



GBIF Work Programme 2020
Annual Update to Implementation Plan 2017–2021

Version 1, 2019-12-04

Table of Contents

Overview	1
GBIF Annual Work Programme 2020	2
Financial situation	2
Planned expenditure	3
Areas to expand activity	4
Secretariat staffing	5
Overview of 2020 Work Items	5
Priority 1: Empower Global Network	12
Activity 1a: Focus on people	12
Activity 1b: Strengthen skills	15
Activity 1c: Equip Participant nodes	21
Activity 1d: Equip data publishers	26
Activity 1e: Expand national participation	29
Activity 1f: Plan implementation	33
Activity 1g: Coordinate resources	37
Priority 2: Enhance Biodiversity Information Infrastructure	42
Activity 2a: Modernize data standards	42
Activity 2b: Deliver names infrastructure	46
Activity 2c: Catalogue collections	51
Priority 3: Fill Data Gaps	54
Activity 3a: Identify priority gaps	54
Activity 3b: Expand data streams	57
Activity 3c: Engage data holders	60
Activity 3d: Rescue datasets	65
Activity 3e: Liaise with journals	66
Priority 4: Improve Data Quality	68
Activity 4a: Ensure data persistence	68
Activity 4b: Assess data quality	69
Activity 4c: Enable data curation	73
Priority 5: Deliver Relevant Data	75
Activity 5a: Engage academia	75
Activity 5b: Document needs	78
Activity 5c: Support biodiversity assessment	80
Activity 5d: Assess impact	83
GBIF Implementation Plan 2017–2021	85
Delivery	85
Structure of GBIF	85
Timeline	87

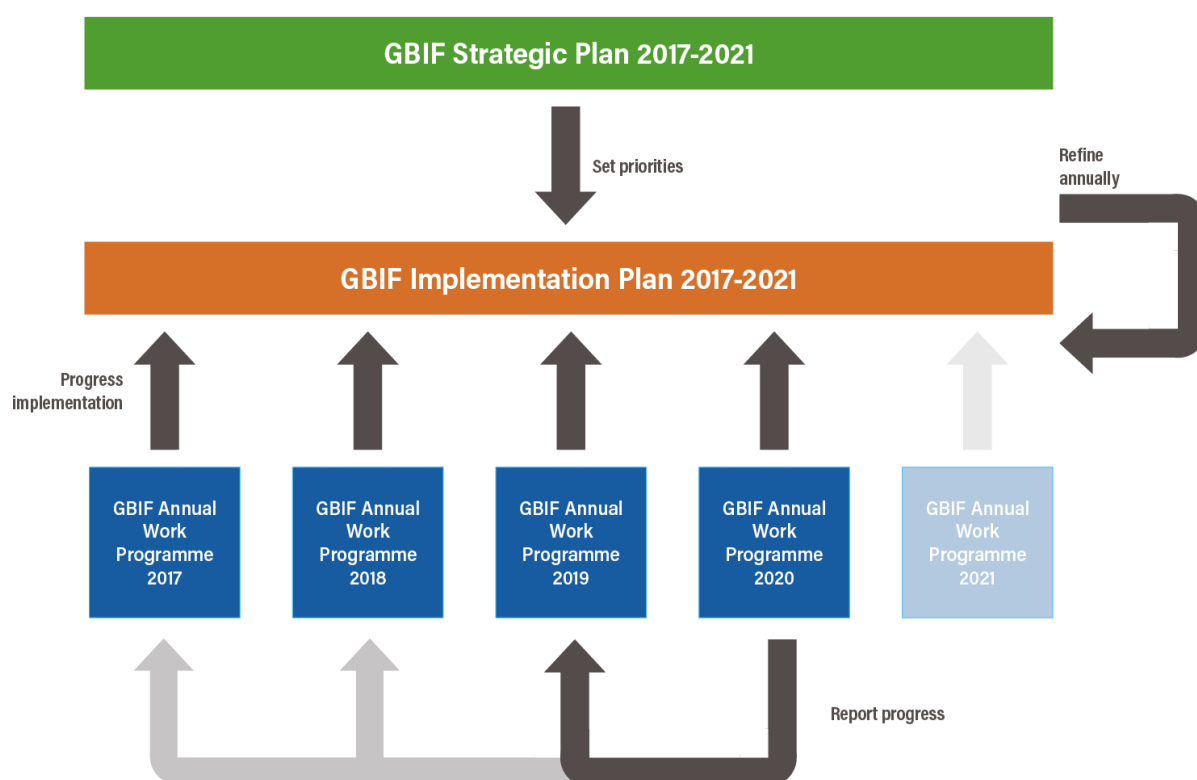
Resources	87
Implementation Plan Structure	88



Overview

This document serves both as the 2020 revision of the [GBIF Implementation Plan 2017–2021](https://www.gbif.org/strategic-plan) and also as the GBIF Annual Work Programme 2020. Both elements are based upon the [GBIF Strategic Plan 2017–2021](https://www.gbif.org/strategic-plan) [<https://www.gbif.org/strategic-plan>].

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



The [Strategic Plan](https://www.gbif.org/strategic-plan) [<https://www.gbif.org/strategic-plan>] (agreed in 2015 at GB22) established five priorities for GBIF's work.

The Implementation Plan presented here in this document identifies 22 activities which GBIF must undertake to address these priorities during the five-year period (and a set of specific tasks under each activity).

Each year, GBIF reviews 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 tasks for work during the year. Annual Work Programmes in the second and subsequent years also report on progress against the plan.

The Implementation Plan is reviewed and refined each year.

The Implementation Plan is presented here with updates reflecting progress during 2019 and planned activities for 2020 (the 2020 Annual Work Programme).

The Annual Work Programme 2020 is summarized below, with the revised [GBIF Implementation Plan 2017–2021](#) to follow.

GBIF Annual Work Programme 2020

Information on activities planned for 2020 is included under each of the Activity sections included in the [GBIF Implementation Plan 2017–2021](#).

For convenience, this information has also been collated as an [Overview of 2020 Work Items](#) of the items for which approval is sought from the Governing Board at GB26.

Financial situation

The Secretariat estimates a carry-over of €992,558 at the beginning of 2020. This amount is smaller than the carry-over from 2018 (€1,556,714) due primarily to Governing Board requested increase spending in 2019 on external contracts and workshops (€719,000) in the work programme. Given the level of capital held by GBIF and the uncertainty of contributions, the 2020 budget includes a smaller allocation (€413,000) for contracted activities in the work programme expenditure (i.e. non-salary costs allocated to calls, workshops, contracts, etc.) than proposed for 2019 at GB25, and €300,000 less than the actual 2019 spend after budget revisions. Staffing levels do not change in the 2020 budget. The planned expenditure for 2020 is €471,806 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 (€520,752 net capital forecast for the end of 2020).

This budget narrative also includes a second scenario that considers a brighter financial situation. It outlines priority areas of spending on contracted and workshop activities that would be submitted in subsequent 2020 budget revisions, should conditions allow.

Scenario 2 differs from the submitted budget in the following ways. It also assumes an additional €300,000 in external funds or increased contributions from voting participants. This budget includes €660,000 for contracted and workshop activities in the work programme expenditures. The planned expenditure under scenario 2 for 2020 is €397,940 greater than the income projected in the budget. This is also considered to represent a responsible balance between advancing GBIF's mission and reserving some funds against future risks (€594,618 net capital forecast for the end of 2020).

Scenario	Baseline (#1)	Optimistic (#2)
Income assumptions	Participants pay at a similar level	Modifications from Baseline Additional Voting Participants or Supplementary funds €320K/year
WP expenditure	€413,000	€660,000
2020 reserve	€520,752	€594,618

Planned expenditure

In addition to salary and operating costs (including support for GBIF Governance mechanisms), funding is reserved in the Work Programme 2020 for the following activities:

Baseline 2020 Work Programme budget for non-salary items:

Priority Area 1

- Following the recommendations of the community-based Documentation Editorial Panel, continue to create new documentation supporting participation in GBIF, based on late 2019 review and prioritization framework (€30,000).
- Develop joint training resources combining GBIF data mobilization and DNA barcoding, support for pilot workshop with Barcode of Life Data Systems (BOLD), Convention on Biological Diversity (CBD) and GBIF (€20,000).
- Consolidate online e-learning platform with training resources, to make the data mobilization, data use for decision making and nodes training modules fully functional as self-instruction modules, using external contracts for some components as necessary; develop a strategy to meet future demand for on-site, moderated and self-instructed training. Work will involve collaboration with key partners: nodes, the pool of trainers and mentors, Nodes Steering Group and numerous international partner projects (€20,000).
- Support for 2020 Regional Nodes Meetings (€60,000). The Secretariat will investigate possible synergies with BID phase 2 regional meetings, if funding is forthcoming, as well as with a BIFA workshop in Asia.
- Support further development of the Living Atlases community (€30,000).
- Support the alliance for biodiversity knowledge by working with steering committee on a governance structure and prioritizing the alliance efforts. Expand communication effort through conferences and the ambassador network (€10,000).
- Launch a new 2020 call for proposals under the Capacity Enhancement Support Programme (€80,000).

Priority Area 2

- Maintain and update processes for constructing the GBIF taxonomic backbone. €108,000 has been allocated in the budget to support GBIF contractor costs.

Priority Area 3

- Carry forward the proposed 2019 activity of data mobilization of vectors and hosts of human diseases to 2020. The Secretariat will establish an expert group (€25,000) to identify priority needs for biodiversity data supporting disease research, critical gaps in availability of such data in GBIF.org, and potential sources of data to fill these gaps.

Priority Area 4

- Explore bidirectional data linking and synchronization with data management systems and

publishers to achieve faster and more accurate mutual updates on data improvements and annotations (€10,000).

Priority Area 5

- Contract a study on use of GBIF-mediated data in the academic literature, with the goal of identifying areas of impact, links to research funding sources and networks of researchers. This information will be used to fine-tune GBIF value proposition message and align future Work Programmes to achieve greater impact (€20,000).

Other Funds

- The initial phase of the Biodiversity Information for Development (BID) programme will be completed by the end of 2019; however, the funder—the European Commission Directorate-General for International Cooperation and Development (EU DEVCO) has agreed to continue the programme with new funding of €1.6m from 2020–2023. This includes approximately €120,000 in funding for work by GBIF Secretariat staff in 2020.
- A fifth call for proposals under the Biodiversity Information Fund for Asia (BIFA) programme will be issued in late 2019 or early 2020, using €131,395 supplementary funds received from the Ministry of Environment, Government of Japan. Subject to agreement by the BIFA Steering Committee, we anticipate that €20,000 of this supplementary funding is to be allocated to Secretariat staff coordination costs during 2020.
- Governance funds: Launch 2020 calls for the Ebbe Nielsen Challenge and Young Researchers Awards (€44,000), considering recommendations from the Science Committee based on the 2019 programmes. The Science Committee is exploring potential external funds for these awards.
- Lastly, the European Commission has funded the SYNTHESYS+ project “to unify European natural science collections, effectively transforming them into an integrated data-driven pan-European research infrastructure”. GBIF is funded for €50,000 in 2020 (as well as €20,000 in 2021 and €20,000 in 2022) to advance joint planning for infrastructure components.

Areas to expand activity

If funding allows, the Secretariat will request that the following items be incorporated into the 2020 budget with approval of the GBIF Executive Committee.

Priority area 1

- Increase alliance for biodiversity knowledge activity to €20,000, Living Atlases to €40,000 and CESP to €100,000.
- Create an additional funding priority area alongside the CESP call for GBIF node regional outreach activities as outlined in the Regional Participation Strategy (€40,000).
- Continue development and implementation of a strategic plan for targeting external funding streams for capacity enhancement projects (€22,000).
- Initiate engagement, data mobilization and capacity enhancement campaign for Arabic speaking regions (€20,000).

Priority area 2

- Redesign the GBIF Integrated Publishing Toolkit (IPT) to support emerging data standards, explore integrations with quality control routines and to address infrastructural needs (ability to install locally, use a GBIF hosted solution or connect to a third-party repository). If funds allow €50,000 for an external contractor.

Priority area 4

- Increase bidirectional linking activity to €20,000.

Priority area 5

- Engage with academic based projects that use or could use better GBIF data in their pipelines and protocols. If funding allows hold a workshop to stimulate use (€25,000).
- Hackathon and workshop to align GBIF with post-2020 Biodiversity Framework (€40,000). The event would develop tools from CESP-CHM project to improve links to data from GBIF.

Secretariat staffing

Staffing levels will remain the same in 2020 with the increases in recent years have enabled the Secretariat to maintain an increasing level of service. We currently have 27 staff (three of whom work part-time) and are in the process of recruiting an additional informatics position. In addition to the Director and Deputy Director, the Secretariat is now structured as four teams, as follows:

- Participation and Engagement (seven staff)
- Data Products (four staff)
- Informatics (eight staff, one contractor, one additional staff position to be filled)
- Administration (six staff, plus student support)

All staff work towards accomplishing the goals of the five Priority Areas in the 2017–2021 Strategic Plan, with most staff working towards multiple areas.

Overview of 2020 Work Items

This following is a summary of all 2020 Work Items proposed in the 2020 update of the Implementation Plan below. It is intended to summarize the work items for which approval is sought from the Governing Board at GB26. Accordingly, it does not include the additional commitments made and reported by Participants.

Activity 1a: Focus on people

- Explore possibilities of providing citation tracking from dataset level to individual level. Provide ability to claim one's data and promote use; explore building on the example of 'Bloodhound tracker' and moving it into core GBIF operations.
- Clearly communicate and demonstrate the benefits of ORCID use through citation tracking.

- Pending a decision on continuing the GBIF Community Forum, take steps to improve interactive engagement with the broader GBIF community, considering the views gathered during the 2019 survey.

