.
- **iDigBio:** iDigBio is an active partner with GBIF on the global collections registry (GRSciColl). iDigBio continues to actively seek new specimen-based biodiversity datasets for ingestion by both iDigBio and GBIF as part of its mission to coordinate a national digitization effort in the United States.
- **International Centre for Integrated Mountain Development:** Sensitize data holders and data providers in ICIMOD's eight regional member countries in the Hindu Kush Himalayan region about the importance of data publishing, and use of global publishing platform. Created a [HKHBIF web platform](#) to sensitize data holders and data providers in the HKH region. Also contributed to an [opinion article to highlight importance of open access to data](#).
- **Korea, Republic of:** Two institutions per year has been engaged in the Korean biodiversity data network as potential data publishers.
- **Mauritania:** Continue the digitization of the botanical collection of the National Herbarium of Mauritania (HNM).
- **South Africa:** The Node will continue hosting of the IPT and will continue its support to data publishers. The Node is part of the pilot Hosted Portals implementation.
- **Spain:** IPT maintenance, publisher support, publishers (factual and potential) training (See activity 1b.), data validation, quality control and associated tasks are operations routinely and continuously carried out by the node.
- **Sweden:** As described above GBIF-Sweden participates in the newly set up working group of nodes to document the various approaches across the community in application of national sensitive species lists. We have NOT contributed, but have followed closely the development of the future "IPT solution".
- **Uganda:** A freshwater biodiversity data portal was developed and launched. This was a result of a grant from the BID programme.

## 2022 Participant plans

- **Andorra:** Capture more data publishers.
- **Argentina:** Argentina has a collection catalogue (and institutions) since 2003. We keep updating it.
- **Belgium:** Helpdesk IPT and Darwin Core users.
- **Brazil:** Training activities and data publication will be continued and strengthened with CESP's project "The role of Zoological Collections Networks to enhance and fill taxonomic and geographic gaps in open access databases" with a focus on training curators of zoological collections for the use of Specify 7.
- **Chinese Taipei:** TaiBIF will continue to provide these services and will actively contact data publishers to help them prepare and clean data for publishing.

- **Finland:** FinBIF is building API connections for sharing occurrence data through API's to replace the IPT installation. This would greatly facilitate the sharing of the rest of the Finnish datasets to GBIF by minimizing manual work needed e.g. in recording metadata.
- **France:** GBIF France will continue to support data hosting and publishing services for national institutions and southern countries. We are willing to continue to help in the development and testing of the future version of IPT.
- **iDigBio:** With renewed funding from the U.S. National Science Foundation, iDigBio will continue its efforts to coordinate a national digitization effort in the United States. iDigBio will continue to work with GBIF on the global collections registry (GRSciColl).
- **Korea, Republic of:** Two institutions which has been engaged in the Korean biodiversity data network in 2021 will be participated in data publishers. We are checking the data quality of them.
- **Mauritania:** We will continue to digitize botanical samples and we will continue our explorations of wetlands that harbour endangered species in this tropical fringe in Mauritania on the southern edge of the Sahara and east of the Atlantic. It will involve exploring the forests of herbaria, mangroves and marine and coastal algae still very little known in Mauritania and which respond to an international concern for carbon fixation and the mitigation of global warming.
- **South Africa:** The Node will continue hosting of the IPT and will continue its support to data publishers.
- **Spain:** Same as 2021, plus updating the IPT to the latest version and hiring a second person as core staff to cope with the work load in this area and provide a better service to the publishers.
- **Sweden:** 2022: In 2022 we will continue our efforts to encompass different kinds of lists of sensitive or otherwise ( specifically appointed species (global, national, regional red, grey, black lists). Our practical/technical support to data providers will continue and include also "new" data types.
- **Switzerland:** Contribute to the sensitive species working group.
- **Taxonomic Databases Working Group:** In general, TDWG activities benefit from Node Member participation and our outputs benefit Nodes. We will continue to align with GBIF efforts and needs as our objectives complement each other. For example, we will have a webinar in September that focuses on recent changes to DwC to help everyone (including Node Managers) understand them, and answer any questions about how to implement them.
- **Uganda:** To increase the usage of data in the decision making process. To assist in use of GBIF mediated data in developing the National redlist for Fish in Uganda.

## Activity 1e: Expand national participation

### Rationale (for the period 2017-2022)

GBIF's national membership has remained largely static in recent years. Active participation in the network is confined to a limited number of regions, with the most dynamic activity in node collaboration and data publication focused in Western Europe, North America, Latin America, Oceania, increasingly sub-Saharan Africa and limited parts of Asia. Thus, significant parts of the world have little or no direct involvement in GBIF's activities, which poses a risk to the long-term credibility of GBIF as a global collaborative network. Secretariat staff and existing Participants

constantly engage potential new members through ad hoc contacts, but the situation calls for a more strategic approach to expand membership.

## 2021 Progress

Looking across the whole implementation period to date, 13 new countries have joined GBIF as national Participants. They include Mali (2017); Cameroon, Liberia, South Sudan, Viet Nam and Zimbabwe (2018); Angola, Belarus and Tonga (2019); Armenia, Cambodia and Uzbekistan (2020); and Namibia (2021). Over the same period, seven regional or international organizations have joined as associate participants: the Secretariat of the Pacific Regional Environment Programme (SPREP), iDigBio (Integrated Digitized Biocollections), LifeWatch ERIC, Specify Collections Consortium, Distributed System of Scientific Collections (DiSSCo), the East Asia Biodiversity Conservation Network, and Observation International. Several of the new members joined GBIF as a direct result of engagement through funded programmes led or supported by GBIF, in particular the BID, BIFA and BioDATA programmes.

Opportunities for direct engagement at international conferences and intergovernmental fora continued to be severely limited in 2021 due to the COVID-19 pandemic. However, a series of regional engagement webinars was planned for later in the year, building on the successful outreach achieved by such webinars in 2020. In addition, targeted communication products were developed during 2021 aimed at strengthening GBIF's value proposition for existing and potential national members, as well as supporting the new GBIF resource mobilization strategy (see Activity 1g).

While expansion in formal participation over the period has been modest, engagement through a range of activities has significantly increased GBIF's global footprint both in terms of the sharing and use of data through the network. By July 2021, institutions from 124 countries were publishing data through GBIF, compared with 61 countries that participate through signature of the GBIF Memorandum of Understanding (MoU). Two countries outside the formal network (India and Italy) were within the top ten locations of visitors to GBIF.org, while China ranked second in the world both for data download requests and peer-reviewed studies citing use of data accessed through GBIF. Perhaps the most dramatic increase in activity outside the formal national membership has been in the case of the Russian Federation, where **data publishing** stood at virtually zero in 2017, and by 2021 had reached more than 3.8 million species occurrence records, being shared through GBIF in nearly 500 datasets published by 117 Russian institutions.

The following specific progress on work items in the 2021 Work Programme may be reported:

- Pro-active engagement in the delayed process for agreeing the post-2020 Global Biodiversity Framework through the Convention on Biological Diversity (CBD) has maintained GBIF's profile in supporting national implementation of global goals and targets to address the biodiversity crisis. Acknowledgement of GBIF in decisions and background documents still under negotiation ahead of the COP15 meeting in China will provide important additional support for expansion of GBIF's membership in the coming decade, especially in relation to GBIF's role to support indicators, taxonomy, knowledge management and capacity development (see also Activity 5c).
- Plans for an annual call for regional outreach proposals were deferred due to the impossibility of travel during 2021.
- Exploration of potential additional funded capacity programmes was initiated through discussions with a number of potential funding partners; this will be carried over into

implementation of the resource mobilization strategy in 2022 (see Activity 1g).

- Further progress in translation of web content and key publications to support internationalization of GBIF.org was planned for the final quarter of 2021. In addition, the contracting of professional live interpretation for virtual events was piloted in the BID Caribbean data mobilization workshop, and will be further considered for the regional engagement webinars.
- A catalogue of capacity enhancement services has been drafted as an internal document, and will be developed further in the final quarter of 2021 to support outreach materials for funders under the resource mobilization strategy (see Activity 1g).
- A set of communication products supporting GBIF's value proposition across key thematic areas was commissioned and will be rolled out in the final quarter of 2021.

## 2022 Work items

- Review the current status and models of participation in GBIF both at national and organization level, with a view to developing recommendations for the new strategic implementation period at GB29
- Take a more proactive role with HoDs through global science collaboration activities such as the Group of Senior Officials
- Use the delayed agreement of the post-2020 Global Biodiversity Framework to further emphasize GBIF's value to countries and promote expanded participation

## 2021 Participant contributions

- **Argentina:** Argentina participates actively in the translation of GBIF content. Support is provided, as a regional representative, to all countries that do not have a node yet, but are interested in joining GBIF.
- **Belgium:** French and Dutch translations to help maintaining and expanding internationalization of GBIF.org.
- **Brazil:** The internationalization of GBIF.org has been expanded, through the communication of notices on the SiBBR website and on social networks, in addition to making GBIF documents and guides available for specific users. We are also structuring and publishing data from long-term research programs (PELD and PPBIO programs).
- **France:** Contribution to French translations on GBIF.org, BID workshop materials and support to francophone countries through documentation and engagement with countries or institutions. Via our contribution as trainer or participant in working group in EU project such as COST MOBILISE or DiSSCO prepare, we try to reinforce awareness on GBIF.
- **iDigBio:** iDigBio, Smithsonian's National Museum of Natural History, and GBIF are collaborating on a two-day virtual event that highlights the collective successes of mobilization and digitization efforts in the US and abroad. The event will be held on 22-23 Sep. Topics will include:
  - Innovations: Strategy & Coordination (Plenary & Lightning Talks)
  - Digitization Spotlights (Poster Session)
  - Community (Plenary & Lightning Talks)

- Natural History Collections in Grand Challenges (Plenary & Lightning Talks)
- What's Next Partnerships, new data concepts (Plenary & Lightning Talks)
- Looking to the Future of Digitization: Conference Wrap-Up (Panellist Discussion)

The event follows the annual ADBC Summit on 21 Sep. Also, the iDigBio Node Manager currently serves as the Regional Representative for participant nodes in North America, and iDigBio has a leadership role with the Biodiversity Collections Network (BCoN), which aims to expand the efforts of the NSF Advancing Digitization of Biodiversity Collections (ADBC) program.

- **Korea, Republic of:** As of September 2021, 61 organizations are participated in the Korean Biodiversity Consortium to strength our data network.
- **Netherlands:** NLBIF has contributed to a successful BIFA application from Indonesia which might result in the future expansion of the GBIF network.
- **Spain:** Maintaining and expanding our ties with LifeWatch, increasing participation with the LTER communities via our participation the eLTER Plus and eLTER PPP, projects (EC funded) Willing to get involved in BID projects pertaining the Caribbean. Progress: Work progressing according to plan. Workshops targeted to specific communities (polar research, LTER) aimed towards this objective.
- **Sweden:** By participating in, and staying updated with several of the most influential organizations (active participation, visiting - virtual - conferences), Our staff is well updated with the general progress on the BI field och activity. Until GB28 we hold a position in their Nodes Steering Group, and have - as always - participated in Global Nodes Meeting and the Governing Board (HoD associate representative).
- **Uganda:** The community of practice for biodiversity informatics enhanced with more open data ambassadors from Uganda.

## 2022 Participant plans

- **Argentina:** Argentina participates actively in the translation of GBIF content. Support is provided, as a regional representative, to all countries that do not have a node yet, but are interested in joining GBIF.
- **Brazil:** Through two CESP Projects, SiBBR will increase integration with curators and publishers in Brazil and other partner countries
- **France:** Help for French translations and support to francophone countries will continue.
- **Korea, Republic of:** Another two institutions will be engaged in the Korean biodiversity data network will be participated in 2022 as potential data publishers.
- **Mauritania:** Work on endangered species, work in progress on medicinal and aromatic plants, work on invasive species.
- **Netherlands:** Continue activities to motivate partners and countries to start sharing open biodiversity data through the GBIF network.
- **Spain:** We plan to continue with our series of workshops targeting specific communities and maintain specific contacts. We signed in support of the Alliance for Biodiversity Knowledge, and would be happy to contribute, but we do not see how.

- **Sweden:** GBIF-Sweden needs to be aware and part of the global community and will remain as participants in all relevant meetings and situations. However we have decided NOT to present a candidate for the new Nodes Steering Group in 2021 as we think this should be someone else's task for some time to come now. We will return to nominate candidates to such positions later on as we intend to shift our present Node Manager for a new colleague late 2021.
- **Taxonomic Databases Working Group:** TDWG plans to continue to strategize and offer ways to participate around the clock, around the world, and in multiple ways and languages. We'll continue to do so. We are in contact with Pensoft, for example, for the last year, about changes we'd like to see in ARPHA so that abstracts in BISS can be in multiple languages. Pensoft assures us that changes like this are coming that will enhance our abilities to make our data and our information FAIR and extensible.
- **Uganda:** Train more open data ambassadors with certifications.

## Activity 1f: Plan implementation

### Rationale (for the period 2017-2022)

Implementing this plan hinges on effective coordination with the plans of individual GBIF national and organizational Participants. GBIF must also ensure that node activities are well recognized and integrated with other biodiversity research and informatics initiatives at national, regional and organizational levels.

### 2021 Progress

As part of the system of reporting on the annual updates of the global GBIF Implementation Plan, the Secretariat has raised the visibility of activities across nodes by including Participant contributions and plans alongside the summaries for each activity within the strategic priority areas. To reinforce this approach, the Secretariat is preparing to capture information from nodes about budgets and in-kind resources dedicated to GBIF activity, so that the total investment in GBIF can be better reflected in financial reporting.

Following the 2nd Global Biodiversity Informatics Conference (GBIC2) in 2018, GBIF has taken a leadership role in convening a number of consultations across the broader biodiversity informatics community through the *alliance for biodiversity knowledge*.

In 2021, as part of the EU-funded **SYNTHESYS+ project** and within the framework of the *alliance*, GBIF convened a workshop focused on Converging Digital Specimens and Extended Specimens - Towards a global specification for data integration. This workshop had been anticipated to be the third in a series of but was moved forward. The consultation covered 11 subtopics and was held in two phases, **Phase 1 in February** and **Phase 2 in June**.

Due to the COVID-19 restrictions, the 3rd Global Biodiversity Informatics Conference (GBIC3) was postponed until 2022. The event will be aligned with the EU-funded **BiCIKL (Biodiversity Integrated Knowledge Library) grant**.

## 2022 Work items

- Organize the 3rd Global Biodiversity Informatics Conference (GBIC3), hosted by the **LifeWatch European Research Infrastructure Consortium (ERIC)**, to address the ongoing operation and governance of the *alliance for biodiversity knowledge* (€10K)
- Organize and run the 2nd SYNTHESYS+ workshop relating to citation of collections

## 2021 Participant contributions

- **Argentina:** Continue contributing to national, regional and global activities.
- **Brazil:** There was a monitoring of activities and services made available at GBIF. The calls for the Ebbe Nielsen Challenge and Young Researchers Awards 2021 were nationally released.
- **Chinese Taipei:** TaiBIF worked with major data holders in Taiwan to develop five-year strategic plans for enhancing data mobilization and developing biodiversity informatics through Taiwan Biodiversity Information Association (see Activity 3c).
- **Finland:** An article describing the structure of FinBIF was published: Schulman, Leif; Lahti, Kari; Piirainen, Esko; Heikkinen, Mikko; Raitio, Olli; Juslén Aino (2021). The Finnish Biodiversity Information Facility as a best-practice model for biodiversity data infrastructures. *Scientific Data* doi: 10.1038/s41597-021-00919-6. FinBIF has been actively participating in SYNTHESYS+ and DiSSCo projects and in CETAF.
- **France:** Continue contributing to national, regional and global activities. Links with the French bon of GEOBON.
- **iDigBio:** iDigBio currently hosts the steering group for the GBIF North America region. In addition, iDigBio has often been hosting the regional meetings of GBIF North America.
- **Korea, Republic of:** Data gap discovery is carrying out for the data from National Science Museum, a record of approx. 5.14 Mil. National Institute of Ecology uploaded total 1,000,134 occurrence data (1 dataset) to GBIF data portal in December 2020. East Asia Biodiversity Conservation Network uploaded total 100 occurrence data (1 dataset, Flora of the Korean DMZ and adjacent area) to GBIF data portal in February 2021. National Marine Biodiversity Institute of Korea uploaded total 68,832 occurrence data (1 dataset) to GBIF data portal in June 2021.
- **Mauritania:** We are counting on the institutional support of our establishment and on the national research and innovation agency for the execution of our work plan.
- **South Africa:** Further engagements with SANBI-GBIF Node Manager and HoD/DSI to evolve the Node planning. GBIF YRA national process established and call distributed.
- **Spain:** Ongoing collaborations and activities with LifeWatch, Synthesys , and CETAF. Increase collaboration and integration of our citizen science activities with iNaturalist.
- **Sweden:** Since 2020, the strategic (5-yr), and annual work plans of GBIF-Sweden are integrated into the same plans for SBDI (the former now spanning over 2021-2024). Where applicable they are heavily congruent with GBIF:s planning, yet of course more detailed and extended where specifically needed. As described we thus follow the development of the ABK and GBIC1, 2 and 3. By being an active partner in the SYNTHESYS+ and DiSSCo we also by necessity have to stay engaged in these and other related organizations.
- **Uganda:** Participated and facilitated in the the biodiversity informatics stakeholder activities.

Data publishing and resource mobilization efforts were supported through provision of recommendation letters and endorsements for support.

## 2022 Participant plans

- **Andorra:** Significantly increase the number of data on the national portal, with an important data reorganization. Then make the GBIF portal more widely available to biodiversity managers in Andorra. All these actions must have to permit a better network's long-term sustainability in Andorra.
- **Argentina:** Continue contributing to national, regional and global activities.
- **Brazil:** Participation and monitoring of GBIF activities and national dissemination of these activities will be continued in the coming year.
- **Chinese Taipei:** TaiBIF will work together with the members in the Association on implementing the plans.
- **France:** Continue contributing to national, regional and global activities. Ongoing collaborations and activities with EU and southern projects. Increase collaboration with national iBOL node.
- **iDigBio:** iDigBio plans to continue attending, supporting, and/or hosting meetings for the GBIF North America Region as required/requested.
- **Korea, Republic of:** We have a plan to complete the data gap discovery for the data from National Science Museum, a record of approx. 5.14 Mil. within next year.
- **Mauritania:** We also count on GBIF to support next phase of certain activities within the framework of the digitization and the training of the agents of the research institutions.
- **South Africa:** Continue contributing to national, regional and global activities.
- **Spain:** Besides continuing with on going collaborations, we plan to participate in GBIC3, and Synthesys+ workshop.
- **Sweden:** See above! GBIF-Sweden will continue in 2022 to encourage its staff and SBDI partners to engage in any relevant joint activity in the BI community.
- **Uganda:** To coordinate the national node participants to take part in the various biodiversity informatics opportunities like capacity enhancement concepts among others.
- **Taxonomic Databases Working Group:** TDWG is participating in BiCIKL. TDWG and BiCIKL will hold a joint TDWG-BiCIKL2022 Conference in Sofia in October (in-person preferred, but may be hybrid, or virtual, depending on Covid). TDWG looks forward to GBIC3 as part of the Alliance for Bio and continuing the work to imagine the future together.

## Activity 1g: Coordinate resources

### Rationale (for the period 2017-2022)

Among its other roles, the GBIF Secretariat coordinates efforts to expand the scale and scope of GBIF activity beyond the levels achievable using only annual core Participant contributions. GBIF Participants commit to establish and operate nodes which serve as significant centres for GBIF activity. Some nodes have sufficient resources to contribute skills and developments which advance GBIF's work, while others may require external support to become fully active. A limited amount of

funding has been allocated each year under the GBIF work programme to support capacity enhancement for GBIF nodes. GBIF or individual Participants may also secure supplementary funds to contribute to particular areas of GBIF work. Improved coordination of these various resources will assist GBIF to advance more rapidly at all scales.

## 2021 Progress

During 2021, the Secretariat selected a total of 42 projects for funding under the second phase of the Biodiversity Information for Development (BID) programme following the calls for the proposals launched during 2020. These included **18 projects in sub-Saharan Africa**, 18 in the Caribbean and 7 in the Pacific. In addition to the top-up funding of €1.6M from the **European Commission's Directorate-General for International Partnerships (EU-INTPA)** to support the new phase of BID, GBIF secured a funding agreement with the **JRS Biodiversity Foundation** to contribute €411K towards BID projects in Africa, enabling approval of five additional projects on biodiversity data mobilization and use.

The continuing COVID-19 travel restrictions inevitably disrupted the original plans for implementing BID activities during 2021. Apart from requiring funded projects to avoid international travel in their plans for this year, the main impact was the need to replace planned in-person training workshops organized by the Secretariat with virtual alternatives for the three target regions. The Secretariat Participation and Engagement team organized virtual data mobilization training for BID grantees in Africa (May), the Caribbean (July/August) and Pacific (August), using the curriculum developed during the first phase of BID and combining a preparatory phase with 'live' sessions over a week and subsequent submission of exercises to demonstrate acquisition of the necessary skills. Additional virtual sessions were conducted for the project teams implementing the data use grants, introduced as an innovation in the second phase of BID.

The BID Steering Committee approved a proposal from the Secretariat to redirect funds originally budgeted for in-person workshops towards regional project support contracts, aimed at providing ongoing technical support for the funded projects within the target regions, and helping to grow the community of practice developed through the BID programme. Following open calls in **Africa** and the **Caribbean**, three individuals in Africa and one in the Caribbean were selected to provide these support services under 2-year contracts starting in August 2021. As well as supporting the projects funded through BID, these contractors will work with regional node representatives to help coordinate and direct technical support for GBIF nodes in the respective regions.

Following a **call for proposals in December 2020**, the Secretariat selected 13 data mobilization projects in Asia, under the sixth funding round of the Biodiversity Information Fund for Asia (BIFA), using supplementary funds contributed by the Ministry of Environment, Government of Japan. As this was expected to be the last BIFA call in its present form, the BIFA Steering Committee agreed to 'double up' two annual contributions into a single call with €220,000 available for grants up to €20,000 each, with projects implemented over 18 months. Particular priority in the call was given to projects combining DNA-based techniques such as barcoding with mobilization of data through GBIF. A virtual data mobilization training workshop for the BIFA projects was scheduled for November 2021, with funds reserved for an in-person engagement meeting in Asia in 2022 if the travel situation permits.

Based on the model developed for the BID programme, the Secretariat prepared a call for proposals for regional project support contracts for Asia, using funds from the EU-funded SYNTHESYS+

programme under an agreement reached with the programme partners. These roles will be recruited by November 2021 from Asian applicants and will support project teams funded through BIFA as well as nodes and data publishers in the region.

The 2021 call for proposals under the GBIF Capacity Enhancement Support Programme resulted in selection of five projects involving a total of 14 nodes, led from Argentina, Armenia, Brazil, Cameroon and South Sudan. The themes of the selected projects included mentoring to support node development, extending knowledge of data quality, and the role of zoological collections in filling taxonomic and geographic data gaps.

A draft GBIF resource mobilization strategy, aimed at broadening the funding base for GBIF's global activities, was developed by the Secretariat's Strategic Partnership Officer and endorsed by the Executive Committee, and will be presented for approval by the GB28 meeting alongside the draft Strategic Plan for the period 2023-2027.

## 2022 Work items

- Launch a call for proposals under the 2022 Capacity Enhancement Support Programme (€80k)
- Following approval of the resource mobilization strategy at GB28, develop implementation road map and begin implementation, with a view to targeting additional supplementary funds as well as consolidating core contributions from Voting Participants
- Subject to an improved travel situation in 2022, organize an in-person engagement meeting in Asia, using supplementary funds from BIFA (€20,000) and the SYNTHESYS+ project (€19,000)

## 2021 Participant contributions

- **Andorra:** We have financial assistance from the government of Andorra to cover part of the operating costs of GBIF Andorra.
- **Argentina:** Participation in 2 CESP programs (as a participant).
- **Belgium:** Help by mentoring BID, BIFA, CESP projects. At national level, the Belgian Biodiversity Platform :
  - collaborates with [LifeWatch.BE](#)
  - acts as IPBES National Focal Point.
- **Brazil:** A new CESP proposal was submitted, and the proposal approved in 2020 is being implemented with the countries of Latin America for the strengthening of regional nodes.
- **Finland:** The research infrastructure funding for FinBIF from the Academy of Finland is used to contribute to GBIF work. The entire work of FinBIF aims to expand and develop the national network and resources. The cooperation with Russian initiatives has been ongoing.
- **France:** We are partners in 1 CESP mentoring project led by Cameroon and 2 BID projects in Africa : 1 regional led by Togo (Progress towards a regional data platform of West and Central African herbaria) and 1 national led by Ivory Cost (Mobilization and strengthening of biodiversity data supporting sustainable development in Côte d'Ivoire) We were trainer and mentor for the virtual data mobilization trainings for BID grantees in Africa (May 2021) and the Caribbean (August 2021).

- **iDigBio:** iDigBio, GBIF, ALA, and DiSSCo are making a concerted effort to coordinate efforts via the Collections Data Infrastructures Working Group (CDIWG). iDigBio became a member of the Research Data Alliance (RDA). iDigBio regularly participates in the collaborative efforts of GBIF, BCoN, and DiSSCo, among others, to forge a path for implementation of Extended Specimens.
- **Korea, Republic of:** East Asia Biodiversity Conservation Network uploaded total 100 occurrence data (1 dataset, Flora of the Korean DMZ and adjacent area) to GBIF data portal in February 2021. National Marine Biodiversity Institute of Korea uploaded total 68,832 occurrence data (1 dataset) to GBIF data portal in June 2021. National Institute of Ecology and National Marine Biodiversity Institute of Korea will upload more than total 700,000 occurrence data from the both institutes within 2021, and also share the coordinates (partially updated) which is already published and potential datasets in GBIF portal.
- **Mauritania:** On the level of the Node.
- **Spain:** Plans: Aiming to contribute to the CESP and BID programs. Progress: Participation in CESP:0, in BID: 1.
- **South Africa:** Draft Resource Diversification Strategy developed and shared with Africa Nodes, for consideration. Continued data mobilization funding through FBIP large and small grants.
- **Sweden:** 2021: Although we cannot support financially external initiatives, by pooling resources allocated to biodiversity informatics in Sweden into the SBDI, and offering expertise thus mobilized, we hope to be able to be more supportive technically to GBIF globally. This goes for 2021 and 2022 (plus in at least the near future thereafter!).
- **Uganda:** The National Node reviewed and endorsed 5 project concepts and 1 was successful and funded. Participated in the proposal development for the more resources regarding biodiversity information for reporting on the 3 UN RIO Conventions for Uganda.

## 2022 Participant plans

- **Andorra:** We have financial assistance from the government of Andorra to cover part of the operating costs of GBIF Andorra.
- **Argentina:** Participation in 3 CESP programs (1 as coordinator, 2 as a participant).
- **Belgium:** Get involve in DiSSCo.BE. Coordinate the BiodivERsA Partnership, a coordinated European initiative of biodiversity research funding agencies.
- **Brazil:** Implementation of the two CESP proposals coordinated by the Brazilian node.
- **Chinese Taipei:** TaiBIF attempts to coordinate biodiversity informatics-related resources among major data holders in Taiwan through Taiwan Biodiversity Information Association (see Activity 3c).
- **France:** Aiming to contribute to the CESP and BID/BIFA programs.
- **iDigBio:** iDigBio will continue its partnerships, discussions, and collaborations via CDIWG, RDA, and BCoN International Partners.
- **Mauritania:** We intend to connect with our partners.
- **South Africa:** National funding process for data mobilization to continue.
- **Spain:** Same as for 2021.
- **Sweden:** 2022: Same as in 2021: Although we cannot support financially external initiatives, by

pooling resources allocated to biodiversity informatics in Sweden into the SBDI, and offering expertise thus mobilized, we hope to be able to be more supportive technically to GBIF globally. This goes for 2021 and 2022 (plus in at least the near future thereafter!).

- **Uganda:** To increase awareness and encourage more stakeholders to increase the response to the call for proposals. Monitor and mentor the successful project implementation to achieve the intended goals of supporting conservation decision making.
- **Taxonomic Databases Working Group:** TDWG looks forward to continued collaboration with GBIF in the coming years to help as we can with coordinating resources, use of those resources, access to them, and contributing to the development of new ones.

## Priority 2: Enhance Biodiversity Information Infrastructure

Provide leadership, expertise and tools to support the integration of all biodiversity information as an interconnected digital knowledgebase.

### Activity 2a: Modernize data standards

#### Rationale (for the period 2017-2022)

The GBIF network participants are able to reliably exchange data thanks to their adherence to a set of standards. As GBIF looks to grow in capability, enable exchange of richer content and improve the quality of data, the standards must be revised and evolve accordingly.

Current standards adopted by GBIF are not yet adequate to accommodate the needs expressed by many potential and existing data publishers. Weaknesses in the model have led to ambiguous or over-complex data representations and unclear documentation, leading to difficulties in data integration and use. The main issues relate to uncertainties around the use of Darwin Core record types, the basisOfRecord element, and the use of Core and Extension vocabularies. Reviewing and updating the core domain model, tightening up the vocabularies and documentation and adopting more robust exchange standards will result in an easier to use, and a wider reaching GBIF data exchange network.

#### 2021 Progress

In collaboration with partners, GBIF under the umbrella of the *alliance for biodiversity knowledge* convened a two phase (**phase I** and **phase II**) virtual global consultation on Digital Extended Specimens (DES) covering the following topics:

- Making FAIR data for specimens accessible
- Extending, enriching and integrating data
- Annotating specimens and related data
- Attributing work done (Data Attribution)

- Analysing/mining specimen data for novel applications
- Robust access points and data infrastructure alignment
- Persistent identifier (PID) scheme(s)
- Meeting legal/regulatory, ethical and sensitive data obligations
- Workforce capacity development and inclusivity
- Transactional mechanisms and provenance
- Partnerships to collaborate more effectively

These consultations have helped bring together the communities and infrastructures that connect with GBIF for targeted discussion and provided some guidance and ideas on pathways forward. With so many stakeholders this remains a complex area and will require continued discussion and global input. Consultation participants are working towards development of a shared roadmap for the Digital Extended Specimen concept.