Activity 1b: Strengthen skills

- Develop joint training resources combining GBIF data mobilization and DNA barcoding, and support pilot workshop with BOLD, CBD and GBIF (€20,000).
- Explore opportunities for additional training workshops on data use for decision making.
- The BioDATA project will continue to train students in Armenia and in Ukraine in 2020, and plans for BioDATA II project are in place to further increase the geographic coverage in Northern and Central Asia and Southern Caucasus.
- Following the recommendations of the community-based Documentation Editorial Panel, continue to commission new documentation, based on late 2019 review and prioritization framework (€30,000).
- Consolidate online e-learning platform with training resources, to make the modules on data mobilization, data use for decision making and nodes training fully functional as self-instruction modules. Use external contracts for some components as necessary. Develop a strategy to meet future demand for on-site, moderated and self-instructed training. Work in collaboration with key partners, including GBIF nodes, the pool of volunteer trainers and mentors, Nodes Steering Group and numerous international partner projects (€20,000).

Activity 1c: Equip Participant nodes

- Support for 2020 Regional Nodes Meetings (€60,000). The Secretariat will investigate possible synergies with BID phase 2 regional meetings, as well as with a BIFA workshop in Asia.
- The Secretariat will continue enhancing guidance documentation for nodes. This guidance will incorporate priorities in data mobilization, gap filling, private-sector engagement and products from CESP projects.
- The Secretariat will design and implement user interfaces and services necessary to support simple hosted portals. The initial deployment will be of the BID programme portal followed by exploration of country portals and an exemplar virtual natural history collection. Discussions will determine whether such services should be restricted to Voting Participant countries and/or institutions that contribute a fee to create institutional portals.
- Continue conversations to align codebases and infrastructure with ALA, iDigBio, DiSSCo and other potential partners. The alliance for biodiversity knowledge will be used as a platform for some of these interactions.
- Support further development of the Living Atlases community (minimum €30,000).
- Consolidate Biodiversity Open Data Ambassador programme with issuing of digital badges, listing ambassadors on country pages and capturing feedback on promotion of GBIF at events.

Activity 1d: Equip data publishers

- Promote wider editing of the GBIF registry and the shared help desk activities; including node staff initiating and diagnosing dataset crawling / ingestion.

- Complete implementation (if not finished in 2019) and develop processes to allow open editing of shared vocabularies used in data interpretations of the GBIF ingestion pipelines (e.g. habitat types, occurrence status, etc...).
- Create system of list management, similar to bulk email, to communicate to a larger section of the GBIF community for compliance and notifications. This includes a twice-yearly mandatory communication with data publishers in compliance with General Data Protection Regulation (GDPR: EU privacy regulations) and explore services for publishers to opt-in to receive push notifications for new citations.
- Provide comprehensive guidance and support services to lower the technical threshold of data-hosting options. Clearly document the benefits and implications of each option including aspects of operational cost, deployment model (local/cloud/GBIF-provided) and expectation of users. Use of volunteer mentors will be promoted to enhance help desk services.

Activity 1e: Expand national participation

- The GBIF Secretariat will continue to identify funding opportunities for regional engagement, including inviting observers to regional node meetings. If budget allows, it will create a specific funding priority area alongside the 2020 CESP call for GBIF regional outreach activities, as outlined in regional participation strategies developed across the network.
- The Secretariat will develop and implement engagement, data-mobilization and capacity-enhancement plans for greater participation in all global regions, with an emphasis on the underrepresented areas. This effort will be aided by capacity enhancement workshops, including through a second phase of BID and other supplementary funding that becomes available.
- The internationalization of GBIF.org user interface, content and guidance materials will continue. Assuming the 2019 commitment for six UN languages is complete, this work will focus on maintenance and updates and will allow the addition of more languages on demand, and translating a wider set of content and materials, funded through both specific projects and volunteer efforts.
- Engagement, data mobilization and capacity enhancement campaign for Arabic-speaking region (€20,000, if funds allow).

Activity 1f: Plan implementation

- Convene SYNTHESYS+ workshops and integrate outcomes with relevant Work Programme activities.
- Continue implementation and refinement of the standardized annual information requests from Participant nodes and display up to date content on revised country pages of GBIF.org.
- Lead some and participate in all alliance for biodiversity knowledge virtual workshops that will be defined in key areas to increased alignment of GBIF with other networks and infrastructures.
- Manage and improve the virtual conferences infrastructure used by the alliance for biodiversity knowledge. Work with steering committee on a governance structure and prioritizing the alliance efforts. Expand communication effort through conferences and the ambassador network (minimum €10,000). Provide technical, communication and administrative support for the alliance.

Activity 1g: Coordinate resources

- Launch a 2020 call for proposals under a renewed CESP (minimum €80,000). If funding allows, this will be supplemented by a special category for regional outreach activities that target increased data mobilization in non-GBIF countries as recommended by Nodes Steering Group ((see Activity 1e).
- Based on a successful negotiation with EU DEVCO, the European Commission funder of BID, the Secretariat will continue with a BID call for proposals that responds to feedback received from community consultations, three regional closing meetings and the BID phase 1 closing meeting.
- Implement a fifth call for proposals (approximately €110,000) under the Biodiversity Information Fund for Asia (BIFA), reflecting priorities agreed by Asian nodes.
- Continue collaboration with the BioDATA (Norway > Eurasia) and Russia support (Finland > Russia) supplementary funding programs.
- Develop and implement strategic plan for targeting external funding streams for capacity enhancement projects. This will be based on 2019 external consultation and using the results of BID Phase 1 closing meetings. If budget allows an additional (€22,000) will be allocated to this effort. The strategy will work with all GBIF regions to target potential funding streams that support additional capacity enhancement for data mobilization and use, building on the BID and BIFA models.
- Make refinements to the newly selected grant management system, FLUXX, to streamline project calls, assessment, selection, implementation, budgeting and reporting, to ensure it meets the needs of the community.
- Launch 2020 calls for the Ebbe Nielsen Challenge and Young Researchers Awards (€44,000), considering recommendations from the Science Committee based on the 2019 programmes. Explore the possibility of soliciting additional donor/sponsor co-funding.
- Begin Secretariat planning to establish a workflow using digital documentation to develop the 2021 work programme and 2022–2026 strategic plan. This planning will address the recommendations of the 2019 20-year review.

Activity 2a: Modernize data standards

- Modernizing data standards is a continuous Work Programme activity for a global infrastructure like GBIF. During 2020 we will focus on advancing and refining data models for Collections, Taxonomic Treatments, Sampling Events, Organisms, Specimens, Organisms, Citations and the linkages between them.
- Provide a set of data-exchange profiles for sharing data within GBIF that conforms with a unified information model that includes both existing and new standards as well as the necessary controlled vocabularies.
- Redesign the GBIF Integrated Publishing Toolkit (IPT) to support these profiles and to address infrastructure needs, such as the ability to support local installations or GBIF-hosted solution. If funds allow, €50,000 for an external contractor.
- Provide documentation for the data model and for the associated services offered through GBIF.org.
- Review and redesign GBIF data management system to accommodate the unified information

model as part of data ingestion, quality control and processing where necessary.

- Continue technical discussions with other data aggregators to seek closer alignment in practice and, as far as possible, implementation of aggregation and indexing processes.
- Demonstrate improvements of information in GBIF.org and hosted national portals in specimen-level information, links to material citations, and links between specimens and sequence data from sources such as BOLD.
- Explore approaches for adding a phylogenetic/evolutionary dimension to the GBIF taxonomic backbone. Pilot phylogenetic browsing capabilities of occurrence data.
- Open discussion with GB participants to provide project funders with an overview of the resulting value relating to their investment (e.g. data mobilization, publications).
- In collaboration with international partners, explore the desirability and scope of “catalogue services” that are targeted specifically at physical specimen collections. Examples could include displaying duplicate or derived specimens across collections, type information, citations in taxonomic treatments and trait data.
- Explore options for displaying occurrence data from long-term sampling sites, piloting with projects like BIOSCAN 2 and/or Norwegian ecological datasets.

Activity 2b: Deliver names infrastructure

- Maintain and update processes for constructing the GBIF taxonomic backbone, 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 the Catalogue of Life.
- Implement a process enabling key checklists to be used in filtering occurrence data, such as Red Listed species and invasive alien species.
- Consult with relevant regulatory agencies, such as the European Environment Agency (EEA), for guidance on which legislative checklists should be incorporated to increase the relevance of COL+ to governments.
- Explore feasibility of supporting national taxonomies for exploring GBIF occurrence data to better enable national level reporting.
- Develop and pilot a process that allows qualified users to collaborate and edit sectors that contribute to the GBIF backbone taxonomy, aimed at reducing the delays before such edits appear on occurrence records from months to days.

Activity 2c: Catalogue collections

- Based on community consultation, build mechanism to synchronize Global Registry of Scientific Collections (GRSciColl) with other catalogues.
- Improve linkages between collections, institutions and occurrences (or specimen) objects indexed by GBIF.
- Develop the user interfaces and services necessary to support a collection catalogue system.
- Work with the community to ensure the content is fit for use, and promote community editing of the registered content.

Activity 3a: Identify priority gaps

- Continue work on items initiated in late 2019, with an emphasis on developing actionable guidance for data publishers and nodes, integration of user needs into prioritization for data mobilization, and data search analysis.
- Continue to improve visualizations of GBIF-mediated data that identify gaps by engaging in interactive community consultations. This work is a candidate for curated discussion through the alliance for biodiversity knowledge prior to implementation on GBIF.org.

Activity 3b: Expand data streams

- Enhance the data exchange standards for sampling-event data, collaborating with partners that generate data to provide sources for filling current gaps. This work aims to establish partnerships with long-term monitoring communities.
- Improve linkages between records originating from museums and BOLD in order to link information that is currently treated as two occurrences.
- Carrying over the proposed 2019 work item, mobilize data on vectors and hosts of human diseases. Establish an expert group (€25,000) to identify priority needs for biodiversity data supporting disease research, critical gaps in availability of such data in GBIF.org, and potential sources of data to fill these gaps. The campaign will use this analysis to engage directly with relevant data holders, support data publication through GBIF and inform data mobilization priorities for use by nodes, publishers and funders (see Activity 3a).
- Continue linking and integration of sequence-based data streams.

Activity 3c: Engage data holders

- Continue work with DiSSCo project team to maximize opportunities for mobilizing collections data from European institutions, including in countries not yet participating in GBIF.
- Work with iNaturalist, iDigBio and nodes community to maximize opportunities for public engagement in GBIF data mobilization.
- Develop private-sector data mobilization guidance and training. Promote revised guidance on mobilizing EIA data and run training programme for private sector consultants at IAIA conference in Spain 2020, and incorporate guidance from CESP project in nodes guidance package.

Activity 3d: Rescue datasets

- Continue to implement workflow for prioritizing and drawing upon potential data sources reported through the ‘dataset catcher’ tool, including involvement of nodes, mentors and crowdsourced solutions.
- Roll out a workflow for ‘Suggest a dataset’ processing.

Activity 3e: Liaise with journals

- Work with journals to make it easier for article authors to deposit supplementary data in formats suitable for GBIF publication.

- Establish model guidelines for data deposition and citation and develop compliance criteria for distribution to publishers. Build on recent developments of COPDESS and Research Data Alliance with GBIF-specific guidance. Develop additional communication materials to describe the benefits of DOI-citation best practices and work with journals on implementation.

Activity 4a: Ensure data persistence

- Continue revision and documentation of flagging routines used in GBIF data ingestion pipelines.

Activity 4b: Assess data quality

- Review, consolidate and update existing documentation for data publishers. In particular, provide clear guidance on minimum requirements for published data.
- Develop metrics to track the completeness of core data elements and the degree to which supplied content is appropriate.
- Supply clear indicator measures for the completeness and usability of data as part of GBIF.org dataset pages, based on examples such as the GEOLabel data branding model.
- Extend data-quality assessment to include aspects only detectable above the level of individual records.
- Assess the patchiness of indexed data (geographical clustering, misleading accuracy or precision of coordinates), including evaluation of the apparent causes of data patchiness and include measures of data patchiness in the data index, at both dataset and record level in the data index.
- Ensure that users of data are able to identify datasets or records that do not fulfil their criteria for geo-accuracy, whether they are accessing data through facets in the GBIF.org, via the API or in downloads.

Activity 4c: Enable data curation

- Continue to explore the use of the GBIF data index to support stable persistent resolvable identifiers for all specimens and occurrence records.
- Explore bidirectional data linking and synchronization with data management systems and publishers to achieve faster and more accurate mutual updates on data improvements and annotations (minimum €10,000).

Activity 5a: Engage academia

- Conduct survey on how GBIF informatics is incorporated into relevant graduate and undergraduate curricula to identify gaps and opportunities and key entry points. Use survey findings to prepare a campaign that could roll out relevant resources through GBIF nodes in 2021.
- Engage with academic-based projects that use or could use better GBIF data in their pipelines and protocols. If funding allows, hold a workshop to stimulate use (€25,000).

Activity 5b: Document needs

- Explore the creation of lightweight, customized website landing pages to address thematic interests on GBIF.org. The pages will increase flexibility to include curated, more informative clustered information, ensuring that users have access to both broad search results and prioritized views of data and information.
- Continue to implement recommendations of past expert user groups. In particular, build on the incorporation of GRIIS checklists to support richer, more targeted information on invasive and alien species, e.g. by highlighting documented occurrences with a relevant IAS status by country. Explore best options to identify and alert users of new occurrences of potential invasive species. In addition, review and address the most

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

Tasks

- i. Develop mechanisms to support and reflect the skills, expertise and experience of individual and organizational contributions to the GBIF network (including revision of identity management system and integration of ORCID identifiers)
- ii. Consolidate use of digital badges to record skills and contributions

2019 Progress

The Secretariat informatics team has begun investigating citation tracking of individual records, and developing the ability for users to ‘claim’ data they have contributed, to promote use and to demonstrate the value proposition of GBIF. User profile work is scheduled to begin in late 2019 and early discussions have been initiated with GBIF Canada around lessons learned in claiming specimens through [Bloodhound Tracker](https://bloodhound-tracker.net) [https://bloodhound-tracker.net] and how to integrate that tool into GBIF.