GBIF participate in the ongoing DiSSCo Prepare project, contributing to the design work relating to the [Open Digital Specimens](#).

The Secretariat have started work that will result in a new information model for GBIF. This model will allow GBIF to better accommodate wider data types in its data infrastructure, potentially including tracking data (organism-oriented data), ecological and monitoring data (site-oriented data), automated detection (e.g. camera traps), preserved physical material (e.g. tissue or environment specimens), DNA-derived data (e.g. metabarcoding projects) and literature-based data (e.g. material citations). Following approval by the Science Committee, Budget Committee and Executive Committee, the Secretariat has commissioned a contracted expert to work during in the second half of 2021 to compile known use cases, and define a data model and initial formats for data exchange beyond the Darwin Core Archive (DwC-A). The outputs of this work will be followed by prototyping work in 2022.

As part of the EU-funded BiCIKL project, GBIF is running a [hackathon topic](#) focusing on enhancements to the clustering algorithm in September. This will bring participants from ALA, Royal Botanical Gardens, Kew and the DiSSCo project to collaborate on software development, with a goal of continued work following the event.

The GBIF IPT 2.5.0 was [released](#) in August 2021 addressing around 33% of issues and feedback identified as priorities in the 2020 consultation, including contributions from the community in translations and software improvements. A [roadmap](#) for the 2.5.x version of the IPT has been created and a lead developer appointed.

Work to revise the GBIF validator is underway (September) to bring this in line with processing with GBIF.org. In addition to consistency, GBIF aim to provide a "data preview" capability and provide an integration with the GBIF IPT, allowing data to be validated before publication in the IPT.

The work of the aligning the Atlas of Living Australia (ALA), Living Atlas community and GBIF Registries has not yet started, but remains an objective. This was delayed as the teams [completed](#) the alignment of the GBIF and ALA data processing. The teams will look to integrate this into the deployments by the Living Atlas community, to further align infrastructure and increase our collective capacity and collaboration. It is anticipated this collaboration will expand to develop the

prototype catalogue services in 2022.

The development work to support the pilot hosted portals (see Activity 1c) has resulted in a new GraphQL-based API, and a collection of reusable ReactJS-based 'widgets'.

Collaborations with academic projects were initiated to display occurrence data within a phylogenetic framework. These projects are expected to be continued and expanded in 2022 (see Activity 5a).

The following activities have not yet started (September)

- Explore standards and interfaces to allow brokering of data clustering annotations to data publishers; broaden the capabilities to include external annotations
- Review the GBIF Metadata Profile to explore the feasibility and benefits of migrating to the latest version of Ecological Metadata Language (EML)
- Revise the GBIF quarterly data analytics, to follow recent changes in data processing, such as handling counts of absence-based records

## 2022 Work items

- Continue to participate in the ongoing Open Digital Specimen and Extended Specimen Network standards development and strive towards a common solution
- Together with the Atlas of Living Australia (ALA), support the Living Atlas community, in adoption of the shared data pipelines
- Following open review of the revised data model, prototype new infrastructure that allows the demonstration of data mapping, publishing and indexing into new catalogues, developed in an open manner. This will involve development in the IPT, new data pipelines, APIs and user interfaces to explore content and communities of data publishers willing to shape data for the pilot. It is envisaged one or more developer workshops will be arranged to encourage more open development and collaboration (€40K)
- Ensure all vocabularies are included in the GBIF Registry and maintained in a collaborative manner

## 2021 Participant contributions

- **Argentina:** Complete the pending 2020 objective and be able to obtain more information and training on Frictionless Data.
- **Belgium:** Further integration of Frictionless Data, EML and Darwin Core, see 2019 tool [Frictionless Darwin Core](#).
- **Brazil:** Several ecological datasets were published on the SiBBr platform, and an ecological data repository was implemented. A continuous effort has been made to improve the visualization of these data with the integration of different variables.
- **Finland:** FinBIF has been actively developing and improving the FinBIF schema also taking into account it's compatibility to fit the data standards. FinBIF has applied for the certificate of a trustworthy data repository (Core Trust Seal). Process to meet the requirements is ongoing. FinBIF has also made preparations to become more FAIR compatible (mostly technical requirements).

- **France:** Contribution on the development of new functionalities in the IPT and support to the LA community.
- **iDigBio:** iDigBio, GBIF, ALA, and DiSSCo are making a concerted effort to coordinate efforts via the Collections Data Infrastructures Working Group (CDIWG). iDigBio regularly participates in the collaborative efforts of GBIF, BCoN, and DiSSCo, among others, to forge a path for implementation of Extended Specimens.
- **Mauritania:** site GBIF Mauritania en veuil.
- **Spain:** Plans: We plan to integrate Plinian Core compliant data in our ALA-based national data portal; linking species level information based on Plinian Core with occurrence-level data based on Darwin Core. Progress: **Plinian Cores** specification and documentation has been refined and put in a form complaint wit the TDWG SDS. Developments works to integrate Plinian Core into the ALA architecture are being delayed. We lead de revamp and update of the TDWG's **World Geographical Scheme for Recording Plant Distributions (WGSRPD)**. We participate in the **TDWG "Biological Interactions Interest Group**.
- **Sweden:** 2021: GBIF-Sweden has continued to participate in the ongoing Open Digital Specimen and Extended Specimen Network standards development. Dissemination of the ODS concept is in progress locally and nationally. We continue to support and participate in the Living Atlas community adopting shared data pipelines. We stay updated but do not participate in practice in the development of the "new IPT". We are prepared to examine and contribute to the pilot.
- **Uganda:** Data quality remains a key issue due to the different data formats available amongst stakeholders. Integrated Publishing Toolkit has been applied in the stakeholder data publication efforts. This was so with technical help from the French node.

## 2022 Participant plans

- **Andorra:** More participation in the Living Atlases community to better understand the new potential of new data standards.
- **Argentina:** Complete the pending 2020 objective and be able to obtain more information and training on Frictionless Data.
- **Belgium:** Further exploration of modern data packaging such as Frictionless Data.
- **Brazil:** Research on tools for ecological data integration will be continued.
- **France:** Aiming to adopt new data pipelines and contribute to the implementation and testing of new development. Willing to take part on the developer workshop if possible
- **iDigBio:** iDigBio will continue its partnerships, discussions, and collaborations via CDIWG and BCoN International Partners. iDigBio will also continue to promote data standards and vocabularies via its many training and workforce development activities.
- **Mauritania:** Plans to collect new biodiversity data and share it with GBIF.
- **Spain:** To continue standard development works (Plinian Ciore, GeoSchemes, Interactions). Impulse, develop and complete the integration of Plinian Core into the ALA architecture.
- **Sweden:** 2022: Still progressing with the setting up of SBDI we will follow, and where appropriate participate/contribute to data model work and implementation.
- **Taxonomic Databases Working Group:** Certainly TDWG members have been part of this effort

and will continue to be. As current Chair, I'm also very interested in how we help people implement and successfully adopt updated, modernized? data standards.

- **Uganda:** Conduct a data management training for potential data publishers in Uganda.

## Activity 2b: Deliver names infrastructure

### Rationale (for the period 2017-2022)

The most significant challenge to improving the quality of aggregated occurrence data is the continuing need for a comprehensive checklist of known species, and even for a comprehensive list of published scientific names. Interpreting and mapping names depends on the quality and completeness of these resources. Even in cases where names in occurrence records are incorrect or misspelled, better names infrastructure can assist by increasing confidence that fuzzy match algorithms or human intervention is required.

Delivering these resources is the focus of a number of GBIF Participants and other stakeholders, including the Catalogue of Life partnership, WoRMS, nomenclators (IPNI, Index Fungorum, ZooBank) and many national, regional or taxonomic databases. A comprehensive resource for scientific names and taxon concepts organized at least as a workable reference classification (but with support for additional classifications as appropriate) would also benefit other infrastructures, including Encyclopedia of Life, Biodiversity Heritage Library, Barcode of Life and GBIF nodes, and improve interoperability between data from these infrastructures. It would also be beneficial to accommodate vernacular names, informal names for undescribed species and other identifiers such as Barcode Index Numbers.

### 2021 Progress

The new Catalogue of Life infrastructure, powered by and hosted at GBIF, was launched in 2021. This includes a new **public website**, **COL ChecklistBank**, and a new **API**. Monthly releases of COL are **available** and the first **annual release** committing to use stable record identifiers.

The outstanding development tasks were completed for the Catalogue of Life to resume **monthly releases**. This included the introduction of stable identifiers for the taxon names exposed in the Catalogue of Life.

Citation practice based on Digital Object Identifiers (DOIs) has been introduced to the datasets that constitute the Catalogue of Life. The COL Checklist 2021 has been issued with associated DOIs for the checklist and underlying species data resources.

Development work is still required around the management of user communities within the Catalogue of Life. Once complete, COL and GBIF data management groups can develop the necessary processes around tasks.

A taxonomy curation pilot project run in collaboration SANBI and Canadensys to update the Fabaceae sector of the GBIF taxonomic backbone. Metrics of occurrence data improvement are being generated.

The March 2021 update of the **GBIF Backbone Taxonomy** included the work of the **Genome**

**Taxonomy Database** to improve the Bacteria and Archaea sections. A new edition of the GBIF backbone will be created during September using the August Catalogue of Life release.

A tool to allow COL data managers to assess the differences between the GBIF and COL taxonomies when treating occurrence data has **been deployed** for early testers and will be revised based on feedback. This will help provide clear priorities on the segments in the taxonomy that require attention.

The focus of the later part of 2021 is on building and integrating an extended Catalogue of Life in GBIF as a replacement for the GBIF Backbone Taxonomy, partly within the framework of the the EU-funded BiCIKL project. This will include assessing and reporting on gaps for organizing GBIF occurrence records, and broadening the community of contributors.

## 2022 Work items

- Complete outstanding tasks to integrate the extended Catalogue of Life in GBIF, partly within the framework of the BiCIKL project
- Maintain and update processes for constructing the GBIF Backbone Taxonomy, including monitoring the content and helping to prioritize editorial effort. €108,000 has been allocated in the budget to support GBIF costs. This work is in collaboration with Catalogue of Life.
- Explore feasibility of supporting national taxonomies for exploring GBIF occurrence data to better enable national-level reporting (carried over from 2020 and likely to be incomplete in 2021)
- Continue and expand the taxonomy curation pilot projects in collaboration with COL and the taxonomic community.

## 2021 Participant contributions

- **Argentina:** Complete the pending 2020 objective.
- **Brazil:** Some reference lists such as the CITES species list and the invasive species list have been added to SiBBr national platform, but the taxonomic reference list for microorganisms has not yet been added.
- **Chinese Taipei:** TaiBIF developed a protocol for exchanging scientific names among databases in Taiwan, and is developing a taxonomic information management tool to facilitate the work of taxonomists and integrate different taxonomic views.
- **Finland:** FINBIF maintains a national checklist of Finnish species, which is publicly available in the portal Species.fi. FinBIF gathers up-to-date information about Finnish species into the checklist, which also functions as a basis for handling observation and specimen data. An annual version of the checklist was published in January. The checklist is maintained with a tool "**TaxonEditor**" developed by FinBIF. Nordic cooperation has been ongoing to harmonize the taxonomy and map the taxon concepts with **tool TaxonID**. More info: Piirainen E., Laiho E.-L., von Bonsdorff T., Lahti T. (2019) Managing taxon data in FinBIF. *Biodiversity Information Science and Standards* 3: e37422. <https://doi.org/10.3897/biss.3.37422>
- **France:** Update of TAXREF, the French national checklist.
- **iDigBio:** iDigBio uses the GBIF Backbone Taxonomy in its data ingestion processes.

- **Korea, Republic of:** We are continuously updating and correcting the incorrect or misspelled names in occurrence records.
- **Netherlands:** NLBIF has financially contributed to the development of the new COL and has funded (30k euro) a new COL application to 1) mobilize national (Dutch) species checklist through COL, 2) develop training materials on the use of COL, and 3) to expand the Dutch user community of COL.
- **South Africa:** Efforts to mobilize taxonomic data and support the names infrastructure has been conducted.
- **Spain:** We will continue working on the activities already ongoing in 2020 in this area.
- **Sweden:** 2021: GBIF-Sweden continues to offer updated versions of national checklist "Dyntaxa" connected to CoL. Also bringing in Micro-fungi and OTU:s/MOTU:s, Bacteria, and Archaea.
- **Switzerland:** National checklists registered for **Butterflies** and **forester moths**.
- **Uganda:** Some of the stakeholders have been trained by the BID project on this aspect of work.
- **United Kingdom:** RBG Kew will continue to work with GBIF/COL+ partnership in providing nomenclatural and taxonomic data for use in the COL+ infrastructure.

## 2022 Participant plans

- **Andorra:** Correction of the plant species names according to the new Pyrenean reference system.
- **Argentina:** Complete the pending 2020 objective.
- **Brazil:** The reference list for microorganisms should be included in the national platform, in addition to updating the current lists of species of flora and fauna.
- **Chinese Taipei:** TaiBIF will expand the current name-based scientific name ID system and integrate a taxon-based ID system.
- **Finland:** FinBIF is planning to map the national taxonomy to GBIF backbone taxonomy (COL+).
- **France:** Publishing of new version of TAXREF.
- **iDigBio:** iDigBio will continue to use the GBIF Backbone Taxonomy in its data ingestion processes.
- **Mauritania:** Develop national taxonomies in order to explore GBIF occurrence data thus enable reporting at the national level.
- **Netherlands:** Support the further development of COL where possible.
- **South Africa:** Efforts to mobilize taxonomic data and support the names infrastructure will continue.
- **Spain:** We participate in the development and maintenance of the national official taxon name lists. We work in coordination with Ministry for the Ecological Transition, and in partnership with other **CSIC's research groups**.
- **Sweden:** 2022: Continued work with checklist data for various groups of organisms.
- **Switzerland:** Publication of national species checklists for red list groups and important invertebrate groups (continued).

- **Uganda:** To continue supporting the use of Digital Object Identifiers by data publishers in Uganda.
- **United Kingdom:** RBG Kew will continue to support GBIF/COL+. We will also release a name registration system for vascular plants based on IPNI.

## Activity 2c: Catalogue collections

### Rationale (for the period 2017-2022)

Natural history collections are the largest source of data on biodiversity outside nature itself. Collectively the world's natural history collections number about 3 billion specimens and document over 300 years of active human exploration of biodiversity on earth. In addition the fossil remains provide us with glimpses into the very far past before humans existed. The collections and their ancillary materials (images, collectors' notes, sequences, measurements, etc.), contain colossal amounts of data that should be digitized and shared. Only about 10% of the world's collections have been digitized and only a portion of digitized collections are shared publicly through the internet. A large number of current GBIF publishers comprise natural history museums and herbaria. The Secretariat will work with Participants to deliver the most comprehensive catalogue possible of collections, including metadata to publicize undigitized collections as a first step towards their digitization and mobilization.

### 2021 Progress

The Secretariat created a [road map](#) in 2021 for the [GBIF Registry of Scientific Collections \(GRSciColl\)](#), focusing on the development necessary to allow wider external contribution, and to mature the processes around editing. Editors may now be given scoped responsibility at institutional or national level. Induction webinars have been held with several nodes, resulting in a global team of 45 editors and 12 mediators actively curating content in the registry as of July 2021. Training videos are being developed.

The [iDigBio](#) collection catalogue is now powered by GRSciColl, through its open APIs. iDigBio data managers edit directly through the online [editing interface](#).

Documentation for editors is in progress and can be expected in Q3 2021. All GRSciColl fields are associated with an English-language description available in the online forms. Capabilities for anyone to suggest a change were deployed in May. Proposed data changes are reviewed by the pool of editors and mediators before being applied.

Capability for user interface translations is set up to support multilingual content. Editors and external communities in support of the catalogue are invited to propose translations to support their work. This has already been taken up by the Society for the Preservation of Natural History Collections (SPNHC) Biodiversity Crisis Response Committee.

A service has been deployed allowing the linking of collections in GRSciColl to specimen records in GBIF. This has resulted in 134 million records being linked to GRSciColl entries. A basic data dashboard is now available for institutions and collections such as [this example](#).

Options for a richer user interface for GRSciColl are being considered within the hosted portal

framework (see Activity 1c). There is work remaining on the data model, and how to make best use of the data clustering to link related data before this can progress. It is anticipated the collections catalogue user interface will continue into 2022.

Adoption of persistent identifiers, such as Research Organization Registry (ROR) identifiers is actively being discussed within European nodes (e.g. through DiSSCo) with some nodes piloting use of ROR IDs on their entries.

Integration with the CETAF registry remains an objective, but has not yet started due to the effort required to enable external editors and focus on content issues. It is expected to extend into 2022. Piloting a profile of the TDWG Collection Descriptions to capture collection-level metadata has also been delayed and will be considered for 2023.

## 2022 Work items

- Complete outstanding tasks to deploy an enriched GRSciColl providing search and access of collections, specimens and people
- Focus on content of GRSciColl: cleanup of existing entries and registration of new ones by promoting use and giving training and support to editors, and promoting consistent use of codes within data shared
- Seek to identify links between journal articles and collections based on the collection codes, within the framework of the EU-funded BiCIKL project
- Support user interface translations for GRSciColl
- Carried over from 2020: Explore synchronization of content with the [Consortium of European Taxonomic Facilities \(CETAF\)](#) Registry (under development)

## 2021 Participant contributions

- **Andorra:** None. This type of data is practically non-existent in Andorra.
- **Argentina:** Keep working with our catalogue and collaborate with the GRBio / GRSciColl initiative presented at GBIF.
- **Belgium:** pilot GRSciColl cleanup for Belgian institutions.
- **Brazil:** It was not possible to add the DOI on the national platform, but the registration of collections on SiBBR was expanded and improved in several aspects, such as the addition of a tool that enables the monitoring of downloads.
- **Chinese Taipei:** TaiBIF is developing a web portal for cataloguing the natural history collections in Taiwan and sharing their information.
- **France:** Support with review and metadata about French Biological Collections to GRSciColl.
- **iDigBio:** iDigBio staff actively participate in TDWG and the Natural Collections Description (NCD) group. iDigBio is an active partner with GBIF on the global collections registry (GRSciColl).
- **Korea, Republic of:** KBIF will support other institution to catalogue their natural history collections, and share the data with GBIF.
- **Mauritania:** Digitization of several data on halophile plants and wetland biodiversity.

- **Netherlands:** NLBIF as DiSSCo-NL national node representative has provided a workshop item for DiSSCo partners on harmonizing institutional and collection identifiers at GRSciColl as a first step to start providing full collection descriptions and to monitor digitization progress.
- **South Africa:** The Natural Science Collections Facility, funded by the Department of Science and Technology plays a key role in the development of catalogues for collections.
- **Spain:** Plans: We plan to work in ways to synchronize our Registry and the GBIF Registry of Scientific Collections (GRSciColl) to make available the most complete, coherent and updated information through these two outlets. Progress: We are finalizing the updating and publication "National Registry of Natural history Collections and Biodiversity Databases" which will be the base for in turn review and synchronized it with the GRSciColl.
- **Sweden:** 2021: Discussion over GRSciColl contribution has been held locally but not yet reached practical work level. Swedish Museum of Natural History (SMNH; SBDI's and GBIF-Sweden's principal organization) member of CETAF.
- **Uganda:** The Fresh water occurrence data sets are being digitized and published in the National freshwater biodiversity data portal.

## 2022 Participant plans

- **Argentina:** Coordinate and improve the data of the institutions / collections / staff of Argentina in the GRSciColl.
- **Belgium:** Maintain GRSciColl for Belgian institutions and collections.
- **Brazil:** Continue the registration of collections with an evaluation to improve the metadata completion form by curators.
- **Chinese Taipei:** TaiBIF will continue the web portal development.
- **Finland:** FinBIF is improving the findability of metadata by providing a publicly available catalogue for metadata search - so far the metadata, although open data, has not been easily browsed or searched except in a closed system (collection management system). Also the process for improving the quality of metadata is ongoing in various collections.
- **France:** Continue support with review and metadata about French Biological Collections to GRSciColl.
- **iDigBio:** iDigBio will continue to support TDWG and the Natural Collections Description (NCD) group. iDigBio will continue to work with GBIF on the global collections registry.
- **Netherlands:** NLBIF as DiSSCo-NL national node representative will provide guidance to the Dutch DiSSCo partners on mobilizing collection description data to GRSciColl following the TDWG CD (Collection Description) and CETAF passport metadata standards.
- **South Africa:** The Natural Science Collections Facility, funded by the Department of Science and Technology plays a key role in the development of catalogues for collections, and this will continue.
- **Spain:** To review and synchronize the "National Registry of Natural history Collections and Biodiversity Databases" with the GRSciColl.
- **Sweden:** 2022: Continued effort to mobilize Swedish natural history collections to add information to GRSciColl, linking also to publications. Swedish Museum of Natural History (SMNH;

SBDI's and GBIF-Sweden's principal organization) member of CETAF.

- **Switzerland:** Update and completion of GRSciColl Swiss institutional records at collection level (joint survey with SwissCollNet).
- **Taxonomic Databases Working Group:** Of course, TDWG members have been actively working on the Collection Descriptions (CD) standard to support this much-needed metadata. We look forward to seeing it implemented and supporting its further development.
- **Uganda:** More data to be digitized and utilized in the conservation decision making in Uganda.

## Priority 3: Fill Data Gaps

Prioritize and promote mobilization of new data resources which combine with existing resources to maximize the coverage, completeness and resolution of GBIF data, particularly with respect to taxonomy, geography and time.

### Activity 3a: Identify priority gaps

#### Rationale (for the period 2017-2022)

GBIF has a range of tools, including fitness-for-use groups, other community consultations, feedback channels, direct communication with authors of scientific studies, and societal demands, to identify and collect data needs. Addressing data gaps may require focus on gaining additional occurrence records, targeting data areas missing from published records, or getting additional metadata elements. By consolidating and prioritizing demands for data content, in the context of already accessible data and knowledge of resources which are not yet available as open data, GBIF will be positioned to inform collection and data holders, funding institutions and political decision makers of the most worthwhile and cost-effective ways to extend the available knowledge base.

#### 2021 Progress

Analytics work in 2021 focused on the exploration of taxonomic gaps, checking for areas in GBIF-mediated occurrences that are insufficiently mapped to the Catalogue of Life (CoL). No major areas were discovered, most gaps being detected in the area of Operational Taxonomic Units (OTUs) rather than Linnean taxonomy. Minor results in that area were fed back to CoL's issue tracking system.

Updated guidance on prioritizing data mobilization based on analysis of data gaps was being commissioned in 2021 as part of the GBIF documentation programme (see Activity 1b). This guidance will take account of previous work carried out by the Secretariat during the current implementation period to analyse gaps and biases in data available through GBIF.org, such as those reported to the 2020 regional meetings and [documented in the GBIF data blog](#).

The Secretariat continued to work with the [Knowledge and Data Task Force of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services \(IPBES\)](#) to catalyse investment in priority research and data mobilization to address knowledge gaps identified

in the IPBES assessment process. This will include participation in a series of regional dialogues with research funders and programmers, to identify potential funding sources to address data gaps.

In the context of discussions on the post-2020 Global Biodiversity Framework carried out through the Convention on Biological Diversity (CBD), the Secretariat has begun discussions with partners including the Group on Earth Observations Biodiversity Observation Network (GEO BON) and Map of Life (MoL) on how indicators of data completeness may be used to prioritize data mobilization, supporting information targets under the new framework (see Activity 5c).

## 2022 work items

- Publish updated best practices on prioritizing data mobilization based on data gap analysis, for incorporation into guidance for nodes and data publishers
- Explore the use of data gap analysis to prioritize support for data mobilization through funded programmes coordinated by GBIF, and to document the value added by existing projects
- Work with partners to address the information needs of the post-2020 Global Biodiversity Framework through filling priority data gaps, including for those countries not yet participating in GBIF

## 2021 Participant contributions

- **Argentina:** Continue with the publication of the national inventory of biodiversity and national lists of species. On the other hand, it is expected to feed the national portal with these national listings.
- **Brazil:** As there are long-term research programs in Brazil, covering different geographic regions, an effort was made to publish these data on SiBBr, with training and help desk directed at researchers. Other researchers working with viruses also had training on structuring and standardizing data on DwC.
- **Finland:** FinBIF has prioritized sharing the sensitive and restricted-use data to decision makers. The data is provided to public authorities through a separate portal in which all the data is available to the same extent as it is stored in the data warehouse. The use of [this portal requires authorization](#).
- **France:** We are working on the connection of DEPOBIO, the official national platform for Environmental Impact Assessment (EIA) data to GBIF.
- **iDigBio:** iDigBio has been working internally and with GBIF to ensure that all data ingested in iDigBio are also ingested in GBIF.
- **South Africa:** Continued engagement of SANBI-GBIF on IPBES task force to support knowledge gaps. Also, the Node will continue efforts to mobilize data through stakeholder engagement. FBIP initiative to continue funding data generation and filling of knowledge and data gaps.
- **Spain:** Plans: Besides continuing the activities with the private sector and environmental administrations, we plan to recover and activate the participation of collections in the light of initiatives such as Synthesys, DiSSCo and CETAF; and increase the involvement of the LTER Community in GBIF through our participation in the eLTER RI projects (eLTER Plus and eLTER PPP)  
Progress: Work goes according to plan
- **Sweden:** 2021: Started work on preparing and publishing information about taxonomically

underrepresented groups (primarily micro-organisms) from hitherto unavailable datasets.

- **Uganda:** The priority has been on embracing open data principles in the Ugandan Public institutions and this is still ongoing.

## 2022 Participant plans

- **Andorra:** We must first collect the potential data from the different Andorran administrations to be able to identify the gaps.
- **Argentina:** Continue with the publication of the national inventory of biodiversity and national lists of species. On the other hand, it is expected to feed the national portal with these national listings.
- **Brazil:** A survey of biological collections in Brazil is in progress, which will be used for a gap analysis, geographical and temporal, with support from the CESP Project "The role of Zoological Collections Networks to enhance and fill taxonomic and geographic gaps in open access databases"
- **France:** Continuing the activities with the private sector and environmental administrations, and pursue the links with the actors of DiSSCo and GEO BON at national level.
- **iDigBio:** iDigBio will continue its work to expand knowledge of what is contained within collections. iDigBio will continue to work to ensure that all data ingested in iDigBio are also ingested in GBIF.
- **Korea, Republic of:** Data gap discovery will be constantly carried out for the data from National Science Museum and some other KBIF data providers.
- **South Africa:** IPBES task force work to continue including discussions on knowledge catalysis and data management plans.
- **Spain:** Same as 2021 plus our targeted contacts with private sector, administrations and LTER researchers, to be expanded with the inclusion of the remote sensing community
- **Sweden:** 2022: Continued work with OTU:s and more - and larger - sets of monitoring data will be provided. Updating already present collections datasets is another priority of GBIF-Sweden - unfortunately slightly neglected due to the heavy focus on organizing SBDI during the recent time period.
- **Uganda:** To contribute to the finalization of the Open data policy for Uganda as we support the establishment of the one stop centre for all freshwater biodiversity data and information for Uganda.

## Activity 3b: Expand data streams

### Rationale (for the period 2017-2022)

GBIF serves as an integration point for any source of evidence of the recorded occurrence of species in time and space. A primary role for the GBIF infrastructure is to serve as a comprehensive single point of access for discovery, access, use and curation of all such evidence. Several classes of data are already well-supported within the GBIF network.

These include collections data, observations from field research, and many categories of citizen

science data. However, there are other new and developing streams of data which should be accommodated if GBIF is to serve as the platform for supporting comprehensive data assessment and modelling (e.g. for GEO BON Essential Biodiversity Variables, IPBES assessments, Red List assessments, etc.). These include efforts to mine historical data records from literature, genomics activities and particularly barcode-driven surveys, and potentially species-level data from remote-sensing systems. More work is also still needed to engage with the full spread of research activities delivering sampling event data of various kinds. GBIF needs to ensure that it provides simple, effective and beneficial ways for researchers to share these and other streams of Darwin Core compatible data.

## 2021 Progress

The guide to [Publishing DNA-derived data](#) has been released in its final version with the accompanying [Darwin Core extension](#). During the remainder of 2021 GBIF will work with early adopters in preparing data following the guidelines. Mobilization of DNA-derived data was a priority in selection of the projects funded through the latest call under the Biodiversity Information For Asia (BIFA) - see Activity 1g. New training resources combining skills needed for DNA barcoding and GBIF data mobilization were developed and tested during 2021, for rollout in workshops planned under the BioDATA programme later in 2021 and 2022.

Efforts in 2021 to expand the mobilization of ecological datasets through GBIF included discussions with the German Centre for Integrative Biodiversity Research (iDiv) on the establishment of a repository for sampling-event datasets from individual researchers that would be published through GBIF. These discussions are connected with GBIF's involvement in the [Humboldt Core task group in TDWG](#) exploring improved standards for exchanging data on species inventories. The Secretariat also engaged with the [EuropaBON](#) and [Tropical Andes Observatory](#) projects to improve connections between GBIF and biodiversity monitoring activities in Europe and South America, respectively. The Secretariat was also invited to join the [IUCN Species Survival Commission Species Monitoring Specialist Group](#).

During 2021, the Secretariat developed a framework for running data mobilization campaigns based on thematic data streams, including engagement of expert groups, identification of potential mobilization targets, and help desk services to support data formatting and publication in collaboration with nodes. Specific mechanisms are being trialled with regard to data on disease vectors and hosts, as well as soil biodiversity data (see Activity 5b), and data mobilized through the private sector (see Activity 3c).