The Secretariat has been carrying out an evaluation of the [GBIF Community Forum](https://discourse.gbif.org) [https://discourse.gbif.org] during its first year of operation, to inform a decision on whether to continue maintaining the platform. Preliminary metrics suggest that use of the forum by the community is well below the targets set as benchmarks when the platform was established as a successor to the GBIF Community Site. 336 posts were made over a year with 87 new topics started by non-Secretariat staff. On the other hand, the time spent by Secretariat staff on moderation/curation was not significant.

2019 Participant contributions

- **Andorra:** The Andorran node of GBIF has scarce human resources to provide data to the GBIF. Nevertheless we have made an effort to increase the data amount by the assistance of our collaborator in Andorra. Also we are constantly in contact with the Spanish node team, the Real Jardín Botánico in Madrid, which gives us an appreciated technical assistance.
- **Benin:** One workshop was achieved in May 2019 to train national partners on data mobilization, data uses as well as registration on GBIF site to become more visible on GBIF site. At 5 least NGOs register on GBIF site this year. The master program in biodiversity informatics is doing great with in-depth capacity building of students from Benin and other African countries. It is just unfortunate that we never have the support of GBIF in that innovative and promising program.
- **Canada:** CBIF contributes to the focus on people via David Shorthouse's development of the Bloodhound application (<https://bloodhound-tracker.net/>) that enables agents to "Claim the natural history specimens you collected or identified, track their use in new science, and help acknowledge your peers, mentors, and organizations." We continue to support David's further development of this important tool and consider how to link this to development of community attribution standards and best practices.
- **Canadensys:** We have followed the discussions in GBIF Community forum, but we have not been part of the survey.
- **iDigBio:** In February 2019, iDigBio and the Office of Academic Support (OAS) and Student Support Services (SSS)/TRiO Program at the University of Florida co-organized a free Biology Career Conference and Fair for undergraduate students called iDigTRiO. TRiO is a federally funded grant program to help first generation students, low income students and students with disabilities navigate and succeed in their undergraduate careers and beyond. The event was targeted to SSS and TRiO students but was open to all undergraduates.
- **Japan:** Provided one workshop to show GBIF Japan can now accept fossils, and to enroll people in this area.
- **Mexico:** Internal process know how extended for people inside publisher CONABIO, know in another best practices on check list, train of trainers in checklist.
- **Netherlands:** NLBIF and Naturalis Biodiversity Center host the Biodiversity Next 2019 conference that will bring together a wide audience of biodiversity data, standards, and infrastructure stakeholders.
- **Spain:** Helped testing and participating in the GBIF Community Forum.
- **Sweden:** GBIF-Sweden has followed the development of GBIF's experiences of GDPR adaptation in relation to expanded user profiles and data publication.
- **United States:** US Node Manager joined GBIF Community Forum and contributes to relevant discussions.
- **Zimbabwe:** In January 2019, GBIF Zimbabwe set up a technical committee composed of people with relevant skills and knowledge required for the effectiveness of the Node. These include, ecologists, entomologists, microbiologists, bioinformaticians, wildlife ecologists, ichthyologists and herpetologists. Launched a website for GBIF Zimbabwe and created twitter account.

2020 Work items

- Explore possibilities of providing citation tracking from dataset level to individual level. Provide ability to claim one's data and promote use; explore building on the example of 'Bloodhound tracker' and moving it into core GBIF operations.
- Clearly communicate and demonstrate the benefits of ORCID use through citation tracking.
- Pending a decision on continuing the GBIF Community Forum, take steps to improve interactive engagement with the broader GBIF community, considering the views gathered during the 2019 survey.

2020 Participant plans

- **Andorra:** We plan to continue working with Spanish node team to continue improving our National Data Portal of Andorra and also increase the volume of data on biodiversity.
- **Benin:** Capacity building through workshops and in the framework of the master program in biodiversity informatics data mobilization data uses.
- **Canada:** CBIF will continue to collaborate with GBIF to incorporate Bloodhound functionality into the suite of GBIF services.
- **Canadensys:** We will continue to follow and be part of the discussions in the forum. We think that the platform and discussions are really useful.
- **iDigBio:** iDigBio will continue to support the series of workshops aimed at broadening participation in the biological sciences. iDigBio is planning a second annual iDigTRiO Biology Conference and Fair for February 2020.
- **Japan:** Provide one workshop to expand the GBIF-related biodiversity informatics community.
- **Naturalis Biodiversity Center:** Naturalis plans to implement augmented user profiles based on ORCID in ELViS, to be developed in SYNTHESYS+
- **Spain:** We will spread through our national communication channels the new and expanded user profile at gbif.org when it is ready.
- **Sweden:** GBIF-Sweden will upgrade data provider's agreements where necessary. Staying informed by and participating in discussions via the GBIF Community Forum (and GBIF webinars) will be encouraged within our regional community.
- **United States:** USGS will work with OBIS to provide an updated memorandum of understanding between OBIS and GBIF.
- **Zimbabwe:** Set up GBIF Zimbabwe community forum by June 2020

Rationale

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.

Approach

The key task is to enhance GBIF.org to connect and display information on its contributors and other users of the site. This may include contact details (subject to individual control), formal roles (e.g. within participant delegations; in connection to publishing/improving datasets; based on training experience) and online participation in help desk discussions. Whether through the use of existing social or commercial platforms or through custom web application development, GBIF.org should absorb activity currently handled through the separate GBIF community site. GBIF should also build on experience during 2015–2016 in using digital badges to identify skills held by individuals. Such community functions will support the operation of GBIF Task Groups, externally funded projects like BID, and engagement of expert groups to curate GBIF data.

Activity 1b: Strengthen skills

Tasks

- i. Develop collaborative help desk capability – including discussion groups, support materials and FAQs
- ii. Articulate training and capacity enhancement activities around a single, global curriculum for GBIF
- iii. Deliver explainer texts and videos to support self-directed instruction and learning

2019 Progress

The Secretariat continues to improve current and develop new training materials as well as standardized arrangements for setting up training workshops based on re-usable curricula in key skills related to GBIF participation. This work is accomplished in close collaboration with trainers and mentors from the GBIF community, who contribute their time and expertise at little or no cost to GBIF's core budget. The capacity enhancement programmes (e.g. BID, BIFA and CESP and GBIF involvement in other externally-led projects) are major motivating factors for the Secretariat to improve materials and support a growing community of practice.

Two ways to measure the value of this investment in generating training materials are, first, the increased number of volunteer mentors and trainers across the network, which now stands at 93 in total, with 58 involved in the BID programme. Second, these programmes have sparked a sharp increase in both the number of data publishers and the amount of data from countries and regions not previously engaged in biodiversity data sharing, with the BID programme adding 90 new publishers from sub-Saharan Africa, the Caribbean and the Pacific. Together, with organizations in these regions that previously shared data, the new publishers introduced through the BID

programme have added nearly 1.3 million occurrence records, filling critical knowledge gaps by increasing in the percentage of data for a country being published within that country and adding records for 1,194 species listed as threatened in the IUCN Red List.

Multiple training events involving the Secretariat and community mentors and trainers occurred or are planned in 2019, including the following:

- ‘Train the mentors’ event to support data mobilization training within the BioDATA project covering Eastern Europe, the Caucasus and Central Asia (Belarus, February)
- Training course on data skills within the BioDATA project (Tajikistan, June)
- Training sessions on establishing a GBIF Participant Node, including a focus on the benefits of GBIF participation, at the BID regional meetings in the Caribbean (Trinidad and Tobago, June), Pacific (Fiji, August) and Africa (Cameroon, September)
- Capacity enhancement workshop on data mobilization for projects selected under the Biodiversity Information Fund for Asia (BIFA), as well as for GBIF nodes in Asia (Viet Nam, July) & Training on development of GBIF Participant Nodes, offered to all node managers as part of the GB26 series of meetings (Netherlands, October)
- Capacity enhancement workshop on data use for decision making for participants in GIZ-funded project to improve biodiversity information management in southeast Europe (Bosnia and Herzegovina, November/December to be confirmed)
- The BioDATA project focused on developing skills in biodiversity data management and data publishing for undergraduate and postgraduate students from Tajikistan, Belarus, Ukraine, and Armenia. The project is run by the University of Oslo and GBIF Norway and exceeded expectations, with more than 8 mentors and 40 students trained on the GBIF publishing and data skills in two events that took place in Belarus and in Tajikistan.

As well as these training events directly involving the Secretariat, a large number of replicated workshops based on the same materials have been carried out in association with various funded projects as well as node activities. For example, BID participants have already replicated the data mobilization and data use workshops 66 times, extending the impact of the BID programme to nearly 1,500 new students within the corresponding regions.

The Secretariat has continued to update training materials for data mobilization, decision making, for use in on-site training and in development of online modules. For the data use curriculum, GBIFS has been working with IUCN to refine guidance on the use of species occurrence data in species Red List assessments. This is one of the areas of collaboration included in a Memorandum of Cooperation between IUCN and GBIF Secretariat, finalized and signed in July 2019.

The Secretariat has established a community-based Documentation Editorial Panel, aimed at prioritizing, developing and updating a core set of electronic guidance documents to support participation in the GBIF community. The panel includes membership from four global regions, and it will have held three meetings by August 2019. A consortium led by the VertNet team is producing an initial package of new and updated documentation in 2019, namely:

- Basic Guide to Using OpenRefine (in Spanish)
- Quick Guide to Georeferencing

- Best Practices in Georeferencing
- Georeferencing calculator
- Best Practices for Generalizing Sensitive Species Occurrence Data

2019 Participant contributions

- **Argentina:** Participation as mentor and trainer (national, regional and global), workshops with the Environment Secretariat for training.
- **Australia:** The Atlas co-coordinated the CESP Paris workshop (March 2019) with GBIF France and GBIF Spain which provided training for the GBIF node staff on spatial components of the Atlas including the Spatial Portal. The Atlas developed training materials to support this workshop that can be used for future events.
- **Belgium:** Completed set of explainer texts in multiple languages: French and Dutch translations.
- **Benin:** One workshop was achieved in May 2019 to train national partners on data mobilization, data uses as well as registration on GBIF site to become more visible on GBIF site. At 5 least NGOs register on GBIF site this year. The master program in biodiversity informatics is doing great in in-depth capacity building of students from Benin and other African countries.
- **Canadensys:** Canadensys, in partnership with iDigBio and GBIF Spain, and partially funded by the CESP program, has developed a series of workshops in Canada in order to strengthen the stakeholders knowledge about data mobilization, data cleaning and data-use. These workshops have been focused on the region, but the documentation developed can be easily adapted and reused outside North America. The documentation is available in both English and French.
- **Colombia:** Biodiversity Data Publication Training Workshop: This workshop provided information and training for publishing biodiversity data through SiB Colombia and GBIF online platforms. The workshop took place on November 30, 2019 at Corporación Autónoma Regional del Risaralda, Quindío. The objective of the workshop was to support GBIF-Colombia's partners (particularly environmental authorities and ONG) achieving greater visibility at the national and international level publishing biodiversity data. A total of 15 participants attended the workshop. Through the workshop aims, participants were informed on the recent activities of GBIF Colombia and the importance of biodiversity data publishing. Participants were invited to keep supporting GBIF-Colombia in its ongoing and future activities. <https://sibcolombia.net/taller-sirap/> Biodiversity Data Publication Training Workshop: This workshop provided information and training for publishing biodiversity data through SiB Colombia and GBIF online platforms. The workshop took place on April 24–26, 2019 at Universidad Industrial de Santander, Bucaramanga. The objective of the workshop was to support GBIF-Colombia's partners (particularly biological collections) achieving greater visibility at the national and international level publishing biodiversity data. A total of 23 participants attended the workshop. Through the workshop aims, participants were informed on the recent activities of GBIF Colombia and the importance of biodiversity data publishing. Participants were invited to keep supporting GBIF-Colombia in its ongoing and future activities. <https://sibcolombia.net/taller-uis-2019/>.
- **France:** GBIF France supplied trainers and mentors for workshops organized by GBIF during 2019 under BID and BIFA for capacity enhancement in data publishing and use. Togo and France : Implementation of a capacity enhancement project to increase the online presence of GBIF Togo. Designing a dedicated, dynamic online site, devoted to biodiversity and news from

Togo, based on ALA tools.