## 2022 Work items

- Building on the framework developed in 2021, establish data mobilization campaigns for additional thematic data streams (possibly freshwater, DNA-derived data and tropical biodiversity) - €40k.
- Seek supplementary funding to support scaling up of support for thematic data mobilization throughout the GBIF nodes network (see also Activity 3c)
- Continue working with partners to increase mobilization and improve visibility of biodiversity monitoring datasets in GBIF

## 2021 Participant contributions

- **Argentina:** Continue with the promotion of the use of sampling event data model for ecological and monitoring datasets
- **Brazil:** Campaigns were carried out to publish long-term research data on SiBBr, with pages for these [datasets](#)
- **Finland:**
  - PROTAX, ASI
  - [GGBN](#)
  - [Red list assessment](#)
  - [Invasive alien species](#)
  - R-package
  - Earth Cape
- **iDigBio:** iDigBio hosted a virtual workshop “Linking and Leveraging Biological Collections: Zoos and Natural History Museums” on 3-5 May 2021 that brought together representatives from living animal collections (zoos), traditional animal preserved specimen collections (natural history museums), universities, a botanical garden, and others.
- **Spain:** Plans: Without abandoning current and consolidated data strings, we plan to focus on LTER communities. Progress: Work goes according to plan.
- **Sweden:** 2021: Preparations were made by SBDI/GBIF-Sweden to mobilize DNA-derived data, including training needed for DNA barcoding and GBIF data mobilization. Another data stream receiving attention by GBIF-Sweden refers to individual ecological research datasets.
- **Uganda:** There is positive coordination among the Clearing House Mechanism under the Convention on Biological diversity with GBIF mediated platforms to harvest data from all possible sources for decision making.

## 2022 Participant plans

- **Andorra:** Start working with partners to increase mobilization and improve visibility of biodiversity monitoring datasets in GBIF.
- **Argentina:** Continue with the promotion of the use of sampling event data model for ecological and monitoring datasets.
- **Belgium:** DNA-derived data project with other francophone nodes.
- **Brazil:** Several partnerships will be maintained with the group of researchers working with viruses and their hosts, with the publication of these datasets on SiBBr. It is expected to advance in the articulation with the private sector, seeking the publication of data from inventories and assessments of biodiversity.
- **France:** We would like to rely on the new training resources combining skills needed for DNA barcoding and GBIF data mobilization to organize a workshop in 2022 for francophone participants.
- **Korea, Republic of:** Illustrated books using KBIF taxonomic data with images has published 2

books in 2021.

- **Netherlands:** NLBIF has funded two projects, one on mobilizing bat sound data and another on mobilizing MOVEBANK data to GBIF.
- **Spain:** Adding the Remote Sensing Community to our portfolio of biodiversity data streams poorly represented to be targeted and incorporated into GBIF.
- **Sweden:** 2022: In addition to continued work of aforementioned topics, long-term forest inventories data (starting in the early 20th century) will be presented in 2022.
- **Taxonomic Databases Working Group:** TDWG will continue work via the Humboldt Core task group in TDWG exploring improved standards for exchanging data on species inventories.
- **Uganda:** Continued cooperation and collaboration with all stakeholders.

## Activity 3c: Engage data holders

### Rationale (for the period 2017-2022)

GBIF has tools and support mechanisms in place to enable publication of several categories of biodiversity data. The key requirement is for increased engagement with, and support for, the communities of institutions and individuals who hold these data. Such engagement is normally most effective at the national level, although international networks also have an important role to play.

### 2021 Progress

The Secretariat has continued to engage with nodes and external partners to support sharing of data from the **private sector**, in particular through monitoring activities connected with environmental and social impact assessments. An updated **guide on publishing primary biodiversity data through impact assessments** was published in partnership with the International Association for Impact Assessment (IAIA).

The **Data4Nature initiative**, designed in partnership with **Agence Française de Développement (AFD)** to engage development finance actors in sharing of biodiversity data, achieved visibility during the virtual meeting of the CBD Subsidiary Body on Implementation (SBI3) with support from the European Union's delegation to the CBD, and further buy-in for this initiative was planned for the IUCN World Conservation Congress taking place in Marseille in September 2021. Interest in supporting this initiative and other activities focused on mobilizing private sector data through GBIF has come from the UN Environment World Conservation Monitoring Centre (UNEP-WCMC) through the Proteus Partnership, bringing together corporations in the oil, gas and mining sectors.

### 2022 Work items

- Building on preparatory work and partnerships developed in 2021, scale up mobilization of primary biodiversity data from the private sector, through the Data4Nature initiative and establishment of help desk services
- Engage additional sectoral associations, financial and other initiatives to integrate biodiversity data-sharing commitments into best practices for Environmental, Social and Governance (ESG) or regulatory requirements

## 2021 Participant contributions

- **Argentina:** Keep promoting the CNC. Keep promoting the CC through RICAP and National Node
- **Brazil:** We have not made significant progress with the private sector. Regarding citizen science data, a tutorial interface was created encouraging the sharing of records through the use of the iNaturalist application.
- **Chinese Taipei:** TaiBIF has been working closely and holding monthly meetings with five major data holders/users in Taiwan through Taiwan Biodiversity Information Association, which will be officially established after signing a MOU among the six organizations/agencies on September 29, 2021.
- **Finland:** FinBIF has launched a service “Data Bank” which consists of tools for data holders so that they can easily share their data to **FinBIF**. FinBIF has also started cooperation with Finnish municipalities - almost 70 have already signed an agreement to start using the Public Authorities Portal.
- **France:** GBIF France is involved in the discussions with the French development agency (AFD) and the GBIF Secretariat in the Data4nature initiative. We organized a training session in August for AFD staff and will take part of the event at the IUCN World Conservation Congress taking place in Marseille in September 2021. Due to the pandemic, we organized the webinar for the private sector linked with the CESP project “OpenPSD - Engage and promote the private sector” led by GBIF Spain in early 2021.
- **iDigBio:** Data mobilization is a core activity of iDigBio, including reaching out to collections.
- **International Centre for Integrated Mountain Development:** Support publishers from the HKH countries- probably next batch of BIFA grantee from the HKH countries. Endorsed research proposal for two BIFA applicants. Publish ICIMOD based biodiversity data from different landscape initiatives. Published 3 datasets from HILIFE landscape and 1 from individual data holder from Nepal. A checklist of Mistletoes of Nepal. List of Medicinal Plants of Gaoligong Mountains. Native Orchids of Gaoligong Mountains. Amphibia and Reptilia of Yunnan. Provide technical services to data publishers in the HKH countries as required- in relation to data publishing in HKH-BIF. Supported BIFA grantee from Nepal to publish six datasets on Mushrooms and Lichens via HKHBIF. **Lichens Flora of Nepal at Central Department of Botany, Tribhuvan University, Nepal** **Lichens Flora of Nepal at National Herbarium and Plant Laboratories (KATH)** **Lichens Flora of Nepal at Natural History Museum, Tribhuvan University, Nepal** [http://rds.icimod.org:8080/hkh-bif/resource?r=lichens\\_nhm](http://rds.icimod.org:8080/hkh-bif/resource?r=lichens_nhm) **Mushrooms Flora of Nepal at Central Department of Botany, Tribhuvan University, Nepal** **Mushrooms Flora of Nepal at National Herbarium and Plant Laboratories (KATH)** **Mushrooms Flora of Nepal at Natural History Museum**
- **Netherlands:** NLBIF has assisted several new data publishers to mobilize their biodiversity data among which a private sector ecological consulting partner.
- **South Africa:** SANBI-GBIF and the Systematics Division will implement the 2021 Foundational Biodiversity Information Management Forum FBIF Forum, which is a key event to engage the data holders in the community. This will be virtual for 2021 and will be a series of smaller engagements.
- **Spain:** Plans: We plan to maintain connections made within OpenPSD and --building upon it-- expand participation of private sector in GBIF. Collaborating with other GBIF nodes and

Secretariat is a given. Progress: Besides maintaining contacts and support with partners identify during the OpenPSD Project, we maintained contacts and discussions with GBIFS (Goodson, Copas) regarding involvement of private sector in GBIF and future plans.

- **Sweden:** 2021: By teaming up in a unified organization (SBDI), a considerable number of "lukewarm" providers have a stronger commitment to arrange, standardize and begin publishing data with SBDI and thus GBIF-Sweden. This includes governmental agencies, universities (mainly individual researchers but also monitoring programmes run on behalf of agencies). We have not yet seriously confronted private or semi-private data holders but expect to start doing this in the near future. Yet, we have to be cautious not to interfere very strongly with the public conservation management sector in Sweden since this domain has so far by itself and by various responsible agencies been expected to be indigenously run a in parallel to GBIF using idiosyncratic tools for data publication and extraction/analysis. As SBDI we now have the mandate to become "the major, single gateway to Swedish biodiversity data".
- **Uganda:** Data holders are encouraged through online help desk support and information sharing.

## 2022 Participant plans

- **Argentina:** Keep promoting the CNC and the GSB. Keep promoting the CC through RICAP and National Node
- **Belgium:** Engage with Belgian data holders to mobilize DNA-derived data
- **Brazil:** Mobilization for citizen science data, as well as from the private sector should be expanded.
- **Chinese Taipei:** The association plans to recruit more data holders in Taiwan.
- **France:** We plan to maintain connections made within OpenPSD and AFD and building upon it, expand participation of private sector in GBIF.
- **iDigBio:** iDigBio will continue its ongoing data mobilization efforts.
- **International Centre for Integrated Mountain Development:** in 2022, we will promote HKHBIF data publishing platform to data holders and also motivate experts, students, and citizens to share information about biodiversity via HKHBIF web platform and contribute stories, research highlight, pictures, species of the month and more.
- **Netherlands:** NLBIF continues to expand the GBIF network through regular node activities and financial support through the NLBIF-call.
- **South Africa:** Pending COVID outcomes the annual Forum will continue either in person or virtual.
- **Spain:** To continue 2021 activities in this regard.
- **Sweden:** 2022: GBIF-Sweden is aware of many data holders which will be approached in due time but as described for 2021, we need to continue soliciting ourselves as SBDI and motivating our existence before being fully accepted by all potential providers. GBIF-Sweden has been accepted but with some suspicion from several sides: collections managers have always complained – without realizing that the responsibility is in fact mainly theirs – about data quality issues, whereas environmentalists have seen GBIF-Sweden – again without being aware of the real situation - as serving mainly collections managers and researchers with far too sophisticated data

for their meaningful use. During the entire period of the present Node Manager holding his position (2009-2021) he has constantly struggled to convince both sides to accept the other!

- **Switzerland:** Publication of DNA-derived datasets. Coordinated activities with SwissBOL and InfoSpecies.
- **Taxonomic Databases Working Group:** Hm. Via TaxonWorks and the Species File Group, certainly I'm in the position of helping to engage data holders to share their data (collections, nomenclature, trait data, etc). And, in this role of Chair of TDWG, certainly we're involved in trying to encourage and expand data contributors.
- **Uganda:** This will continue and more data holders will be brought on board.

## Activity 3d: Rescue datasets

### Rationale (for the period 2017-2022)

Many researchers hold potentially valuable data which are not yet in a suitable digital format for integration into GBIF. Historical publications are a similar source of valuable data which remain inaccessible. This offers an opportunity to establish a community platform to capture information on such datasets where the researcher or owner lacks the time or capability to make the data available as a GBIF-compatible dataset, and to enable interested individuals to volunteer time to collaborate with the owner to publish a dataset, potentially in conjunction with a data paper credited to all parties. Such a model may address a key bottleneck in bringing valuable data online.

### 2021 Progress

Converting the growing collection of suggested datasets, submitted through the **user interface**, is work in progress, with most of the 228 entries (as of July 2021) categorized for geographic coverage, country of origin, and broader subject area. All entries are available in the adjoined **GitHub repository**, and relate to data from and about 40 countries, including datasets with a wider geographic scope.

Support for the uptake and data mobilization by the wider GBIF community requires re-evaluation of this work program item in the second half of 2021. With wide variations in the level of detail supplied with submissions, prioritization remains a major challenge, and requires more in-depth evaluation of individual suggestions. A more targeted approach, working with nodes and thematic communities interested in the uptake of suggestions (see Activity 3b), is required to make more efficient use of the resources available.

### 2022 Work items

- Continue to implement workflow for prioritizing and drawing upon potential data sources reported through the 'suggest a dataset' tool, including involvement of nodes, mentors and crowdsourced solutions (carried over from 2021)
- Based on the outcome of the re-evaluation, work with interested parties to roll out a workflow for processing data mobilization requests arising from the 'suggest a dataset' tool

## 2021 Participant contributions

- **Brazil:** This activity had no progress.
- **Finland:** FinBIF has mapped out valuable datasets and has continued the work to store them in the database.
- **iDigBio:** iDigBio hosts the Small Collections Network and participated in the Organization of Biological Field Stations annual meeting.
- **Spain:** Plans: Using our National Registry of Biodiversity Collections and Databases, we plan to review, check and help those not publishing data, or that published long time ago to update their participation in GBIF. Progress: National Registry of Biodiversity Collections and Databases, reviewed and updated. Collaborations with GBIFS in tracking and finding about the situation of inactive data publishers
- **Sweden:** 2021: No targeted activity within this field at present – but as described under 3c individually approached researchers known to hold valuable/interesting data are approached and asked to consider spending the extra effort needed to arrange and standardize "old" data for publication. In fact this is urgently needed as researchers retire, but also very time- and effort-consuming.

## 2022 Participant plans

- **Andorra:** Start to implement workflow for prioritizing and drawing upon potential data sources.
- **Brazil:** For the next year we hope to bring data from several collections that do not have their datasets available. An ongoing national survey will contribute to get data from different stakeholders.
- **iDigBio:** iDigBio will continue to provide digitization and workforce development activities and events that are open to all. iDigBio will also continue to provide assistance with digitization, data standardization, and data mobilization within the scope of its mission.
- **Spain:** To Issue National Registry of Biodiversity Collections and Database in various forms and review activity levels of data publisher.
- **Sweden:** 2022: Continued efforts approaching researchers and other data holders as described for 2021, but not highly prioritized in spite of some urgency.
- **Taxonomic Databases Working Group:** Again, as TDWG, we're in a support role. But from a TaxonWorks perspective, we're certainly involved in rescuing datasets that would otherwise languish.
- **Uganda:** Plans to have a citizen science program envisaged.

## Activity 3e: Liaise with journals

### Rationale (for the period 2017-2022)

Journals are the traditional established avenues for scientific communication. They not only disseminate research findings and other scholarly communications but are increasingly helping to disseminate research data. These data may be provided as supplementary materials or deposited in biodiversity data repositories as a precondition for publication of the paper. However, such data

publishing data does not necessarily facilitate its integration with other related data or make the data discoverable and reusable. In order to benefit from data disseminated through journals, the Secretariat will lead or coordinate activity to engage directly with publishing houses, journal editors and authors to promote GBIF-compatible approaches to publication and the use of GBIF-operated repositories as accepted or preferred destinations for supporting data.

## 2021 Progress

Following the successful funding of the **Biodiversity Community Integrated Knowledge Library (BiCIKL)** grant through the European Union Horizon 2020 programme, the Secretariat began collaboration with partners **Pensoft Publishers** and **Plazi** to improve integration of data extracted from research literature (see also Activity 2a).

As use of GBIF-mediated data tracked by the Secretariat in research articles continued to increase, the quality of citations passed a significant milestone in 2021, with 50% of articles now correctly citing the Digital Object Identifiers (DOIs) provided with data downloads. The use of DOIs in citations is expected to improve further with the introduction of **a new tool to generate DOIs from derived datasets** cited by researchers following further filtering or cleaning of data downloaded from GBIF.org, or accessed via public cloud services (see Activity 4a).

Following the successful call in 2020 for **sponsored data papers mobilizing data from European Russia**, using supplementary funds provided by the Finnish Biodiversity Information Facility (FinBIF), the model was repeated with a **second call** for Russian-based papers in 2021, and planned as a means of mobilizing datasets relating to vectors, hosts and reservoirs of human diseases (See Activity 5b).

## 2022 Work items

- In consultation with journal publishers, explore solutions to enable academic authors to submit and publish datasets to GBIF alongside research articles
- Further develop the model for sponsored data papers, in connection with mobilization campaigns from thematic data streams
- Continue BiCIKL collaboration to improve integration of data extracted from research literature

## 2021 Participant contributions

- **Argentina:** Continue promoting the publication of data papers in Argentine journals.
- **Brazil:** This activity made no progress in 2021.
- **Finland:** Articles published recently:
  - Schulman, Leif; Lahti, Kari; Piirainen, Esko; Heikkinen, Mikko; Raitio, Olli; Juslén Aino (2021). The Finnish Biodiversity Information Facility as a best-practice model for biodiversity data infrastructures. Scientific Data doi: 10.1038/s41597-021-00919-6.
  - Hyvärinen M-T, Väinölä R, Väre H, Ståhls-Mäkelä G, Sihvonen P, Kuusijärvi A, Myllys L, Kröger B, Heikkinen M, Juslén A, Oinonen M, Schulman L (2020) General Collections Policy of the Finnish Museum of Natural History. Research Ideas and Outcomes 6: e58167. <https://doi.org/10.3897/rio.6.e58167>

- **iDigBio:** iDigBio is continuing to develop and refine citation guidance for data obtained from its portal, which is available in the iDigBio Portal.
- **Netherlands:** The NLBIF node manager/DiSSCo-NL NN has accepted the role of liaison officer for the European Journal of Taxonomy (EJT). This European open source taxonomic journal is looking (among others) for mechanisms to persistently link specimen derived data shared through GBIF and DiSSCo in taxonomic treatments.
- **Spain:** Idem 2020
- **Sweden:** 2021: GBIF-Sweden continuously encourage data holders (including potential providers) to publish in Open Access publications, and to always prepare data papers to enhance dataset visibility. We also point out the significance of, and expect correct usage of DOI:s for citation purposes.
- **Uganda:** No progress.

## 2022 Participant plans

- **Argentina:** Continue promoting the publication of data papers in Argentine journals.
- **Brazil:** TBD
- **iDigBio:** iDigBio will continue to update and refine its citation guidance in response to user feedback.
- **Korea, Republic of:** Biodiversity research paper using the mobilized KBIF data has published in JAPB (Journal of Asia Pacific Biodiversity) published by National Science Museum of Korea.
- **Netherlands:** Continue as liaison officer for EJT.
- **Spain:** Idem 2020.
- **Sweden:** 2022: Work is in progress, but again not a high priority (e.g. we have not committed ourselves to BiCIKL) due to our primary focus on organizational matters with SBDI.
- **Taxonomic Databases Working Group:** TDWG works closely with Pensoft and has been contributing some needs (e.g. supporting multi-language publishing, using bioschemas to improve discoverability) and will be holding our joint 2022 Conference with BiCIKL.
- **Uganda:** This will be explored as well.

# Priority 4: Improve Data Quality

## Activity 4a: Ensure data persistence

### Rationale (for the period 2017-2022)

There exists a significant portion of data available through GBIF.org that is not actively curated by a data host. In some cases, there are no resources or desire to make further edits to the datasets. These datasets are effectively orphaned and the GBIF.org version of the dataset is often the last remaining version available on the internet. As GBIF develops mechanisms to provide feedback to data publishers and support curation of datasets, we need to consider that these orphaned datasets will not be updated with corrections or migrated to adhere to modern data standards.

## 2021 Progress

Work earlier in the current implementation period addressed the issue of **orphaned datasets**, with a major effort to ensure that datasets no longer connected to their original host were transferred to a 'live' version that could be successfully crawled by the Secretariat, with many nodes 'adopting' these datasets for the purposes of hosting and curation.

Following **open discussion on the Community Forum**, GBIF now provides monthly exports of GBIF occurrence records and makes them available on Microsoft Azure in the **Planetary Computer** and on Amazon AWS as **Public Datasets**. Tutorials and guides have been developed, and support will be given to early adopters. The Secretariat anticipates that GBIF-mediated data will be available in a similar manner on the Google Cloud, and made available in Google BigQuery during Q3 2020.

When implementing the new service for exports to cloud environments, the Secretariat was very conscious of the risk of losing the connection between users of GBIF-mediated data and the data publishers whose records are included in applications using these services. To address this, a new citation tool was developed in support of the cloud data services, known as the **Derived Dataset** service. This ensures users of cloud environments, or after further filtering of any GBIF download, can still provide a DOI-based citation for the content used (see Activity 3e).

GBIF has been accepted as a **Data Provider within the European Open Science Cloud**. The Secretariat will now explore opportunities to strengthen integration of the GBIF network within the wider European research data infrastructures.

## 2022 Work items

- Support adopters of cloud technologies for accessing GBIF-mediated data, and revise procedures and formats as necessary.
- Explore opportunities to strengthen integration of the GBIF network within research data infrastructures

## 2021 Participant contributions

- **Brazil:** For institutions that do not update their data for a significant time, a diagnosis was made and it was verified that one of the problems was not updating the IPT version. Thus, for these institutions, the datasets were migrated to the SiBBR's IPT.
- **France:** We spread the news on the new service for exports to cloud environments. Currently we are not directly involved in EOSC activities but we are working with people involved in EOSC projects.
- **iDigBio:** iDigBio offers a Strategic Planning online short course twice per year - one targeted at herbaria, and the other targeted at all biodiversity collections. The goal is to produce a strategic plan for each represented collection, including vision, mission, stakeholders, strategies, values, goals, objectives, evaluation, and sustainability, among other things.
- **Spain:** We will reactivate this activity set taking into account the lessons learned. Progress: We refined the way our **supported applications** provide persistent identifiers to occurrence data and other data elements handled and relevant for biodiversity data management.
- **Sweden:** 2021: GBIF-Sweden do not publish any orphaned datasets - although we have noted

some that were not recently updated by the provider. In the latter situation the provider has been contacted. We are aware of further providers that need to be pushed for updates. We have not searched for, nor received information regarding 'Derived datasets' citations.

- **Uganda:** There has been no progress regarding orphaned datasets.
- **United Kingdom:** Nicky's work will continue until March. hopefully she will be able to spend time in Copenhagen with the GBIF team. Unfortunately this work was delayed by COVID.

## 2022 Participant plans

- **Andorra:** Identified orphaned datasets.
- **Brazil:** The work of migrating the datasets of institutions that do not update their IPT will be continued in the next year
- **France:** Strengthen integration of the GBIF network within research data infrastructures is of great interest and we are eager to collaborate with fellow nodes and Secretariat in this area.
- **iDigBio:** iDigBio will continue to refine its data mobilization efforts and workflows. iDigBio will continue its work with the community on strategies and solutions for long-term data storage.
- **Spain:** To launch a campaign among GBIF Spain data providers to increase the proportion of records that are published in GBIF with persistent identifiers assigned in origin.
- **Sweden:** 2022: As part of our SBDI data mobilization strategy we will continue to push providers and update datasets as required.
- **United Kingdom:** Nicky Nicolson will continue working with GBIF staff on the resolution of duplicates and hopefully spend time with the GBIF team in Copenhagen.

## Activity 4b: Assess data quality

### Rationale (for the period 2017-2022)

Assessing data quality includes applying data validation tools to capture and monitor suspected and confirmed errors and ambiguities in data, highlighting useful areas for additional information (metadata and qualifiers) that would improve usability and enhance processing options, and documenting completeness and standardization of information both within a dataset and within aggregated data. A number of validation tools exist in the wider community, and should be brought together to mutually profit from investments and to more efficiently plan future distributed development efforts. This will benefit data publication frameworks as well as individual data holders, giving concrete feedback on best gains in data management.

### 2021 Progress

Documentation to support better and more targeted interpretation of issue labels has been continued from 2020. A popular post in the [GBIF data blog](#) from October 2020 on "[Issues and Flags](#)", supporting the interpretation and proposed action on diagnosed data issues, is at the centre of communications and further planning. The post has been maintained through additions in 2021, and will be part of the technical documentation under development (see Activity 1d). Other components are development of training resources and materials targeted to alert data users to [post-processing](#)

requirements when using GBIF-mediated data.

Ongoing work in 2021 on aligning the GBIF [data validator tool](#) with the GBIF.org data processing pipeline, will be supplemented with a community consultation on data validator functionality from Q3 2021.

GBIF participates in the EU-funded [DiSSCo Prepare](#) project, supporting the group with data exports to assess existing content. This has included the creation of a [DiSSCo network entity](#) to easily download the related content that is shared from the European institutions, and support for using Google BigQuery to assess the current [Minimum Information about a Digital Specimen \(MIDS\)](#) level of content. The current focus is on assessing the level of completeness of data shared by institutions, with clear guidelines of what level of information needs to be provided to achieve the next level of MIDS.

## 2022 Work items

- Implement improvements to data validation reports identified in community consultation
- Promote integration of the data validator in tools connected to GBIF, including the IPT, through use of the validator API

## 2021 Participant contributions

- **Argentina:** Training workshops in different areas of the country for data providers. Permanent help desk in the publication and data quality. Participation on a CESP project of Data Quality.
- **Brazil:** This activity made no progress in this year.
- **Finland:** In addition to the process to assign each dataset with a quality label, FinBIF has introduced an annotation system in which users can annotate observations in Species.fi portal. All registered users of Species.fi can make annotations. In addition, experts have more rights to make annotations that affect the visibility of observations. Read more of the [quality](#).
- **France:** GBIF France help to run data quality tests before publishing datasets through GBIF. We organized data mobilization workshops including data quality and promotion of the data validator.
- **Korea, Republic of:** Data gap discovery will be carried out, and constantly updating and correcting the KBIF data.
- **Netherlands:** NLBIF has supervised a MSc student who analysed GBIF download records in order to assess the use and importance of predicates. One of the key findings was that the presence of geographical coordinates without geospatial issues (flag) is among the most.
- **Spain:** plans: Continuing with the Darwin-Test checking of datasets and training workshops as the normal operations of the Spanish Node. Progress: According to plan.
- **Sweden:** 2021: Assuring and maintaining data quality is still primarily considered by GBIF-Sweden to be the responsibility of the providers, feedback in regard to issues found by users being forwarded manually to contacts at host institutions. We have not in our planning expected to take any measures to automatize such information loops but look forward to the integration of the data validator in tools connected to GBIF, including the IPT, using the validator API.
- **Uganda:** Data providers are urged to ensure the fitness for use of their datasets

## 2022 Participant plans

- **Andorra:** None. We must first take stock of our data.
- **Argentina:** Training workshops in different areas of the country for data providers. Permanent help desk in the publication and data quality. Participation on a CESP project of Data Quality.
- **Brazil:** Tutorials on the data quality tools available in the Living Atlas infrastructure will be made available and disseminated.
- **Finland:** FinBIF will improve automatic validations of the data.
- **France:** GBIF France will continue to run data quality tests before publishing datasets through GBIF. We can test the integration of the data validator in tools connected to GBIF if needed.
- **Netherlands:** Further implement the use of "Issues and Flags" to improve the data quality.
- **Spain:** Same as 2021: this is an ongoing core operation performed continuously.
- **Sweden:** 2022: GBIF-Sweden plans to contact and inform existing data providers about the progress regarding the GBIF data validator tool but primarily for the time being, such matters will continue to be dealt with via the support system developed within the SBDI.
- **Taxonomic Databases Working Group:** TDWG members are interested in data quality (as in 'adherence to standards' or not). I'm particularly interested in how we can implement the BDQ tests and assertions "locally" to improve and standardize data before it gets to the aggregator. We plan to do this (as much as is feasible) in TaxonWorks. It would be good to continue to pursue how to show those who generate these data, what it 'looks like' in their world. (e.g. "What do [paleo, your group here] folks put in this field?")
- **Uganda:** Support the need to transform mobilized primary data into information and knowledge products.

## Activity 4c: Enable data curation

### Rationale (for the period 2017-2022)

In a global network, curation of the shared data pool is increasingly becoming a joint responsibility of aggregators, publishers, experts and data users. The goal is to integrate corrections, improvements, additional information and analysis results in a timely manner, with better visibility to all network participants and data users. Expanding the existing knowledge base requires improved communication channels and workflows for collaboration between all actors, tools to capture and rapidly display new or improved information, commentary and data, and not least tools, credit systems and support to engage expert activities.

### 2021 Progress

Feedback received during the 2021 consultation on digital extended specimens through the *alliance for biodiversity knowledge* (see Activity 1f) supported the need for GBIF to accommodate a more expressive data model for specimens, and to consider how communities can assist in annotating, linking and cleaning data (see also Activity 2a). This will help to inform GBIF's future work on enabling data curation.

Further work on the data validator tool in 2021 (see Activity 4b) will flag issues with stable persistent resolvable identifiers, supporting data curation.

The **data clustering algorithm** developed in 2020 has gone through iterative improvements, and by July 2021 was detecting candidate links for 23 million records, including 12 million specimens. A hackathon is being scheduled for Q3 2021 within the European BiCIKL project that will invite additional developers to contribute improvements to this algorithm. The Secretariat has presented the outputs at various conferences, and there is interest in brokering these annotations back to source systems.

## 2022 Work items

- Explore what can be enabled to offer annotation services in GBIF.org following ideas offered during the **consultation topic** during 2021 (see Activities 1f and 2a).
- Continue to explore the use of the GBIF data index to support stable persistent resolvable identifiers for all specimens and occurrence records (see Activity 2a)

## 2021 Participant contributions

- **Argentina:** Develop strategy and support mechanisms for expert communities to curate sections of GBIF data
- **Brazil:** A curatorship work has been implemented and is ongoing, whenever suspicious records, geographic outliers are identified, publishers are informed at the time of publication.
- **Chinese Taipei:** TaiBIF has developed a web system for managing monitoring and research projects using camera traps and managing the images, video and annotated data, and is developing a desktop program for data annotation and uploading for the use in field stations without stable Internet connection.
- **Finland:** FinBIF will continue to use and develop the annotation system. FinBIF also has active daily-basis cooperation with collection curators and is providing tools for them for their work. FinBIF participates in DiSSCo and **ELViS** which enhances the possibilities to curate the collections.
- **iDigBio:** iDigBio regularly participates in the collaborative efforts of GBIF, BCoN, and DiSSCo, among others, to forge a path for implementation of Extended Specimens.
- **International Centre for Integrated Mountain Development:** Explore possibilities with China Node (IB) a regional training on enhancing data quality of herbarium specimen This was initially thought as in-person training, so we did not push this forward this year. However, this can be embedded in the capacity strengthening programme we hope to formalize in 2022 after inputs from our data mobilization need survey.
- **Spain:** Continuing 2020 activities in this area.
- **Sweden:** 2021: GBIF-Sweden participated in the consultation on digital extended specimens but understand that our providers have in general less advanced knowledge of its contents and expectations/potential.
- **Uganda:** The National Node has been guiding the data holding institutions to co-opt experts in the processes leading to a publication. Particularly the IT, Biologists and Curators.