- **iDigBio:** iDigBio partnered with Environmental Data Initiative (EDI), Earth Science Information Partners (ESIP), DataONE, GBIF, NEON and Arctic Data Center on a Data Help Desk at the 2019 meeting of the Ecological Society of America. The goal was to engage participants one-on-one about their questions, comments and concerns about using these data in their discipline and area of research. iDigBio has developed a strong relationship with The Carpentries. iDigBio staff participate in Carpentries workshops throughout the year. iDigBio also has several staff members trained as Carpentries instructors.
- **Japan:** Provided two workshop for data carpentry and GBIF/S-Net website use.
- **Mexico:** Development website snib.mx for accessing and download occurrences data and national list of species with distribution in Mexico. Preparation tutorials for use geoportal.conabio.gob.mx it is a geographical portal with thematic maps and occurrences maps. In process, collaboration CESP2018_011 “Increasing capacities to develop National Species Checklists in the Latin America and the Caribbean Region” with the publish “Aves de México” and “Helmintos parásitos de vertebrados” national list of Mexican species.
- **Naturalis Biodiversity Center:** Naturalis contributed to training and capacity enhancement activities using GBIF training materials in the COST Action MOBILISE.
- **Netherlands:** NLBIF has motivated data providers to take more responsibility for their data and manage their data on the NLBIF IPT or start hosting their own IPT.
- **Norway:** BioDATA is a three-year project (2018–2021) managed by GBIF Norway in collaboration with the Norwegian Research School in Biosystematics (ForBio) and the GBIF Secretariat, that is funded by the Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education (DIKU). BioDATA build biodiversity data management skills across the former Soviet Union region and includes partners, students, and training events in Belarus, Tajikistan, Armenia, Ukraine and Norway (see also Activity 1e). BioDATA is reusing GBIF BID training materials and have completed the translation of the training curriculum to Russian together with the GBIF.ru team. A light-weight Raspberry Pi IPT server was developed for BioDATA training purposes at venues with unstable and low bandwidth Internet connections (See also activity 1e and 3b).
- **South Africa:** SANBI-GBIF is recruiting a Biodiversity Informatics Support Officer to support training and engagement portfolio of work. This will support national and regional efforts and tie into GBIF helpdesk requirements as needed.
- **Spain:** The e-learning platform of GBIF.ES has been used to host training modules for BIF and BIFA programmes, as well as for other capacity programmes within the network such as Biodiversity Data Management Skills for Students (BioDATA) and Biodiversity Information Management and Reporting (BIMR), reaching communities in Eurasia and South-East Europe respectively. GBIF Argentina has also used our e-learning platform to offer a course on biodiversity data quality and publication. We trained staff from Secretariat in using the Spanish e-learning platform. GBIF Spain is organizing an online workshop on R open to be run in November 2019 addressed to the national and international community. GBIF Spain co-coordinated with ALA and GBIF France the Living Atlas International Workshop that took place in Paris (March 2019) and used our informatics infrastructure to test the exercises.
- **Sweden:** Representatives from GBIF-Sweden has participated in BID activities as Trainers and mentors (South Africa, Russia, Trinidad/Fiji/Cameroon, the Netherlands (GB26).

- **United States:** Contributed to Marine Biodiversity Observation Network Pole to Pole workshop to train Central and South American country representatives in aligning data to Darwin Core and sharing via IPT.
- **Zimbabwe:** Trained stakeholders (data holders and data users) on GBIF data access – January 2019.

2020 Work items

- Develop joint training resources combining GBIF data mobilization and DNA barcoding, and support pilot workshop with BOLD, CBD and GBIF (€20,000).
- Explore opportunities for additional training workshops on data use for decision making.
- The BioDATA project will continue to train students in Armenia and in Ukraine in 2020, and plans for BioDATA II project are in place to further increase the geographic coverage in Northern and Central Asia and Southern Caucasus.
- Following the recommendations of the community-based Documentation Editorial Panel, continue to commission new documentation, based on late 2019 review and prioritization framework (€30,000).
- Consolidate online e-learning platform with training resources, to make the modules on data mobilization, data use for decision making and nodes training fully functional as self-instruction modules. Use external contracts for some components as necessary. Develop a strategy to meet future demand for on-site, moderated and self-instructed training. Work in collaboration with key partners, including GBIF nodes, the pool of volunteer trainers and mentors, Nodes Steering Group and numerous international partner projects (€20,000).

2020 Participant plans

- **Australia:** Attendance and involvement in future workshops and efforts to improve documentation.
- **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.
- **Belgium:** Complete set of explainer texts in multiple languages: French and Dutch translations.
- **Benin:** “Capacity building through workshops and in the framework of the master program in biodiversity informatics Data mobilization Data uses.
- **Canadensys:** Canadensys will continue to work on developing the documentation about data-cleaning, publication and data-use. This documentation will be published on our community page (<https://community.canadensys.net/>) but will be available to reuse for other institutions or nodes. This documentation will probably be published as short blog post on specific topics.
- **France:** GBIF France will continue support for capacity enhancement activities and workshops.
- **iDigBio:** iDigBio will continue its collaboration and relationship with The Carpentries. iDigBio has several workshops, webinars, symposia, and other events planned.
- **Japan:** Provide two workshop for data carpentry and more use cases to be collected.
- **Naturalis Biodiversity Center:** Further training and capacity enhancement activities will be

organized through the COST Action MOBILISE.

- **Norway:** BioDATA will during 2020 organize data mobilization training events in Armenia (April 2020) and in Ukraine (October 2020).
- **South Africa:** SANBI-GBIF will continue to conduct training workshops and develop communities of practice in identified biodiversity informatics areas.
- **Spain:** We will replicate the Data Use For Decision Making workshop for the Iberoamerican community in coordination with other nodes from the region. We will continue to provide support for GBIF coordinated training through our e-learning platform, as well as for other members of the network (e.g. GBIF South Africa). We will work together with the Secretariat to expand the functionalities of the GBIF.ES e-learning platform. We will attend and support future workshops and efforts to improve documentation around the Living Atlases platform.
- **Sweden:** Experiences gained from the above will be integrated in educational and outreach activities nationally and elsewhere. Participation in the development of biodiversity informatics training curricula will continue.
- **United States:** USGS will continue to participate as a teacher in workshops and meetings to advance data curation skills globally and across domains.
- **Zimbabwe:** Data mobilization workshop in January 2020 Data access workshop in March 2020.

Rationale

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.

Approach

To strengthen key skills, GBIF needs to develop and maintain a comprehensive set of clear reference information and training materials that support all GBIF audiences. Since 2016, the Secretariat has been upgrading documentation to deliver this comprehensive resource, first via concise explanatory text materials (including numerous translations by network members into French, Portuguese, Spanish, Chinese, Japanese and Russian). Some of these text materials will be supported by short videos that present key functional and operational aspects of GBIF. The aim is to deliver a structured set of short, clear, current and complete documents that enable GBIF stakeholders to gain an understanding of relevant aspects of GBIF's work quickly, including formal Participation, establishing a national or thematic node, planning digitization, mobilizing data, discovering and using data, citing data, among others. This work complements significant activity among Participants to produce information and training materials relevant to particular audiences.

GBIF will organize all these materials to form a curriculum that builds and supports the skills and activities identified in GBIF's self-assessment tools for nodes and data holders. A core set of these resources should provide new stakeholders with an effective introduction to GBIF's structure, roles,

approaches and solutions. Other more technical resources will provide the additional detail needed by those working in specific areas like digitization, data publishing and use of GBIF-mediated data. The Secretariat will coordinate ongoing maintenance and updates to this curriculum with contributions and translations from the network.

Activity 1c: Equip Participant nodes

Tasks

- i. Perform ongoing nodes capacity self-assessment
- ii. Build reusable national/regional platform tools centred on ALA software
- iii. Develop online collaboration through GBIF.org helpdesk to assist and mentor node managers
- iv. Explore alternative options for lightweight national portal solutions, e.g. CKAN

2019 Progress

Based on discussions alongside the GB25 meeting in Ireland in 2018, the Kilkenny Accord developed a roadmap towards an open governance and funding model for the Living Atlas community. GBIFS established and has overseen contracts with an Administrative and Technical coordinator to support the community (€40,000) in 2019. GBIFS assisted in the ALA Advanced Workshop in Paris to help adopters move towards Spatial Portal capabilities. The Secretariat informatics team held a technical meeting with the ALA at Canberra in March to explore possibilities to collaborate, with the Registry/Collectory and a hosted data indexing process identified as an opportunity to reduce the costs of operation for Australia and show the value proposition of GBIF.

In 2019 the Secretariat started exploring the potential for a lightweight portal solution that has been requested by several participants to lower the technical threshold. A concept of the simple hosted solution (static site generator and occurrence search) has been demonstrated to the Science Committee and Nodes Steering Group, and a visual design is in development for GB26. A showcase portal containing data generated through projects funded by the BID programme in Africa, Caribbean and the Pacific has been chosen as the initial project in 2019.

The informatics team opened the [GBIF Registry administration](https://registry.gbif.org) [https://registry.gbif.org] and piloted editing by external users (starting with GBIF Norway node staff). The administration console provides editing and registration capabilities of the core entities (institutions, datasets, etc.) and also provides oversight and control of the dataset crawling infrastructure and diagnostics. This is part of a strategy to progress towards becoming a more open platform for others to use by lowering the technical threshold through enabling infrastructure.

A global nodes meeting and training event is taking place in conjunction with GB26 in Leiden, Netherlands in October (see 1b above). The training is a two-day course focused on the functional capacities involved in node development: positioning a node in the broader landscape, identifying key stakeholders, developing strategic and implementation plans and running effective training workshops. The online materials are being made public by the end of the 2019.

The GBIF network has initiated a programme creating a new role, the Biodiversity Open Data Ambassador, aimed at supporting a corps of experts who promote free and open access to

biodiversity data. During its initial development phase without significant promotion, 42 professionals from 19 countries, representing each of GBIF's six regions, volunteered to become Biodiversity Open Data Ambassadors. The programme is intended to recognize and support the active role that ambassadors have alongside the network's node managers and staff, formal delegates, committee members and Secretariat staff. The GBIF Secretariat continues to develop resources to help ambassadors' efforts, including presentation slides, talking points, posters, and support materials in both digital and print formats.

2019 Participant contributions

- **Andorra:** On February 2018, Atlas Living Andorra was launch and presented to media. The ALA-Andorra was undertaken by a company specialized in informatics systems and the assistance of our colleagues of the Spanish Node.
- **Argentina:** Capacity self-assessment ready, mentor to new nodes.
- **Australia:** The Atlas assisted in the recruitment of two Living Atlas Coordinator positions with GBIF's funding; a Technical and Administrative position. These positions will help develop the community further by organizing the governance arrangements around the Atlas tools and improving support and documentation for nodes attempting to use the Atlas tools for their biodiversity informatics portals.
- **Benin:** In the framework of our master program, courses are being registered and will be provided to JRS Biodiversity Foundation, our donor. It will be then popularized and disseminated to be available for the general public.
- **Canada:** CBIF is reviewing use cases and considering options for a new Canadian biodiversity portal. We participated in the workshop "Living Atlases workshop (CESP) 2019" held in Paris, France in April. CBIF participates in the Biodiversity Open Data Ambassador and has included GBIF materials in talks at conferences/meetings on biodiversity informatics topics.
- **Canadensys:** Our node manager (Carole Sinou) is involved as a trainer in the BID and Global Nodes Training. The Canadensys team is also involved in the Living Atlases Community and Carole is part of the interim steering group. Anne Bruneau (HoD) and Carole will participate in the GB26 and Biodiversity_Next conference.
- **Colombia:** Ventanas departamentales / Regional windows: Starting from a need for regionalization of SiB Colombia that has been latent for some time, during the first quarter of 2019 the first instance to explore biodiversity at departmental level in the country was launched. The first department is Santander. This window is a tool for exploring the data and information available about a specific department and sow the seed to develop a very useful tool for decision-making and territorial management at the regional level, through appropriation, awareness and empowerment. Data repatriation: from September 2018, the SiB Colombia has been working in the data repatriation process to the national data portal. In the first quarter of 2019, we repatriate 3,773,350 occurrences from eBird, iNaturalist, Xeno-canto and GBIF.org (<https://sibcolombia.net/integracion-datos-2018/>). In the third quarter we hope to support the hosting of a strategic GBIF dataset from CIAT - A global database for the distributions of crop wild relatives. <https://doi.org/10.15468/jyrthk> Participation in CESP: Increasing capacities to develop National Species Checklists in the Latin America and the Caribbean Region.
- **France:** GBIF France helped to develop a portal based on the ALA software for GBIF Togo. GBIF

France coordinated the CESP project : “International Living Atlases workshop: how to improve data use with Atlas of Living Australia modules” and co-organized with ALA the Paris workshop in April 2019. GBIF France will supply mentor for workshops organized by GBIF under BID and GNM on managing a GBIF participant node.

- **iDigBio:** iDigBio has been participating in the monthly GBIF webinars when feasible. If live participation was not possible, the recordings have been utilized. iDigBio will be participating in the Node Training at GB26.
- **International Centre for Integrated Mountain Development:** Joint organization of 2018 Global Biodiversity Information Facility (GBIF) Asia Regional Nodes Meeting in Kathmandu on 17–18 September 2018 (<https://www.gbif.org/event/EEvUGytWxw48WQgkAwCkg/2018-gbif-asia-regional-nodes-meeting>). Reactivated HKH-BIF (re-registration with GBIF). Participated in GBIF Network Webinar#1: Understood some elements of sequenced based data mobilization, and received introduction on Biodiversity Information for Development and BIFA.
- **Japan:** Continuous improvement of S-Net system (domestic website for searching natural history specimens).
- **Mexico:** Develop websites Enciclovida (<http://enciclovida.mx>), SNIB's geoportal (<http://geoportal.conabio.gob.mx/>) and SNIB (<http://www.snib.mx>)
- **Netherlands:** NLBIF made a financial contribution to the Living Atlases project and started to look into implementation of Living Atlases for NL. The NLBIF node manager has joined the Open Data Ambassador Programme.
- **Norway:** GBIF Norway contributes as mentor to the regional GBIF network meeting in the Pacific in August 2019.
- **Spain:** We have contributed to translations of training modules on managing a GBIF Participant Node, and we have hosted the online activities for these workshops. We have been mentors in the training activities on managing a GBIF Participant Node carried out in Trinidad. We will be mentors in the Global Nodes Meeting training in Leiden. We have been supporting GBIF Zimbabwe in establishing its GBIF Node. We have supported the development of the Living Atlases (LA) community. We are part of the LA interim committee and we have been active with the LA communication and outreach, promoting LA in different forums and conferences. We have supported and assisted new LA coordinators that will be helping the nodes that intent to use the Atlas tools. We have worked together with Nodes from UK, Australia, France, Sweden, Portugal and Canadensys to create video tutorials and online exercises for LA users around the world. We have supported other data managers of other Living Atlases installations.
- **Sweden:** Reiterated capacity self-assessments have been made and the result added to the European nodes regional report (and we do well!). Additional ALA platform tools are being developed and offered to the Living Atlases Community. Apart from data publication (see below) the latter is probably the most significant contribution from GBIF-Sweden to the global biodiversity informatics community.
- **South Africa:** SANBI-GBIF will consider mechanisms to engage the scientific community around the new and evolving concept of the GBIF Open Science Ambassador Concept. Here the Node will take forward the initial advocacy actions with the marine science community, conducted at the South African Marine Science Symposium. SANBI-GBIF has highlighted research aimed at developing the field of biodiversity informatics in South Africa through the use of primary data (big data –more than 2 million records) and informatics tools (niche

models) to address key biodiversity challenges (food security and climate change). This demonstrates the use and relevance of GBIF mobilized data across the value chain. Showcase examples for Southern Africa will be derived.