## 2020 Participant plans

- **Argentina:** Develop strategy and support mechanisms for expert communities to curate sections of GBIF data.
- **Brazil:** TBD
- **Chinese Taipei:** TaiBIF plans to further develop the camera-trap data management system to better support metric calculation and data visualization.
- **iDigBio:** iDigBio will continue to streamline its data mobilization workflows and will continue to work with GBIF on the global collections registry. In addition, iDigBio will continue to work with the international community on the Extended Specimen implementation.
- **Mauritania:** Creation of the Gbif Mauritania webpage.
- **Spain:** Same as 2021
- **Sweden:** 2022: We intend to take up a discussion with our providers to inform both parties about the awareness of the requirements and potential of common data curation processes. In earnest, the concept 'digital extended specimens' is not well known neither among collections curators nor with managers of observational or research datasets.
- **Taxonomic Databases Working Group:** From a TDWG, and personal perspective, I note just how hard it is for data providers to update their local datasets (for all the usual reasons). So, we need a) to see future grants that include "integration" of feedback included in the grant costs, and b) software and tools that simplify this process. On a recent grant (just now finishing up), I put in that we would meet with the data providers at the start to explain what we were doing with their data, keep the data providers in-the-loop (i.e. transparency) for the entire grant and meet with them at the end to help them move forward to integrate what ever information they find useful (that we added or enhanced). We will have a paper coming out that discusses some of what we learned in trying to do this that may help enable future efforts to close this currently "one-way" path where we really need loops or a network.
- **Uganda:** Efforts for teams that have a diversified expertise put in place.

## Priority 5: Deliver Relevant Data

### Activity 5a: Engage academia

#### Rationale (for the period 2017-2022)

The most significant user community for GBIF is academic researchers. Even policy-related uses of GBIF often derive from the work of such individuals. It is accordingly important for GBIF to understand the needs of researchers and academic societies and to communicate clearly regarding the tools and services GBIF can deliver. Communication should include information and support materials for students and early-career researchers, on both publication and use of data, including citation, use tracking and data papers. University faculties and libraries may be important channels for this information. In addition, GBIF needs to engage more closely with taxonomic societies and other academic bodies which could be key collaborators in curating and improving data. Achieving such an outcome depends on understanding how GBIF can become a more central tool for their

work, so that work on digital knowledge directly benefits those who contribute

## 2021 Progress

The following progress can be reported on the work items included in the 2021 Work Programme:

- The Secretariat developed new training resources in 2021 to support academic uses of GBIF-mediated data, trialled in a virtual workshop for a graduate programme in the United Kingdom coordinated by the **Natural Environment Research Council**. These resources focus on guidance for finding and downloading data through GBIF.org, data cleaning and best practices for citation. Once consolidated into a standard training package, this material will be incorporated into the GBIF training curriculum for use in GBIF-coordinated workshops and self-instruction (see Activity 1b).
- The lack of face to face meetings has limited progress in growing and deploying the corps of Biodiversity Open Data Ambassadors. However, the Secretariat has improved the means of tracking events at which ambassadors may be deployed, and continues to work on the resources available to ambassadors to communicate and promote GBIF.
- GBIF continues its engagement with the BioDATA programme, funded by the Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education (DIKU), focused on developing biodiversity data management skills for students in Eastern Europe, South Caucasus and Central Asia. During 2021, the Secretariat worked with GBIF Norway to secure follow-up funding for a four-year 'BioDATA Advanced' project which aims to accelerate biodiversity research through DNA barcodes, collection and observation data with academic partners in Norway, the Russian Federation and South Africa.
- Work on engaging professional societies to advance knowledge of GBIF functionalities in user communities did not progress significantly in 2021. Components of this work are carried over into the 2022 Work Programme.

## 2022 Work items

- Consolidate training resources aimed at academic users and publishers of GBIF-mediated data, for use in funded programmes and guidance for nodes
- Support implementation of the BioDATA and BioDATA Advanced projects through partners in Norway, Ukraine, Armenia, the Russian Federation and South Africa
- Promote GBIF at relevant fora including the 9th International Barcode of Life Conference, scheduled for August 2022, and the International Congress of Entomology, scheduled for July 2022
- Begin pilot project to integrate phylogeny with occurrence data (€20K)

## 2021 Participant contributions

- **Argentina:** Keep continue attending scientific meetings to continue promoting the national biological data system and GBIF.
- **Brazil:** Analysis and publication tools were disseminated to the academic community
- **Chinese Taipei:** TaiBIF has been helping researchers publish data papers and holding mini-

symposium on ecological informatics. TaiBIF regularly organizes promotion activities in academic conferences.

- **Finland:** FinBIF is part of an academic institution, University of Helsinki. FinBIF has been providing instructions and trainings for students and researchers such as courses in R. FinBIF also maintains a species learning environment Pinkka for students. FinBIF also has cooperation and projects with research projects (such as LIFEPLAN) and is providing tools for the researchers for collecting observation data in various projects (Glow worm project, Fungi atlas).
- **France:** Contributions to modules in Biology and Systematics master and doctoral modules at MNHN and Sorbonne University.
- **iDigBio:** The Florida Museum of Natural History in collaboration with iDigBio and the Natural Science Collections Alliance held the fifth annual Digital Data in Biodiversity Research Conference, virtually 7-9 June 2021. This was the second year the conference was held virtually. The virtual format again proved successful with 439 registrants from 40 countries.
- **International Centre for Integrated Mountain Development:** Conduct a survey with academia of the HKH region in relation to the capacities and challenges for data holders for biodiversity data mobilization. Biodiversity data mobilization capacity survey has been designed, and currently being drafted in lime survey. We will be sending the survey to the universities in the HKH region that are under HUC network, in September. Results will be synthesized by end of November 2021. The survey result is expected to guide our action for 2022.
- **Korea, Republic of:** Commercial companies collaborating with National Science Museum of Korea, such as NAVER, KAKAO, and WOONGJIN THINKBIG provide an e-dictionary to general public and children using the data that KBIF have mobilized.
- **Mauritania:** Master programme on the regeneration of biodiversity.
- **Netherlands:** NLBIF node manager has provided several lectures that include the importance of GBIF mediated data for scientific purposes.
- **South Africa:** SANBI-GBIF has developed an MoU with the University of Sol Plaatje for Partnership and collaboration with regards to training and capacity development in biodiversity informatics.
- **Spain:** Planned: Engaging with professional societies: specially with the Spanish-Portuguese Botanical Collections Association, and with the Spanish LTER network. Progress: According to plan.
- **Sweden:** 2021: Being funded by the Swedish Research council (SRC) as a national research infrastructure (with 50% co-funding by the partners in SBDI) GBIF-Sweden's predominant users are by definition expected to be found within academia. It is mandatory for us to produce key performance indicators regarding the use of the infrastructure in terms of academic users alone. altogether disregarding use by any other group (this in fact being mandatory!). Against this we struggle to convince both the SRC and particularly the Swedish Environmental Protection Agency (SEPA) whose influence on environmental management practitioners otherwise advised to use indigenous tools for data extraction and analysis is inevitably high, of the advantage of instead using and collaborating in developing SBDI and GBIF tools. We obviously have strong connections to academia, yet we still need to find more opportunities to contribute to basic and advanced education, training at least the next generation of researchers and management officials in the use of SBDI and GBIF. As described under the first major heading of the annual work plan update, GBIF-Sweden, and by the SBDI collaboration also our colleagues, participate

in, or run a set of courses and workshops in biodiversity informatics in Sweden. Outside of our sphere, interest is unfortunately alarmingly low!

- **Uganda:** Engagement with the academia is still a key priority for the Uganda node. The UNCST is key in the development and sustenance of the national biodiversity data bank housed at Makerere University. Plus more collaboration in the many faculty scholars at the various universities in Uganda.

## 2022 Participant plans

- **Andorra:** Promote GBIF at local level but there are mostly online universities in Andorra, which makes the task quite difficult (the assistants are not in Andorra, sometimes nor in neighbouring countries).
- **Argentina:** Keep continue attending scientific meetings to continue promoting the national biological data system and GBIF.
- **Brazil:** The dissemination of analysis and publication tools will be continued.
- **France:** Participations to master and doctoral modules to be continued. Possibility to translate some material if needed.
- **iDigBio:** iDigBio is currently planning the 2022 Digital Data conference. iDigBio will continue to publish newsletters, including Biodiversity Spotlight and Research Spotlight articles.
- **Mauritania:** Research on animal and plant biodiversity.
- **Netherlands:** Continue to highlight the importance of GBIF mediated biodiversity data and promote their further use.
- **South Africa:** Will continue efforts with academia.
- **Spain:** Same as 2021: This is a core ongoing operation that is performed continuously.
- **Sweden:** 2022: In 2022 we will continue our efforts to convince all potential present and future users of our services about the advantages and potential of advancing biodiversity informatics to find, attend to and solve problems within their different areas of interest. We seek to participate in conferences and other similar arrangements as frequently as possible, marketing ourselves (SBDI, GBIF) and related organizations.
- **Switzerland:** Coordination with Swiss Research Infrastructures.
- **Taxonomic Databases Working Group:** By its very nature, TDWG engages academia and will continue to do so. Certainly we also have a focus on enabling those in the southern hemisphere and other under-served parts of the planet.
- **Uganda:** Continued collaboration with academia especially in the valorisation of Uganda's genetic resources.

## Activity 5b: Document needs

### Rationale (for the period 2017-2022)

GBIF-mediated data are aggregated from many sources and are consequently heterogeneous, varying in fitness for various uses. During 2015–2016, GBIF established three task groups on data

fitness-for-use, in agrobiodiversity research, in distribution modelling and in research on invasive alien species, to document how these communities use GBIF data and to understand their data quality demands. The resulting reports inform data mobilization, data processing and improvements to GBIF.org. Further expert groups are considered during the current Strategic Plan period. Depending on resources, these groups will operate through a combination of face-to-face meetings and online or remote collaboration.

## 2021 Progress

The **GBIF task group on mobilization and use of biodiversity data for research and policy on human diseases** was formally established in November 2020, and during 2021 expanded its membership to include experts from Africa and Asia as well as Europe, North America and South America. Among the early activities of the group was to approve commission of a **systematic review of literature using GBIF-mediated data relating to human health**. Following an open call, Dr Francisca Astorga of Universidad de Chile was selected to carry out this review, which will help to inform GBIF's future work in meeting the biodiversity data needs of the biomedical research community.

Additional work requested by the task group is an initiative to promote mobilization of biodiversity datasets relevant to human health, through support for data papers to be published in a special journal issue following a call later in 2021. This will be enabled through a financial contribution from the **Tropical Disease Research centre (TDR)** of the World Health Organization, with GBIF providing funded help desk support to ensure that datasets associated with the data papers are correctly formatted and published through GBIF.org (see also Activity 3e).

With the completion and improved bibliographical content of checklists available in GBIF.org through the **Global Register of Introduced and Invasive Species (GRIIS)**, the Secretariat continued discussions with the IUCN Invasive Species Specialist Group on how data on invasive alien species can be better organized and curated, including through an improved research database managed by La Trobe University in Australia. An updated research article describing GRIIS is in preparation, and the discussions will focus on how the collaboration can best support information and indicators addressing invasive species within the CBD post-2020 Global Biodiversity Framework (see Activity 5c).

With additional funding approved by the Executive Committee for the 2021 Work Programme, the Secretariat is scoping activities for the second half of the year to engage experts to promote mobilization of soil biodiversity data, recognized by the Food and Agriculture Organization (FAO) as a significant information gap to support links between biodiversity and food security. This work involves collaboration with the **Global Soil Biodiversity Initiative (GSBI)** and the **Soil Biodiversity Observation Network (SoilBON)** - see also Activity 3b.

## 2022 Work items

- Coordinate work of the expert group on mobilization and use of data on zoonotic diseases based on outputs from the systematic review commissioned in 2021
- Continue programme for mobilization of soil biodiversity data through partnerships and work streams initiated in 2021
- Continue partnership with the IUCN Invasive Species Specialist Group and the Convention on Biological Diversity Secretariat to improve content and curation of data relating to invasive alien species

## 2021 Participant contributions

- **Brazil:** We are monitoring the work of a network of researchers who are collecting viruses in vertebrates, providing support for the DwC standard and subsequent publication in SiBBr and GBIF.
- **Finland:** Compiled user case dossier for DiSSCo. Participation in The Synthesys Virtual Access development and piloting for proper documentation (Digitation on demand)
- **iDigBio:** iDigBio hosts a large number of workshops, webinars, symposia and other event on digitization and data related topics, many of which are community-requested topics.
- **Spain:** Exploring ways on how to engage in these activities in a productive way. First dataset from GBIF Spain on soil diversity published.
- **Sweden:** 2021: Beyond the mentioned focus areas (diseases/human health – introduced and invasive species – soil biodiversity) GBIF-Sweden is aware of several existing and potential needs where our data is useful. We try to approach them all by raising interest in our offerings but have not coordinated nor participated specifically in work towards any of the three classes mentioned. If resources (staff time) were at hand, and provided that it be considered to be within the scope of our task as seen by other organizations in Sweden, we would probably be best suited to continue to explore the opportunity to mobilize further research and monitoring data on under-represented organisms, and from areas in Sweden (including marine environments and for that matter soil) where we know information is still scarce.
- **Uganda:** The node has been involved in the development of the capacity for National Reporting on the 3 UN RIO conventions for the national focal points.

## 2022 Participant plans

- **Brazil:** Continuing monitoring of a network of researchers working with viruses, as well as documentation of the publication of viruses x hosts datasets.
- **iDigBio:** iDigBio will continue to be responsive to community needs by offering training and workforce development activities via workshops, webinars, symposia, and other events.
- **Mauritania:** We need documents on the value of biodiversity.
- **Netherlands:** The NLBIF node manager as Open Biodiversity Data Ambassador will document user needs for GBIF data wherever relevant.
- **Spain:** To work with this recently contacted community to expand availability of soil biodiversity data in GBIF.
- **Sweden:** 2022: The above (see 2021) is a long-term undertaking, and we will as possible try to commit ourselves to any such activity but have not yet planned for it in our work plan.
- **Taxonomic Databases Working Group:** TDWG members participated actively in the CETAF-DiSSCo COVID19 Task Force as well as by developing the ViralMuse Task Force [https://www.idigbio.org/wiki/index.php/ViralMuse\\_Task\\_Force](https://www.idigbio.org/wiki/index.php/ViralMuse_Task_Force). Members of ViralMuse would be happy to contribute our expertise to the GBIF task group on mobilization and use of biodiversity data for research and policy on human diseases. Note that Pam Soltis (UF, iDigBio) and Deborah Paul (TDWG, INHS) have an NSF grant to work on exactly what this GBIF group is working on. We'd be pleased to collaborate on activities if deemed appropriate.

- **Uganda:** More continued support and participation in the documentation and mentor-ship of data provider documentation and dissemination efforts.