- **United States:** Continued collaboration with iDigBio building nodes capacity and cooperation.
- **Zimbabwe:** Purchased and installed server for hosting data for GBIF Zimbabwe.

2020 Work items

- Support for 2020 Regional Nodes Meetings (€60,000). The Secretariat will investigate possible synergies with BID phase 2 regional meetings, as well as with a BIFA workshop in Asia.
- The Secretariat will continue enhancing guidance documentation for nodes. This guidance will incorporate priorities in data mobilization, gap filling, private-sector engagement and products from CESP projects.
- The Secretariat will design and implement user interfaces and services necessary to support simple hosted portals. The initial deployment will be of the BID programme portal followed by exploration of country portals and an exemplar virtual natural history collection. Discussions will determine whether such services should be restricted to Voting Participant countries and/or institutions that contribute a fee to create institutional portals.
- Continue conversations to align codebases and infrastructure with ALA, iDigBio, DiSSCo and other potential partners. The alliance for biodiversity knowledge will be used as a platform for some of these interactions.
- Support further development of the Living Atlases community (minimum €30,000).
- Consolidate Biodiversity Open Data Ambassador programme with issuing of digital badges, listing ambassadors on country pages and capturing feedback on promotion of GBIF at events.

2020 Participant plans

- **Andorra:** We will attend to improve Capacities and Tools of ALA-Andorra
- **Argentina:** Mentor to new nodes
- **Australia:** Continued support for the Living Atlases Community, providing support and guidance to the Living Atlas coordinators.
- **Benin:** In the framework of our master program, courses are being registered and will be provided to JRS Biodiversity Foundation, our donor. It will be then popularized and disseminated to be available for the general public.
- **Canada:** CBIF will contribute to discussions/evaluations about lightweight national portals and potentially collaborate with GBIF to further develop/support/sustain an option that would be suitable for Canada and other countries.
- **Canadensys:** We will continue to be involved as much as we can in the different initiatives developed by GBIF. It is important to give back to the community as much as we can.
- **Colombia:** Ventanas departamentales / Regional windows: During the first quarter of 2020 will be released the version 1.1. The second department will be Boyacá. This window is a tool for exploring the data and information available about a specific department and sow the seed to develop a very useful tool for decision-making and territorial management at the regional level,

through appropriation, awareness and empowerment. Data repatriation: Data repatriation of new occurrences between January – December 2019.

- **France:** GBIF France will continue to participate in the activities of the Living Atlases Community.
- **iDigBio:** iDigBio finds the monthly GBIF webinars valuable and plans to continue attending. iDigBio is currently in a sustainability planning process. As part of these efforts, iDigBio will consider capacity self-assessment as an information tool. In addition, iDigBio plans to evaluate Living Atlases as a potential long-term portal infrastructure. Lastly, iDigBio will consider taking advantage of the node mentorship program.
- **Japan:** Continuous improvement of S-Net system (domestic website for searching natural history specimens).
- **Mexico:** “Continue to update data occurrences in published datasets. Continue to gather new occurrences for publish in Mexican node”
- **Norway:** As part of the BioDATA project, GBIF Norway will provide node mentoring services to partner countries who make progress towards joining GBIF as new members (see also Activity 1e).
- **Spain:** We will keep on supporting the Living Atlases community: support coordinators to find a model of governance for the community; assist other data managers of the different Atlas around the world. We will work to develop online exercises on different modules of the Atlas to continue improving LA documentation for users.
- **Sweden:** Continued development and addition of ALA tools will be shared through the new Swedish “Bioatlas” and by sharing code at GitHub. Support to other nodes employing techniques where we may be of help will continue/increase.
- **South Africa:** SANBI-GBIF will highlight the role of GBIF and engage the scientific community at the annual Biodiversity Information Management and Foundational Biodiversity Information Programme Forum, on the Biodiversity Open Ambassador programme

Rationale

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.

Approach

GBIF promotes capacity self-assessment as a tool to help nodes with planning and tracking progress at all stages in the development of national biodiversity information facilities. The results of these assessments can also identify important capacity needs to address with learning materials, tools and collaborative projects. The reuse and adaptation of the Atlas of Living Australia’s open-source

software tools by a growing community of nodes between 2013 and 2016 is an example of successful collaboration based on the mutual needs of the countries involved. Additional support for this international partnership will allow still more nodes to benefit from the collective resources and expertise to build websites that provide biodiversity data and services to a range of national and regional user communities. The nodes' teams will also engage in collaborative helpdesk activities to enable them to share knowledge across the scientific and technical domains handled by the nodes.

Activity 1d: Equip data publishers

Tasks

- i. Promote and support capacity self-assessment for data holders
- ii. Promote publication of collection metadata
- iii. Simplify data publication pathways (spreadsheet-level publishing)
- iv. Manage IPT feature upgrades
- v. Operate hosted IPT infrastructure
- vi. Consolidate NSG-led endorsement process
- vii. Develop online collaboration through GBIF.org helpdesk to assist and mentor data publishers
- viii. Provide clear online reporting of the use of data for data publishers
- ix. Promote data management plans as key tool for data publishers (Added 2018)

2019 Progress

The Integrated Publishing Toolkit (IPT) is being maintained to deal with bug fixes, translations etc., and version 2.4 was released in 2019. Review of requirements for a revised IPT is on hold until there is greater clarity on how broad a data model for GBIF should be (see Activity 2a). With a new model in place, GBIF will design appropriate tools to map and publish data. The cloud-based IPT infrastructure maintained by the Secretariat continues to provide a fall-back solution for publishers unable to host their own installation or to find a third-party hosting option. As of June 2019, this option was being used by nine publishers associated with the BID programme and one publisher associated with the BIFA programme, sharing a total of 100 datasets. Cloud-hosted IPTs are expected to become more widely used in the coming months, especially with the use of volunteer mentors to provide help desk support for users of this option.

2019 Participant contributions

- **Andorra:** Increase biodiversity data amount related with Andorra.
- **Argentina:** Argentina has a collection catalogue (and institutions) since 2003. We keep updating it.
- **Australia:** The Atlas worked with the collections to identify the trait data available in existing database held in data managements systems by natural history collections. In addition, the Atlas engaged with members of the research community in Australia to discuss use of traits.

- **Belgium:** Hosted IPT installations for two African nodes.
- **Benin:** Capacity building is achieved to data publishers during our workshops in achieving data format, data cleaning, and data publishing as well. They are also assisted and encouraged to register on GBIF site. When they are ready to publish their data, they are intensively assisted by members of GBIF Benin to have data in adequate format, and cleaned before publishing.
- **France:** GBIF France offers data hosting and maintains IPT instances for 14 southern countries.
- **Germany:** Mobilizing biodiversity data: By mid of 2019 more than 47 million occurrences within 35,000 datasets covering 252 countries and areas have been made available by 40 publishers from Germany. By end of 2019 occurrence data, especially for the vascular plants for Bavaria (“Flora von Bayern Initiative”) will be significantly extended and the dataset for German fungi will be updated. German GBIF-Nodes continue to develop the DCOLL initiative with the aim of digitizing all German natural history collections (large-scale funding approval pending) and collaborate in the corresponding European Initiative (DiSSCo). Continuous support for Data providers using the BioCAsE software for data publishing in GBIF, Europeana, GGBN, and other special interest networks. Implementation of further routines for data quality control are implemented in the Diversity Workbench software suite and in JACQ.
- **iDigBio:** As of the end of May 2019, iDigBio had approximately 71 datasets in various stages of mobilization. As part of our data mobilization efforts, iDigBio staff periodically provide IPT help, support, and training to data publishers. iDigBio partnered with Environmental Data Initiative (EDI), Earth Science Information Partners (ESIP), DataONE, GBIF, NEON and Arctic Data Center on a Data Help Desk at the 2019 meeting of the Ecological Society of America. The goal was to engage participants one-on-one about their questions, comments and concerns about using these data in their discipline and area of research.
- **Japan:** New IPT server established, and started to provide data to GBIF. Tools to convert/check data from Excel to DwC provided.
- **Mexico:** Increase number of occurrences through Mexican node at least 500,000 records planned for the last quarter. In progress publish two national checklist “Aves de México” and “Helmintos parásitos de vertebrados” with approximately 2,000 valid species.
- **Naturalis Biodiversity Center:** Naturalis has discussed with GBIF Secretariat to provide an IPT through DiSSCo for serving collection data from countries in Europe participating in DiSSCo that have no GBIF node yet.
- **Netherlands:** Through several workshops NLBIF has engaged with most of the Dutch Natural history museums and explained the ease of DwC and the possibility publication in the Biodiversity Data Journal. In the second half of 2019 this will have follow up and expected contributions of data to GBIF. Furthermore, NLBIF supports and promotes data publishers in hosting their own IPTs. And NLBIF is supporting data publishers to add multimedia files to their GBIF records.
- **Norway:** GBIF Norway provides a helpdesk for data publishers in Norway, and will continue to host datasets for selected data publishers outside of Norway upon request – including data hosting requests from BioDATA partners. GBIF Sweden assisted GBIF Norway with the assessment of proposals and selection of data publishing co-funding grants for mobilizing Norwegian biodiversity data in GBIF.
- **South Africa:** Strategic engagements/meetings between SANBI-GBIF Node Manager and South African Head of Delegation will continue in order to evolve the South African Node planning,

and Africa portfolio of work and to elaborate the Science Diplomacy role SANBI-GBIF can play. Phase 2 (of national data platform) commenced in October 2018 (24 months), which looks at the implementation phase of the NBIS.

- **Spain:** We are using our IPT installation in test mode to help other Nodes and providers in the data publication process.
- **Sweden:** GBIF-Sweden hosts IPT installations of successively more data providers nationally.
- **United States:** Worked directly with new data publishers to align their data to Darwin Core and share it via IPT and provide support and guidance on the use and implementation of the IPT, data preparation, and Darwin Core implementation as needed/requested.

2020 Work items

- Promote wider editing of the GBIF registry and the shared help desk activities; including node staff initiating and diagnosing dataset crawling / ingestion.
- Complete implementation (if not finished in 2019) and develop processes to allow open editing of shared vocabularies used in data interpretations of the GBIF ingestion pipelines (e.g. habitat types, occurrence status, etc...).
- Create system of list management, similar to bulk email, to communicate to a larger section of the GBIF community for compliance and notifications. This includes a twice-yearly mandatory communication with data publishers in compliance with General Data Protection Regulation (GDPR: EU privacy regulations) and explore services for publishers to opt-in to receive push notifications for new citations.
- Provide comprehensive guidance and support services to lower the technical threshold of data-hosting options. Clearly document the benefits and implications of each option including aspects of operational cost, deployment model (local/cloud/GBIF-provided) and expectation of users. Use of volunteer mentors will be promoted to enhance help desk services.

2020 Participant plans

- **Andorra:** We plan to keep working of gather biodiversity data of Andorra as much as possible. Another goal will be make to know GBIF in Andorra and so increase the number of entities interested in publishing their data in the GBIF portal.
- **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.
- **Australia:** Develop a community of practice in the management of trait information. Develop a roadmap for trait mobilization activities.
- **Belgium:** Host IPT installations.
- **Benin:** Capacity building in data mobilization and data uses.
- **Canadensys:** If needed, our team can help as mentor or as tester for the cloud-hosted IPT.
- **France:** GBIF France will continue to support data hosting and publishing services for southern countries.
- **Germany:** Independent of the success of the DCOLL Initiative, German GBIF Nodes will continue to support digitization efforts in collections and publications of observation datasets.

Expected are significantly increased numbers in occurrence data of vascular plants, further digitization of German collections, fully referenced diatom data from the of the German Barcode of Life Initiative and further improvements in the JACQ (Virtual herbaria) and Diversity Workbench software. Continuing BioCASE Helpdesk. Implementation of ABCD 3.0 in BioCASE and beyond.

- **iDigBio:** iDigBio is currently in a sustainability planning process. As part of these efforts, iDigBio will evaluate data holder capacity self assessment as a method of improving our ongoing data mobilization processes.
- **Netherlands** NLBIF will continue the 2019 activities.
- **South Africa:** SANBI-GBIF hosts an IPT which supports both national and regional data publishing. This includes helpdesk support.
- **Spain:** GBIF Spain will continue to offer online support to data publishers also outside Spain in using IPT. We will assist GBIF Zimbabwe to configure and maintain its own IPT.
- **Sweden:** GBIF-Sweden will continue to offer services and support for data publishers covering all kinds of data. We expect progress also within the field of molecular data publication.

Rationale

Data publishers are an essential component of the GBIF network as they share their content through the common infrastructure. More than 800 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.

Approach

Following the model of the self-assessment tool for node managers, the Secretariat has developed a self-assessment tool for data publishers as part of the support for the BID programme, which will guide the work with collection managers and other data holders to assess and prioritize areas for capacity enhancement or investment. The Secretariat already operates instances of the IPT that data publishers lacking their own infrastructure may use, and Participants are encouraged to deploy instances of the IPT or other GBIF-compatible data publishing software to support data holding institutions. Planned enhancements to the IPT will simplify publishing pathways using spreadsheet templates as an alternative for the less advanced data publishers. GBIF will improve reporting to data publishers on both quality aspects of their data and uses of data documented through download DOI citations.

Activity 1e: Expand national participation

Tasks

- i. Ensure GBIF.org interface and key content supported in relevant languages
- ii. Improve coordination and mechanisms for engaging new countries
- iii. Provide Governing Board with annual review and progress assessment for GBIF status and engagement in all countries

2019 Progress

The internationalization of the GBIF.org user interface and core content elements, in the six UN languages as well as selected additional languages, is expected to be complete by the end of 2019.

Technical adjustments have been made to GBIF.org to accommodate right-to-left script needed for Arabic and other languages. A volunteer Arabic-speaking network is being developed. Translations of GBIF.org into Russian and Chinese (both simplified and traditional) are far progressed and are being launched later in 2019, adding to the existing interfaces and core content in English, French, Spanish, Portuguese and Japanese.

Three 2019 regional meetings within the BID program are providing opportunities to strengthen and expand national participation in the Caribbean, Pacific and sub-Saharan Africa. The Caribbean meeting in Trinidad and Tobago was attended by over 40 participants from across the Latin American and Caribbean region, including a strong delegation from the host country coming from various institution and decision-making bodies. Strong attendance is expected in the other regional meetings later in 2019.

The main outputs of the meetings are regional statements on regional priorities (e.g. from the Caribbean), and roadmaps for ensuring long term sustainability of biodiversity data sharing. The final day of the event focuses on node management training including the benefits of GBIF participation and being part of the GBIF nodes community.

During 2019, Angola and Belarus joined the community as Associate Participants, and Chile returned to Voting Participant status after a period as an Associate Participant. Brazil communicated its intention to become a Voting Participant. Israel, Indonesia and India dropped to observer status after passing the maximum period as Associate Participants. Argentina, Ghana, Kenya, Uganda and the United Kingdom moved from Voting to Associate status on 1 January 2019 due to non-payment of their contributions for two consecutive years. A complete and updated progress assessment of the GBIF status and engagement in all countries is being presented at the GB26 meeting in the Netherlands.

2019 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.
- **Benin:** We are invited to attend workshops organized by partners to make presentations on GBIF, how to publish data, how to register on GBIF site and the usefulness of data uses.
- **France:** Help for French translations: GBIF.org, workshop materials. Co-organized with Norway and the NSG chairs, invitation of the project leader of a Balkan project and few observers of

Balkan and eastern countries to the EU nodes meeting in Oslo, Norway (May 2019).

- **Germany:** [National and International] German GBIF Node institutions are participating in numerous international collaborations and are constantly promoting participation in countries that have not yet joined GBIF.
- **iDigBio:** The iDigBio Node Manager currently serves as the Regional Representative for participant nodes in North America. 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. iDigBio will be hosting the annual ADBC Summit in October 2019, which brings together representatives from NSF's Advancing Digitization of Biodiversity Collections (ADBC) program.
- **Japan:** Several new organization started to provide data to JBIF (currently 100 organizations).
- **Mexico:** The SNIB (Sistema Nacional de Información sobre Biodiversidad) process was published on the portal snib.mx, the summary process and phases diagram from the collection of specimen information to its publication on a web page. Include 8 documents describing subprocesses inside general process SNIB.
- **Netherlands:** NLBIF facilitated the contribution of all data from observation.org, the largest citizen science portal of the Netherlands, with more than 26 million records. NLBIF has funded a project aimed at connecting the Ukrainian UkrBin data <http://www.ukrbn.com> to GBIF and further improve the engagement of citizen scientists in Ukraine. The NLBIF node manager has presented the requirements of becoming a GBIF participant at an online conference meeting in Ukraine.
- **Norway:** GBIF Norway funded and organized the 2019 European GBIF Nodes meeting in Oslo in May 2019 and provided funding for inviting observers from Armenia and Belarus, and co-funding together with GIZ for an observer from Croatia to join the European Nodes meeting.
- **South Africa:** The ABC project follows an incentivised approach for data mobilization. As a final outcome of the project, a prize-giving ceremony will be conducted in February 2019 based on a number of criteria. This approach will demonstrating the use and application of data which has been mobilized, for aspects of societal need and/or policy relevance. Through SANBI-GBIF, new countries in Africa will be identified to further the biodiversity informatics agenda in support of science and sustainable development. National networks will be developed in the region through the implementation of further BIMFs and GBIF advocacy will be conducted.
- **Spain:** We have helped translating gbif.org and <https://www.biodiversityinformatics.org/> to Spanish. We have run data mobilization onsite and online training activities with national and international participants to engage new participants.
- **Sweden:** GBIF-Sweden has shown interest in supporting presumptive new participants in the community.
- **Zimbabwe:** Data holding institutions were identified.

2020 Work items

- The GBIF Secretariat will continue to identify funding opportunities for regional engagement, including inviting observers to regional node meetings. If budget allows, it will create a specific funding priority area alongside the 2020 CESP call for GBIF regional outreach activities, as outlined in regional participation strategies developed across the network.

- The Secretariat will develop and implement engagement, data-mobilization and capacity-enhancement plans for greater participation in all global regions, with an emphasis on the underrepresented areas. This effort will be aided by capacity enhancement workshops, including through a second phase of BID and other supplementary funding that becomes available.
- The internationalization of GBIF.org user interface, content and guidance materials will continue. Assuming the 2019 commitment for six UN languages is complete, this work will focus on maintenance and updates and will allow the addition of more languages on demand, and translating a wider set of content and materials, funded through both specific projects and volunteer efforts.
- Engagement, data mobilization and capacity enhancement campaign for Arabic-speaking region (€20,000, if funds allow).

2020 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
- **Benin:** Data publication, Registration of data publishers, data uses
- **Canadensys:** Carole is deputy regional representative for the North American region and will continue to be involved in both the region and, if possible, in the BID program, as mentor.
- **France:** Translations to be continued
- **Germany:** Continued promotion of GBIF in national and international initiatives.
- **iDigBio:** iDigBio is willing to continue offering its Node Manager as the Regional Representative if re-elected. iDigBio will continue its participation and involvement with BCoN. iDigBio is planning a 2020 Biodiversity Summit, which will be a collaborative meeting of the ADBC community, GBIF governing board, TDWG, and the National Museum of Natural History (Smithsonian). The meeting will feature the evolution and accomplishments of specimen-based science and the impact of digitization. This international event will be an excellent opportunity to highlight to a global audience the excellent work that our ADBC community is doing on behalf of U.S. collections.
- **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 HKH-BIF regional platform for data publishing or use of available in-country platforms such as in India and China.
- **Japan:** Provide meeting occasions to enroll more organizations.
- **Mexico:** Prepare documents about National Catalogue of species.
- **Netherlands:** Continue same work programme.
- **Norway:** The BioDATA project managed by GBIF Norway contributes to promoting expanded GBIF country membership in the former Soviet Union region (including strategic partners from Belarus, Tajikistan, Armenia, and Ukraine). In addition to the primary goal of providing academic training (Activity 1b), BioDATA aims to build capacity needed for establishing national GBIF Participant Nodes (Activity 1c and 1e).

- **South Africa:** A service provider has been secured and SANBI-GBIF will consider the ideal strategic approach to support expanding participation, and will engage with the GBIF Secretariat to ensure coordination.
- **Spain:** Translation of gbif.org to be continued. Explore ways to expand participation in North African countries in collaboration with Ministry of Science.
- **Sweden:** GBIF-Sweden will primarily be engaged in the Living Atlases Community as a means to contribute to the expansion of the global network of GBIF.
- **United States:** Planned collaboration with US marine community participants on a workshop to foster uptake of Darwin Core in marine biological data communities, draft controlled vocabularies for this community, and assess overlap with other ocean observing standards like NetCDF and Climate and Forecast.
- **Zimbabwe:** Identification of more data holding institutions.

Rationale

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.

Approach

The Secretariat will compile and maintain an up-to-date register of the status of GBIF engagement and relevant contacts in all countries, providing the Governing Board with an annual review that analyses progress and remaining gaps in membership across each region. Following the development of a coordinated engagement plan for each region, Participants and regional groups with strong contacts to relevant agencies in other governments may take the lead in engagement. The Secretariat will continue to develop strategic engagement through the CBD and other intergovernmental networks, as well as through the BID programme and other capacity-related activities. Adapting the GBIF.org interface to support key content in multiple languages will seek to address language barriers to national expansion, with targeted translation carried out both through external contracts and in-kind services provided by Participants.

Activity 1f: Plan implementation

Tasks

- i. Develop better integration between nodes activities and Governing Board processes
- ii. Expand regional meetings to serve as GBIF community Implementation Planning conferences and opportunities to develop fundable programmes of work
- iii. Increase alignment between GBIF work programmes and activities of each GBIF Organizational

- iv. Explore increased alignment and joint organization of workshops with other networks such as TDWG, SPNHC, GEO BON, RDA, etc.

2019 Progress

GBIF leads work package five of the European initiative SYNTHESYS+ project, which includes running three workshops. The focus of these workshops, defined by the global community, is to be specified in Q3 2019, but is anticipated to include activities around collection cataloguing, richer specimen models and tracking of citations of specimens, collections and the individuals, all of which are also priorities for GBIFS (see also activity 1a).

The Secretariat has been involved at many organization levels (Steering Committee, Programming committee) of the Biodiversity_Next conference held in conjunction with GB26. Secretariat staff are involved in over 20 talks and organizing several sessions that relate directly to the Work Programme. Involvement has begun with a similar meeting scheduled to take place in September 2020 in Virginia, USA.

The Secretariat is developing standardized annual information requests based on the GBIF Strategic Plan and Implementation Plan structure aimed at capturing relevant updates and work plans from Participant nodes. The work is intended to improve the visibility of the GBIF Participants, and especially of the nodes, on GBIF.org and to better reflect the network as a whole, particularly with regard to the active contributions toward work programme implementation.

The *alliance* for biodiversity knowledge was established as an outcome of the 2018 GBIC2 conference hosted by GBIF in Copenhagen. The Secretariat has reconstituted the GBIC2 steering committee as the steering committee for the *alliance*. In conjunction with external stakeholders, the Secretariat is developing a virtual workshop infrastructure and will be hosting the first of a series of virtual workshops under this infrastructure later in 2019. The Secretariat is also driving efforts to develop and disseminate *alliance* communication materials and to build expertise networks for managing virtual workshops and networking models.

2019 Participant contributions

- **Argentina:** Coordination, as a regional representative, in the reports from each node for the joint regional report.
- **Australia:** Analysis of GBIF pipelines was undertaken, and project plans for further work with GBIF's code base have been made.
- **Benin:** We advanced in data mobilization and data uses, we encouraged institutions to register on GBIF site so as to become more visible.
- **Canada:** CBIF participated in the North American Nodes Meeting where discussions of ways to better coordinate regional, country and continental activities with GBIF plans took place. We currently also contribute to plan implementation through supporting the efforts of the HOD for SPNHC and TDWG (Dr. Macklin). Dr. Macklin is also the current Chair of TDWG and is working with GBIF to advance mutual goals in standards development and biodiversity informatics. We also contribute as a member of the Steering Committee (via TDWG) for Biodiversity_Next 2019, a critical forum for discussion of the goals of GBIF and the *alliance*.

- **Canadensys:** Our team has helped in the translation of the documentation for the *alliance*.
- **France:** Co-organization with Norway of the EU nodes meeting in Oslo. Election of the new EU representative and Deputy.
- **iDigBio:** iDigBio hosted the regional meeting of participant nodes in North America in association with the 2019 Digital Data in Biodiversity Research Conference.
- **Naturalis Biodiversity Center:** A meeting was held between the DiSSCo coordination office and GBIF Secretariat to discuss coordination of implementation activities. Naturalis provided D2.3 Design of a collection digitization dashboard in ICEDIG which provides input for further development of GRSciCol hosted by GBIF.
- **Netherlands:** NLBIF has adopted the DiSSCo-NL NTF lead and as such integrates GBIF and national DiSSCo activities.
- **Norway:** GBIF Norway funded, hosted and organized the 2019 European GBIF Nodes meeting in Oslo in May 2019 (see also Activity 1c and 1e). GBIF Norway and the GBIF Norway network partners serve on GBIF committees as elected vice-chair for the GBIF Science committee and as vice-chair for the European GBIF Nodes committee.
- **South Africa:** Strategic engagements/meetings between SANBI-GBIF Node Manager and South African Head of Delegation will continue in order to evolve the South African Node planning, and Africa portfolio of work and to elaborate the Science Diplomacy role SANBI-GBIF can play. Phase 2 (of national data platform) commenced in October 2018 (24 months), which looks at the implementation phase of the NBIS.
- **Spain:** The GBIF.ES web page hosts activity reports from European Nodes that can be useful for other reporting activities Nodes. We have been trying to coordinate efforts on how to integrate EU Nodes services in the EOSC. We have communicated the progress done by GBIF in developing a global *alliance* for biodiversity knowledge in our national context.
- **Sweden:** GBIF-Sweden has participated in regional meetings, and at conferences adding to the implementation of prescribed GBIF activities. By providing the 1st vice chair of the Nodes Steering Group, GBIF-Sweden contributes to leading the work of connecting interested world-wide.

2020 Work items

- Convene SYNTHESYS+ workshops and integrate outcomes with relevant Work Programme activities.
- Continue implementation and refinement of the standardized annual information requests from Participant nodes and display up to date content on revised country pages of GBIF.org.
- Lead some and participate in all *alliance* for biodiversity knowledge virtual workshops that will be defined in key areas to increased alignment of GBIF with other networks and infrastructures.
- Manage and improve the virtual conferences infrastructure used by the *alliance* for biodiversity knowledge. Work with steering committee on a governance structure and prioritizing the *alliance* efforts. Expand communication effort through conferences and the ambassador network (minimum €10,000). Provide technical, communication and administrative support for the *alliance*.