## Activity 5c: Support biodiversity assessment

### Rationale (for the period 2017-2022)

One of GBIF's key roles is as organizer of global evidence for species distribution, based on point records for species in time and space. Expansion of data publishing to accommodate sampling event data enables this evidence base also to mobilize and organize basic data on species populations and abundance. As a result, GBIF is positioned to serve as a critical resource for supporting biodiversity assessment at all scales. In particular, GBIF should serve as the data foundation for GEO BON to deliver Essential Biodiversity Variables (EBVs) for species distribution and population abundance. These EBVs represent a continuum from modelled representation of species occupancy (presence-only) in defined units of space and time through to richer assessment of species abundance in those units. GEO BON should serve as a forum for addressing the challenges of modelling such variables and interpolating sensibly between existing data points. This includes determination of appropriate scale at which modelled variables are adequately supported by current data. GBIF needs to ensure that it delivers the data foundations required for these activities, thereby supporting the requirements of IPBES, species Red Listing through IUCN and national authorities, the CBD and the Aichi Targets and Sustainable Development Goals.

### 2021 Progress

Following the contract awarded in 2020 to the VertNet consortium in response to the [call for proposals to analyse the biodiversity data needs in the post-2020 Global Biodiversity Framework](#), the Secretariat worked with the authors to finalize the document which will be published later in 2021. The analysis presents a number of conclusions and recommendations for more consistent and transparent use of species occurrence data in indicators monitoring the goals and targets of the framework. Plans are being developed for a workshop to be held early in 2022 bringing together key partners to help implement the recommendations.

The Secretariat has engaged actively in the delayed preparations for the COP15 UN Biodiversity Conference, through virtual sessions of the Convention's subsidiary bodies in May and June, as well as the final meeting of the Open Ended Working Group in August/September 2021 addressing recommendations for the detail of the framework. Through collaboration with GEO BON and a number of national governments, these meetings have provided an opportunity to highlight the importance of primary data mobilization in successful implementation of the framework, and this is reflected in wording of several of the draft recommendations arising from these meetings.

In April ahead of the subsidiary body meetings, the Secretariat organized and chaired a [virtual round table on biodiversity monitoring and data](#) as part of the [Fifth Science-Policy Forum for Biodiversity and Eighth International Conference on Sustainability Science](#). Including speakers from GEO BON, the Chinese Academy of Sciences, UNEP-WCMC, NatureServe, iDiv and the Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, the session provided an opportunity to reach a broad science/policy audience regarding the need for increased investment in biodiversity data and monitoring. The key messages from the round table were included in an [information document](#) provided for the virtual SBSTTA24 meeting.

Work continued in 2021 to implement several of the collaborations included in the [Memorandum of Cooperation between GBIF and IUCN](#). This included a survey of experts involved in the IUCN Species Survival Commission to understand better the way that GBIF-mediated data is used in the Red Listing process, and issues identified by IUCN users. The survey will help to inform planned engagement activities including webinars to improve understanding of GBIF's services and functionality, as well as emphasizing best practices for data citation.

The Secretariat implemented a new connection with the IUCN Red List web services to ensure that the latest version of the extinction risk category is available on the species pages of GBIF.org. Development has begun to implement filtering of occurrence data by Red List extinction risk category, but issues related to the handling of synonyms are being addressed before this functionality can go live on GBIF.org.

The Secretariat continued to work with the IPBES Task Force on Knowledge and Data across a number of areas to improve the use and citation of data in assessments, as well as prioritizing data gaps (see also Activity 3a). This included contributing to guidance videos provided by the Task Force to assessment experts in the application of the [IPBES Data Management Policy](#), and contributing suggestions for the handling of data originating from indigenous and local knowledge.

## 2022 Work items

- Through the *alliance for biodiversity knowledge*, hold a workshop on the outcomes of the analysis of primary data needs for the post-2020 Global Biodiversity Framework, involving partners in GEO BON, the Biodiversity Indicators Partnership, CBD Secretariat and others
- Implement a process enabling key checklists to be used in filtering occurrence data, such as Red Listed species and invasive alien species (carried over from 2020 and likely to be incomplete in 2021)
- Work with the new secretariat of GEO BON to improve connections between Biodiversity Observation Networks (BONs) and nodes at national, regional, thematic and global levels
- Continue collaboration with IUCN to implement the Memorandum of Cooperation, including exploration of publication of reference datasets associated with Red List assessments, and development of training material for use of data in the Key Biodiversity Area (KBA) process
- Continue to engage with the IPBES Data and Knowledge Task Force to support use of GBIF-mediated data in ongoing and upcoming assessments

## 2021 Participant contributions

- **Brazil:** Monitoring the tools available in GBIF for the assessment of biodiversity has been carried out, in addition to the dissemination of the tools available in the Living Atlas infrastructure, and also the availability of various thematic lists on SBBR, such as threat lists, invasive lists, sociobiodiversity list.
- **Chinese Taipei:** Biodiversity data has been used to calculate biodiversity indicators through the Taiwan Biodiversity Observation Network (TaiBON) project and some indicators has been used for writing Taiwan's first national biodiversity report.
- **Finland:** FinBIF has developed tools for the Finnish assessment of threatened species. The Regional Threat Assessment 2020 was published in the [Web Service of the Red List of Finnish](#)

## Species.

- **iDigBio:** iDigBio participated in WeDigBio, Worldwide Engagement for Digitizing Biocollections, a twice-annual event that brings together biodiversity collections, educators and researchers using collections data, and community scientists around the world to mobilize collections data via online platforms. The 6th annual event took place entirely online. During the four days of WeDigBio 2020, over 83,000 digitization tasks were completed, making this the most productive October event yet! The only event that has been more productive was WeDigBio Lite, the online-only event in April 2020 when 135,000 tasks were completed – a goal we look forward to breaking in the future.
- **International Centre for Integrated Mountain Development:** Review of fungi The Landscape Initiative for Far-eastern Himalaya (HILIFE) of ICIMOD is supporting a review work of fungi in far-eastern Himalayan Landscape, based on published literature. The dataset on the diversity of fungi will be published via HKHBIF; Likewise, checklist of flora and mammals is also expected.
- **Mauritania:** Awareness campaigns have been carried out for the protection of the biodiversity of continental wetlands.
- **South Africa:** SANBI-GBIF has supported biodiversity assessment efforts through work on the IPBES knowledge and data task force.
- **Spain:** Exploring ways on how to engage in these activities in a productive way. Progress: We are in contact with the Ministry for the Ecological Transition and regional governments to understand their data and science expertise needs regarding biodiversity, and finding ways to support them as well as to assist them in publishing the data they produce and gather in GBIF.
- **Sweden:** 2021: Biodiversity assessment/analysis for assessment is one major focus of SBDI, and although GBIF-Sweden has not been actively involved in the work on producing EBV:s we consistently strive to furnish data users with more and better data for environmental analysis, and with efficient tools to carry out assessments. We contribute by offering SBDI4R (web-enabled version of RStudio with ALA4R pre-installed) and Dockerized R-software bundle Mirroreum for use in the cloud or locally as desired.
- **Switzerland:** Publication of annotated national checklists for selected beetle families (Cantahridae, Lycidae).
- **Uganda:** The node is providing technical backstopping to number of biodiversity assessment efforts within the stakeholder ecosystem. Notably was the development of the national offset guidelines for Uganda.

## 2022 Participant plans

- **Brazil:** Continue monitoring GBIF activities and disseminate tools and services for biodiversity assessment at the national level.
- **Chinese Taipei:** TaiBIF plans to improve the data pipeline from the data repository to indicator calculation and trend interpretation.
- **iDigBio:** iDigBio plans to continue participation in WeDigBio events.
- **International Centre for Integrated Mountain Development:** In 2022, we will explore small grants to support assessment and review of key biodiversity resource.
- **Mauritania:** Field mission awareness campaign and training of populations on approaches to

conservation and restoration of ecosystems.

- **South Africa:** SANBI-GBIF will continue efforts to support assessments through IPBES task force of knowledge and data.
- **Spain:** Same as 2021.
- **Sweden:** 2022: In 2022 work on SBDI4R continues, and the recent update of Mirroreum is promoted. GBIF-Sweden has been represented in GEO BON since the early 2010's but due to resource shortage unable to commit to any work in practice. Yet we follow advances made with EBV:s, KBA:s, and with IUCN Red and Black listing.
- **Switzerland:** Continuation of data digitization activities in relation with running programs of the Federal Office of Environment.
- **Taxonomic Databases Working Group:** Please note from recent NSF grant, we developed a tool for use by the IUCN to visualize the georeferences against the shape files for the Horseshoe Bats and relatives. We successfully worked out a way to give them access to this information. For your purposes, we note they shared with us that most often, groups come to them asking for data, but rarely do groups come to them, offering them data. They were pleased to receive the contact and the data — for use by their IUCN group that manages these particular taxa. (Note of course we selected these taxa for their alleged link to Covid19).
- **Uganda:** To continue provision of the technical support to stakeholders.

## Activity 5d: Assess impact

### Rationale (for the years 2017-2022)

GBIF Participants require clear evidence of the benefits arising from investments in national- scale content mobilization and from GBIF global activity. At present, the main source of evidence presented derives from monitoring of published literature to identify uses of GBIF within research. This activity has been reported through annual GBIF Science Reviews and clearly demonstrates growing use of GBIF in research. Monitoring the literature in this way is time-consuming, and becomes more so as the relevant literature increases. A sustainable approach is required for future monitoring of this kind. GBIF now issues and promotes Digital Object Identifiers for data downloads and expects that these can be used both to simplify discovery and to improve the detail offered to Participants and data publishers on some uses of data. Participants also require more information on non-research uses of GBIF infrastructure, particularly in various kinds of government or industry assessments. A broader review of costs and benefits arising from GBIF investment would be valuable for Participants arguing continued engagement within GBIF and other countries considering Participation.

### 2021 Progress

The outcomes of the systematic review of research enabled by access to GBIF-mediated data, commissioned by the Secretariat in 2020 and carried out by Dr Mason Heberling were published in the [PNAS journal](#), providing a valuable reference on the range of science across many disciplines that is made possible by the GBIF collaboration and investment.

The total number of [research papers citing use of GBIF as a data source](#) passed the 6,000 milestone

in July 2021, and the [2020 Science Review](#), published in February 2021, summarized 67 articles reflecting the range of research and policy applications supported by GBIF. Based on the citation trends observed by July 2021, the number of research papers citing GBIF use was anticipated to exceed 1,200 during the whole year, compared with 987 in 2020.

To support broader demonstration of GBIF's impact, the Secretariat secured approval from the Executive Committee during 2021 to commission a robust quantitative analysis of the value and impact of the GBIF network. Terms of reference for this study will include examining GBIF's investment and use values, contingent value, efficiency impacts, innovation value and return on investment. Responding to a recommendation of the GBIF 20-year review, this work is expected to begin in the last quarter of 2021 and be completed by the middle of 2022.

## 2022 Work items

- Publish quantitative analysis of the value and impact of the GBIF network, and promote in engagement with governments, funders and other partners
- Launch Ebbe Nielsen Challenge and Young Researchers Awards competitions for 2022 (€25k), and generate communications connecting past awardees with subsequent work and impact
- Explore options for a 'Nodes Award Scheme' to recognize progress, performance and innovation across the community
- Produce communication materials including short videos demonstrating the impact of GBIF's funded programmes such as BID and BIFA
- Complete and promote communication materials demonstrating GBIF's impact across thematic communities, thereby supporting the case for increasing the number of Voting Participant countries

## 2021 Participant contributions

- **Brazil:** Activity did not progress in 2021.
- **Finland:** FinBIF has developed tools for [automatic tracking of citations](#). FinBIF has also launched Restricted Data Request Service for managing data requests for sensitive data or data that have other restrictions preventing open access. The data compiled will shed light on what is restricted-use data needed for (mining, wind farms, power lines, construction etc.) and will give unique data on impact of making restricted data available and accessible.
- **iDigBio:** iDigBio conducts an annual survey of its internal team as well as the collections community, broader scientific community, partners, stakeholders, and others interested in the national digitization effort. The survey is conducted to help track progress toward goals, identify community needs, solicit community feedback, and measure the impact of iDigBio's activities. iDigBio tracks data use in literature and publishes graphs on its website showing publications resulting from ADBC funding.
- **Korea, Republic of:** We have translated and published the 2020 GBIF Science Review to distribute these to researchers and decision-makers in related fields.
- **Mauritania:** Resilience of local communities.
- **Netherlands:** Naturalis Biodiversity Center together with NLBIF has embarked on a study on the physical and digital use of its collections. This has provided insights on number of downloads

from shared data and the importance of data aspects for data users. The most requested predicates included taxon key, the presence of coordinates and the absence of geospatial issues. These metrics, next to the actual use of data in scientific papers, are important KPI's that assess the impact of data sharing through GBIF.

- **South Africa:** SANBI-GBIF is developing a showcase to support the efforts relating to impact.
- **Spain:** Plans: Implementing the conclusion drawn from the activities mentioned for 2020 in the Nodes' plans. Progress: Data and figures from GBIF Science Reviews and GBIF Overview Slides are routinely used in our communications and reports. Ebbe Nielsen Challenge and Young Researchers Awards are widely promoted and supported. GBIF Spain dataset and publisher performance tracking system was dismantled in 2016 is being recovered. This was the basis for assessing performance and impact of institutions and projects involved in GBIF from Spain, and we aim to recover this capacity.
- **Sweden:** 2021: Using Matomo software, SBDI/GBIF-Sweden monitor publications using GBIF-mediated data, data downloads and SBDI services usage on behalf of the Swedish Research Council for annual reports KPA:s).
- **Uganda:** The data and information from and about Uganda in Uganda are increasingly utilized by the scientists and policy makers in Uganda for informed decision making. More response to calls have been reinforced by the available facts to support funding decision making.

## 2022 Participant plans

- **Belgium:** Reaching Integrated and Prompt Action in Response to Invasive Alien Species (RIPARIAS) LIFE project develops an innovative evidence-based workflow for decision making on IAS management, **translating national management objectives to concrete actions by maximizing cost efficiency of management, piloted in Dijle, Mark and Zenne river sub-basins of the Scheldt river basin district.** Belgian expert from the Platform to review IAS assessment for IUCN.
- **Brazil:** TBD
- **iDigBio:** iDigBio plans to conduct a community survey in early 2022 to assess the impact of its efforts and to gather community input. iDigBio will continue to track and publish statistics for its website, data portal, data use, and events.
- **Mauritania:** Conduct field work.
- **Netherlands:** Standardize the R-script to generate the download KPI's for general use by data publishers.
- **South Africa:** The efforts to reinforce the data-science-policy value chain through initiatives like the Foundational Biodiversity Information Programme of SANBI, will continue into 2022.
- **Spain:** To continue and conclude plans started in 2021.
- **Sweden:** 2022: This year represents a continuation of previous year's monitoring procedures, and to some extent an even more detailed account of research interest's infrastructure use. On the other hand use by other categories would be valuable and time and effort permitting we will set out to find out more about public management and private enterprise usage of SBDI.
- **Taxonomic Databases Working Group:** TDWG conducts yearly surveys post-conference, and otherwise as needed to track progress and note needs. We will continue to do so of course, and

the challenge of continuing to do all things virtually, adds to our need to listen, and improve how we do our work.

- **Uganda:** Increase data use in decision making and encourage from the Lab to the world concept of availing data and information.