2020 Participant plans

- **Argentina:** Continue contributing to national, regional and global activities.
- **Australia:** Review GBIF's data pipelines work and data registry as part of an Atlas infrastructure review and refresh. Aim to align with GBIF where possible.
- **Benin:** Capacity building, data mobilization, data use.
- **Canada:** CBIF will continue to build on the progress made in 2019. We will also contribute through membership on the Steering Committee (via TDWG) for the Biodiversity Summit 2020 in Washington DC, a critical forum for discussion of the goals of GBIF and the *alliance*.
- **Canadensys:** We will continue to help and be involved in the *alliance* as much as we can.
- **iDigBio:** iDigBio plans to continue attending, supporting, and/or hosting regional meetings in North America. iDigBio also plans to continue offering the annual Digital Data Conference as a venue for the regional meetings.
- **Naturalis Biodiversity Center:** DiSSCo is coordinating with GBIF, iDigBio and IGSN to establish a global solution for Digital Specimen identifiers and other infrastructures. In SYNTHESYS+ ELViS development it will be investigated whether linkages can be achieved between ELViS and GBIF collection data (specimen datasets and collection metadata). Naturalis would like to share its developments using AI in species identification and work on using environmental DNA for biomonitoring with the GBIF community.
- **Netherlands:** Continue work programme.
- **Norway:** GBIF Norway contributes staff resources to the DiSSCo National Task Force (NTF) representing Norway in the DiSSCo Prepare process. GBIF Norway maintain liaison with international organization in the agrobiodiversity community to promote and actively support the harmonization with GBIF provided and solutions endorsed standards (see also Activity 2a, 3c and 5b).
- **South Africa:** Further engagements with SANBI-GBIF Node Manager and HoD to evolve the Node planning and Africa portfolio.
- **Spain:** Collaboration with EU Nodes and EOSC to explore best ways to integrate Node's services in the EOSC.
- **Sweden:** GBIF-Sweden will continue to take part in leading and contributing to the development of interconnections at national and international levels (GBIC 2 follow-up etc. incl. participation in SYNTHESYS+).
- **United States:** USGS will continue to be involved in and contribute to advancing and contributing to the ideas around the GBIC2 workshop sharing implementation and infrastructure suggestions as they pertain to the US and the various global communities of practice.
- **Zimbabwe:** Develop and finalize a five year strategic plan and 2020 action plan.

Rationale

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.

Approach

To make national-scale implementation more central across the GBIF network, the Secretariat will consult with the Governing Board on models for unifying aspects of the work of the Nodes Committee into Governing Board meetings. The Secretariat also seeks funds to increase the scope of nodes meetings to include work to develop proposals for fundable activities aligned with this Implementation Plan, particularly by addressing regional priorities. The GBIF Secretariat will also engage organizational Participants to develop specific memoranda of cooperation in key areas and to identify and support alignments with this Implementation Plan and opportunities for joint organization of meetings and workshops.

Activity 1g: Coordinate resources

Tasks

- i. Secretariat to coordinate participant-led activities aligned with Work Programme (process and models to support participant commitments)
- ii. Development of concept notes/proposals for supplementary funding at regional or global scales
- iii. Operate annual Capacity Enhancement Support Programme (CESP) with core funds, aligned with supplementary investments
- iv. Ensure that BID programme and other supplementary funds align with and reinforce CESP and GBIF community, online resources and infrastructure
- v. Align Ebbe Nielsen Challenge and Young Researcher Awards with GBIF priorities

2019 Progress

The initial five-year phase of the BID programme reached its final year with a flurry of activity, closing all of the 63 projects in sub-Saharan Africa, Caribbean and the Pacific supported by the European Union-funded programme. The Secretariat organized closing meetings and training events in all three regions focused on showcasing the achievements of BID, planning regional coordination of future biodiversity data mobilization and use, highlighting the benefits of GBIF participation and establishing a GBIF Participant Node (see Activity 1b). A review of the impacts of BID in terms of data mobilization and use, capacity enhancement and sustainable outcomes, was commissioned following a competitive call and will be finalized later in 2019. An event to close the first phase of BID is being held in Brussels in late 2019, with the aim of summarizing the programme's results to the funders and other stakeholders, as well as setting priorities for future biodiversity informatics development using the model of BID to strengthen and expand data availability in all global regions.

Based on the success of the first phase of BID, the Development and Cooperation Directorate of the European Commission (EU DEVCO, or EuropeAid) expressed interest in extending the programme with additional funding. Details of this extension are nearly complete, but it is expected that GBIF will have flexibility to expand the geographic scope of the programme beyond the three regions targeted in the first phase. The BID event of the first phase of BID will include engagement with

other providers of overseas development assistance to explore the potential to supplement the funding from EU DEVCO for future phases of the programme.

During the April 2019 budget revision, €22,000 was added to hire an external consultant to identify potential funding streams from philanthropic organizations, and to help the Secretariat develop a strategic plan for a funding stream for these programmes. This work commences in late 2019.

The Biodiversity Information Fund for Asia (BIFA), based on supplementary funding provided by the Japanese Ministry of Environment, attracted a strong response in its fourth call for proposals for data mobilization projects. After a thorough two-stage selection process, nine projects were selected for funding. As with the last BIFA call, this year's call replicated the model developed under BID in providing a training workshop, held in Viet Nam, for all funded projects to acquire basic data publishing skills to support project implementation.

A further call under the Capacity Enhancement Support Programme (CESP) was launched in 2019, resulting in selection of five projects for funding, managed by nodes in Europe, North America, Africa and Latin America. Each of these projects contribute to activities in the 2020 work programme items:

- Using the CBD Clearing-House Mechanism to strengthen biodiversity data acquisition and data sharing aims to establish technical links between GBIF national nodes and the network of CBD national Clearing-House Mechanisms (5c. Support biodiversity assessment)
- OpenPSD: Engaging the private sector to promote biodiversity data publication and use will establish deeper connections between the participating nodes and private-sector sources of data (3c. Engage data holders)
- Data Use for Decision Making Workshop: an Iberoamerican community call replicates and applies training curriculum developed under the BID programme outside the ACP regions (1g. Coordinate resources)
- Strengthening Zimbabwe's GBIF node through mentoring by GBIF Spain follows a mandate to improve mechanisms for engaging new countries (1e. Expand national participation)
- Tackling the spatial challenge in the Southern Cone through georeferencing training is applying updated documentation on best practices for georeferencing data (1b. Strengthen skills)

The 2019 calls for GBIF's annual awards programmes—the Young Researchers Award (YRA) and the Ebbe Nielsen Challenge—are attracting many high-quality entries. The results of the YRA are being announced in late August or early September, while the winners of the Challenge incentive prize are being announced during the GB26 meeting.

The GBIF Secretariat has invested in Fluxx, a grants management system. The system will help streamline and automate processes in relation to the management of the grant programmes GBIF is involved in. This should save time for all involved and ensure that the processes are made more efficient. The GBIF deployment of Fluxx is currently in the development phase. A small group of Secretariat staff is testing and offering feedback to the Fluxx developers. Once all issues have been addressed, the testing will be opened up to a wider test group before the final sign-off. The process has somewhat longer than expected as the investment of time from the Secretariat side has been much larger than originally anticipated.

The Secretariat continues to offer support to BID data publishers at the technical and data content

level, which improves data consistency and quality control of all data submissions, assuring adherence to the data publishing and quality guidelines. This work supports the external BID impact study. The help desk supports BID 'cloud' IPT data hosting options for publishers.

2019 Participant contributions

- **Argentina:** Participation, as a jury, of the evaluations of the IDB and BIFA projects. As well as a jury of the Ebbe Nielsen awards. On the other hand, participation in 2 CESP programs (one as a participant, the other as a coordinator). Trainer at the last BID meeting in Caribbean and at the Global Nodes Meeting.
- **Australia:** ALA hosted a week long workshop/hackathon with developers from GBIF Brazil in Canberra (December 2018) to help them in the development of the Atlas based portal for Brazil.
- **Benin:** We coordinated the implementation of the BID regional and national projects. In the framework of BID Regional project, we coordinated data mobilization and data uses in a consortium of 6 African countries. In the framework of a CESP program with Nigeria, we received and trained in Benin, the GBIF node of Nigeria.
- **Canadensys:** We have participated in the five-year CESP review but we have not applied for funding during the 2019 call. The CESP program is a great incentive to promote projects between different nodes.
- **Consortium of European Taxonomic Facilities:** CETAF is a driving force behind the European DiSSCo Initiative, which aims at creating a Europe-wide research infrastructure to mobilize, unify and deliver bio- and geo-diversity information. The CETAF Secretariat is actively participating in several projects related to the implementation of DiSSCo, inter alia SYNTHESYS+ and MOBILISE.
- **Germany:** Several German GBIF Node institutions actively participate in projects contributing to the European DiSSCo Initiative, that aims at creating a Europe-wide research infrastructure to mobilize, unify and deliver bio- and geo-diversity information. Strong engagement also in CETAF, GGBN and BioCASE.
- **Netherlands:** Evaluated CESP 2019 call.
- **South Africa:** Further engagement with regards to funding allocation will continue. The Africa Coordination Mechanism (ACM) 5 year strategy is seen as a catalytic investment. Sustainability will be built into the 1st phase, so there will be less concern about what happens subsequently, as this would have been built into the model. The ACM Business Case and Regional Engagement Strategy is a model that could be adopted in other regions of the world. In the 1st three years, through the ACM, GBIF-Africa, through SANBI-GBIF will explore the need and feasibility of developing a bigger platform that can provide for a strengthening of a community of practice for all African biodiversity informatics initiatives, to enhance and make more efficient use of the data. This will be done through leveraging additional regional and global partnerships and initiatives. Data mobilization funding through FBIP large and small grants ± €645,000 will continue in 2019. 4. ABC Prize money to be awarded in 2019.
- **Sweden:** Contribution by staff participation in-kind at meetings, conference sessions, workshops etc.
- **Zimbabwe:** Applied and obtained CESP funding for GBIF Zimbabwe in collaboration with GBIF Spain.

2020 Work items

- Launch a 2020 call for proposals under a renewed CESP (minimum €80,000). If funding allows, this will be supplemented by a special category for regional outreach activities that target increased data mobilization in non-GBIF countries as recommended by Nodes Steering Group ((see Activity 1e).
- The Secretariat will continue with a BID call for proposals that responds to feedback received from community consultations, three regional closing meetings and the BID phase 1 closing meeting.
- Implement a fifth call for proposals (approximately €110,000) under the Biodiversity Information Fund for Asia (BIFA), reflecting priorities agreed by Asian nodes.
- Continue collaboration with the BioDATA (Norway > Eurasia) and Russia support (Finland > Russia) supplementary funding programs.
- Develop and implement strategic plan for targeting external funding streams for capacity enhancement projects. This will be based on 2019 external consultation and using the results of BID Phase 1 closing meetings. If budget allows an additional (€22,000) will be allocated to this effort. The strategy will work with all GBIF regions to target potential funding streams that support additional capacity enhancement for data mobilization and use, building on the BID and BIFA models.
- Make refinements to the newly selected grant management system, FLUXX, to streamline project calls, assessment, selection, implementation, budgeting and reporting, to ensure it meets the needs of the community.
- Launch 2020 calls for the Ebbe Nielsen Challenge and Young Researchers Awards (€44,000), considering recommendations from the Science Committee based on the 2019 programmes. Explore the possibility of soliciting additional donor/sponsor co-funding.
- Begin Secretariat planning to establish a workflow using digital documentation to develop the 2021 work programme and 2022–2026 strategic plan. This planning will address the recommendations of the 2019 20-year review.

2020 Participant plans

- **Argentina:** Participation in 2 CESP programs (one as participant, one as coordinator), coordinating national co-financing for CESP, coordinated from Argentina.
- **Belgium:** Help establishment of a second phase of BID in collaboration with the European Union
- **Benin:** Continue to exert the mentorship among partners
- **Canadensys:** We will not lead a proposal for a CESP grant in 2020, but, depending on the needs, we could act as mentor in another proposal. We are also opened to act as mentor for a project in the BID program, but we do not have a project in mind currently.
- **Consortium of European Taxonomic Facilities:** Participation in DiSSCo-related projects continues.
- **Germany:** Participation in DiSSCo-related projects, CETAF, GGBN and BioCASE continues.
- **iDigBio:** iDigBio will continue its integration and coordination efforts with GBIF, ALA, and

DiSSCo.

- **Norway:** The GBIF CESP OpenPSD project developed by the GBIF Nodes from Spain, Portugal, Norway, Colombia, and France will engage and document best practices for mobilizing biodiversity data from the private sector (see also Activity 1d, 3b and 3c). All of the Nordic and Baltic GBIF nodes contribute to the Nordic-Baltic collaboration on e-infrastructures for Biodiversity Informatics (DeepDive2 project proposal) funded by the Nordic e-Infrastructure Collaboration (NeIC) of the Nordic Research Council (NordForsk) (see also Activity 2a, 2b, 3b, 4a, 4b and 4c).
- **South Africa:** Develop funding strategies to support the Africa Coordinating Mechanism Business Case. Continued data mobilization funding through FBIP large and small grants of ± €645,000 will continue in 2020. The ABC2 project has been approved by the JRS and will be initiated in 2019 and continue into 2020.
- **Spain:** GBIF Spain will be involved in 3 CESP projects that will take place from June 2019 to June 2020. We will coordinate the OpenPSD CESP project to mobilize data from private sector. We will be available to help in the second phase of BID programme if this takes place.
- **Sweden:** GBIF-Sweden will continue to contribute by offering opportunities for its staff members to participate in coordinated activities aimed at expand the scale and scope of GBIF.
- **Zimbabwe:** Implementation of CESP project.

Rationale

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.

Approach

This Implementation Plan itself provides a framework for organizing information on Participant and supplementary fund resources to complement GBIF's core funding. The GBIF Secretariat seeks information from all Participants on planned activities and commitments which may help to advance the work of other Participants or GBIF globally – examples may include funding for workshops; committed resources to develop tools, standards or best practices; mentoring actions; etc. These will be recorded as part of the overall GBIF Work Programme and updates will be presented in the GBIF annual report. A shared Implementation Plan also offers the opportunity for GBIF (either the Secretariat or Participants) to develop concept notes to seek supplementary funding to target currently unfunded or underfunded areas. The annual Capacity Enhancement Support Programme budget includes only limited funds, but existing supplementary funds (in particular BID and BIFA) align with these funds and leverage CESP tools and processes. Future supplementary funds should follow a similar model.

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

Tasks

- i. Promote development of a shared domain model for sharing and linking all components of biodiversity information
- ii. Lead a review of the Darwin Core vocabulary and associated extensions to ensure consistency and full alignment with a shared domain model
- iii. Explore opportunities to increase accessibility of biodiversity data through evolution of Darwin Core Archive formats to W3C CSV on the Web formats
- iv. Explore models to enable GBIF and other biodiversity infrastructures to deliver comprehensive global catalogues of instances of key data classes
- v. Improve management of trait data of relevance to GBIF

2019 Progress

The alliance for biodiversity knowledge is beginning to act as a platform to engage the biodiversity informatics community around community standards. GBIF is active in numerous significant existing networks that seek to address these needs in parallel. This work is ongoing into 2020 and beyond.

At the core of this is the Biodiversity Information Standards (TDWG) community. GBIF continues to participate in open TDWG discussions around [ABCD/DwC alignment](https://doi.org/10.3897/biss.3.37491) [https://doi.org/10.3897/biss.3.37491] and recognize many other complementary activities and discussions on a biodiversity knowledge graph by partners. Notably:

- Through Plazi taxonomic treatment data are becoming connected
- Pensoft have piloted the [OpenBioDiv system](https://doi.org/10.3390/publications7020038) [https://doi.org/10.3390/publications7020038] to better connect information in scholarly publications with specimens
- The [Ozymandias biodiversity knowledge graph pilot](https://ozymandias-demo.herokuapp.com) [https://ozymandias-demo.herokuapp.com] by Rod Page

Proposed SYNTHESYS+ workshops are to focus on modernizing standards activities that seek to improve the representation of information such as

1. Citation and Provenance models for collections data, including links between specimens, individuals and literature and their relationships to DOI, ORCIDs and other important open

identifier issuing systems

2. Information model for representing Natural History Collections including the TDWG natural Collections Descriptions standard and integrations with other collections catalogues
3. Information model for representing Specimens; reviewing collection management systems and developing an information model with data classes representing all the asset types and linked information of importance in building a fully interconnected virtual natural history collection

During 2019, the informatics team has continued the redesign and implementation of GBIF data ingestion pipelines. Data growth required significant changes to the backend to ensure GBIF can 1) continue to grow with data volume, 2) accommodate new feature deployments that require full data reprocessing, and 3) look to expand data content types.

2019 Participant contributions

- **Argentina:** Promote development of a shared domain model for sharing and linking all components of biodiversity information.
- **Australia:** “Work is ongoing, and the Atlas continues to contribute to the work of the TDWG Biodiversity Data Quality Interest Group and Citizen Science Interest Group. The Atlas continues to be a seminal contributor to the international collaboration project which is developing and implementing the PPSR-Core data and metadata standard for citizen science. This project aims to provide a robust mechanism for data and metadata standardization and exchange for the citizen science domain, which is also compatible with, and leverages, other existing standards and protocols.
- **Canada:** CBIF plays an active role in developing, reviewing and implementing standards. Staff contribute to the Collections Description (CD) standard, Darwin Core, Taxon Concept Schema (TCS), Data quality and controlled vocabulary standards, annotation and attribution standards, genomic standards, and others. Dr. Macklin is the current chair of TDWG and is working closely with GBIF to insure alignment to GBIF requirements and to seek funding to further mutual work.
- **Germany:** Two German GBIF Nodes finalized the new Version of the TDWG Standard ABCD (Access to biological collection data, <https://abcd.tdwg.org/3.0/>). The Botanical Node houses the technical Secretariat of the Global Genome Biodiversity Network, significantly contributing to standard development for DNA- and Tissue samples.
- **iDigBio:** “Several representatives of iDigBio participated in the GBIC2 conference. iDigBio, GBIF, ALA, and DiSSCo are making a concerted effort to coordinate efforts. Several iDigBio staff are involved with the Natural Collections Description (NCD) group.
- **Naturalis Biodiversity Center:** Naturalis participates in modernizing data standards through TDWG and active participation in the TDWG CD taskgroup.
- **Netherlands:** NLBIF is engaged in TDWG CD (Collection Description) which will likely feed into GRBio/GRSciColl which is adopted by GBIF.
- **Norway:** The wider Norwegian GBIF community contributed to the implementation of the sampling event data model for environmental monitoring and survey-based data with focus on national implementation while contributing to the international standardization process (see also activity 3b).

- **Spain:** We have started to explore how to serve species information under the Plinian Core Standard in the Living Atlas platform. Involved in the TDWG Species Information Interest working group.
- **Sweden:** Being part of a national research infrastructure for biodiversity data provisioning, GBIF-Sweden has started working on taxonomic and other data standards for various kinds of “new” data types (molecular data, tracking data, sensor data).
- **United States:** Continued to advance sampling event with extended measurement or fact extension for documenting data associated with biological observations but not appropriate for occurrence only Darwin Core.

2020 Work items

- Modernizing data standards is a continuous Work Programme activity for a global infrastructure like GBIF. During 2020 we will focus on advancing and refining data models for Collections, Taxonomic Treatments, Sampling Events, Organisms, Specimens, Organisms, Citations and the linkages between them.
- Provide a set of data-exchange profiles for sharing data within GBIF that conforms with a unified information model that includes both existing and new standards as well as the necessary controlled vocabularies.
- Redesign the GBIF Integrated Publishing Toolkit (IPT) to support these profiles and to address infrastructure needs, such as the ability to support local installations or GBIF-hosted solution. If funds allow, €50,000 for an external contractor.
- Provide documentation for the data model and for the associated services offered through GBIF.org.
- Review and redesign GBIF data management system to accommodate the unified information model as part of data ingestion, quality control and processing where necessary.
- Continue technical discussions with other data aggregators to seek closer alignment in practice and, as far as possible, implementation of aggregation and indexing processes.
- Demonstrate improvements of information in GBIF.org and hosted national portals in specimen-level information, links to material citations, and links between specimens and sequence data from sources such as BOLD.
- Explore approaches for adding a phylogenetic/evolutionary dimension to the GBIF taxonomic backbone. Pilot phylogenetic browsing capabilities of occurrence data.
- Open discussion with GB participants to provide project funders with an overview of the resulting value relating to their investment (e.g. data mobilization, publications).
- In collaboration with international partners, explore the desirability and scope of “catalogue services” that are targeted specifically at physical specimen collections. Examples could include displaying duplicate or derived specimens across collections, type information, citations in taxonomic treatments and trait data.
- Explore options for displaying occurrence data from long-term sampling sites, piloting with projects like BIOSCAN 2 and/or Norwegian ecological datasets.

2020 Participant plans

- **Andorra:** It is little bit hard for us to contribute in improve the participation on Building of data standards. Nevertheless we are open to adapt our data to the newest standards.
- **Argentina:** Promote development of a shared domain model for sharing and linking all components of biodiversity information.
- **Australia:** Align with international projects in establishing and using standardized tests and reporting.
- **Belgium:** Help documenting a unified information model that covers the scope of GBIF content.
- **Biodiversity Heritage Library:** Review options for implementing IIIF.
- **Canada:** CBIF will continue to contribute to the standards work outlined in the 2019 progress report.
- **Germany:** ABCD 3.0: German GBIF Nodes contribute to a new working group which aims to integrate ABCD and Darwin Core. Continued activities in GGBN.
- **iDigBio:** iDigBio will continue its integration and coordination efforts with GBIF, ALA, and DiSSCo. iDigBio will continue to support the Natural Collections Description (NCD) group.
- **Naturalis Biodiversity Center:** DiSSCo seeks to join forces with GBIF and other infrastructures to work on interoperability standards for natural scientific collections. Naturalis plans to revive the BD Integration IG in RDA and to involve the the GBIF community in rewriting the chapter to create recommendations for modernizing GBIF community data standards in a multidisciplinary setting.
- **Netherlands:** Continue TDWG CD activities.
- **Norway:** Dependent on continued stable funding for GBIF activities in Norway, GBIF Norway will contribute to modernize and expand the support in GBIF for new data types for genome and eDNA data linked to “material samples”; and data types for ecological data sets based on the “sampling event” model (see also Activity 3b).
- **Sweden:** GBIF-Sweden will further work on taxonomic and other data standards for various kinds of “new” data types (molecular data, tracking data, sensor data).

Rationale

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.

Approach

GBIF will work with TDWG and other key stakeholders to review existing solutions for a common domain model, working towards agreement on a model to adopt with key partners. This conceptual model should cover the main components of biodiversity information (the domain “classes” such as Specimen, Collection, TaxonName, TaxonConcept, Publication, Sequence) and document the mandatory and recommended properties expected for each component and the vocabularies that should control the properties. A review of existing vocabularies and their current uses will be undertaken and revisions and new vocabularies will be proposed where necessary. A revision of the Darwin Core Archive mechanism and supporting tools, such as the publishing toolkit (IPT) and the data validator, will be undertaken to accommodate the richer content model and the new recommendations from the W3C CSV on the Web working group. GBIF should continue discussions with other key global biodiversity data infrastructures to develop comprehensive catalogues to support discovery and normalization of instances of the most critical domain classes (particularly TaxonName, TaxonConcept, Collection, Specimen, TaxonOccurrence).

In addition to completing this knowledge graph, GBIF should be equipped to link between people, datasets, cited use and funding agencies through the correct attribution chains using e.g. Digital Object Identifiers (DOIs) and Open Researcher and Contributor ID (ORCID) as potential mechanisms.

Activity 2b: Deliver names infrastructure

Tasks

- i. Partner with other biodiversity informatics initiatives and taxonomic database holders to plan and deliver a comprehensive nomenclatural dataset and working consensus classification for all life
- ii. Promote publication of species checklists through GBIF network
- iii. Explore potential use of checklists to assist with data validation or derive augmented data products
- iv. Explore integration of Linnaean nomenclature of formally described taxa with provisional names and species hypotheses and OTU naming

2019 Progress

A new IT infrastructure for building and maintaining GBIF backbone taxonomy is near completion. This new infrastructure came from a collaboration with the Catalogue of Life (COL) and includes a clearinghouse for nomenclature and taxonomy. This infrastructure was designed to replace the GBIF Checklistbank. This infrastructure will be used by GBIF and COL and is available service for other initiatives. GBIF is working with COL to provide tools, including a web-based console for editors and technicians. The checklist is being monitored and will replace the GBIF backbone when suitably complete in coverage. The design and deployment of the new public Catalogue of Life website is scheduled for later in 2019.

The adoption of Operational Taxonomic Unit (OTU) identifiers from the UNITE fungal database and the Barcode of Life Database (BOLD) in the Catalogue of Life is scheduled for the second half of

2019 Participant contributions

- **Argentina:** Promote publication of species checklists through GBIF network, as a coordinator of the CESP2018-011: 11 nodes/institutions work together to develop a manual to improve the publication of checklist. Also to boost the publication of national checklist.
- **Australia:** Continued work on improving access to the Australian checklists. These data are available now in Darwin core archives to feed into Catalogue of Life plus.
- **Benin:** “Many checklists have been published on medicinal plants and agroforestry species:

1. <https://www.gbif.org/dataset/81a257d9-9522-410e-b921-8c9283dbb939>
2. <https://www.gbif.org/dataset/6ac09c4d-bf7b-4e47-9c2d-f5abf6e89aa0>
3. <https://www.gbif.org/dataset/dba21dbc-ea38-4bbf-b50f-ba3d256e1f73>

Many checklists published also concerned, invasive alien species and threatened species:

1. <https://www.gbif.org/dataset/3d4ac01a-51a9-4ec9-8599-230a27e58758>
2. <https://www.gbif.org/dataset/de5e3fb2-3922-40ba-9773-e45a6850ffeb>
3. <https://www.gbif.org/dataset/0a20bf65-4c0f-42ce-b6e4-2218c0ea5833>

- **Biodiversity Heritage Library:** Participated in GlobalNames Workshop and Catalogue of Life Plus meetings.
- **Canadensys:** The Canadensys team has maintained Vascan (<https://data.canadensys.net/vascan/>), the Database of Vascular Plants of Canada, which is used in the GBIF backbone, and is an important resource for biologists and botanists in Canada and elsewhere. Vascan continues to be actively curated and up to date based on recent publications.
- **Colombia:** Publication and update of different national species checklists, supported by specialist groups and biological collections.
 - Birds: https://ipt.biodiversidad.co/sib/resource?r=aco_listaavescolombia2017
 - Mammals: https://ipt.biodiversidad.co/sib/resource?r=mamiferos_col
 - Freshwater Fishes: https://ipt.biodiversidad.co/sib/resource?r=ictiofauna_colombiana_dulceacuicola
 - Endangered species: <https://ipt.biodiversidad.co/sib/resource?r=resolucion1912-2017mads>.
- **France:** Update of TAXREF, the French national checklist.
- **Germany:** A project implementing the registration system for algal names (PhycoBank) was concluded and the fully functional application is now being tested within the phycological community. The global Caryophyllales Network is coordinated at the Botanical Node. It will produce a complete taxonomic backbone for this order, comprising about 5% of flowering plants.
- **Integrated Taxonomic Information System:** Through the Catalogue of Life Partnership and through the CoL Global Team, ITIS was integral to the planning and implementation of the Catalogue of Life Plus initiative.
- **Japan:** Additional data and revision to be continued. Endangered Species (National and local in

Japan) checking application to be developed.

- **Mexico:** Consolidation of the first two national species lists for GBIF Mexican node: “Aves de México” and “Helmintos parásitos de vertebrados”. Publish in process through participation in CESP2018_011 “Increasing capacities to develop National Species Checklists in the Latin America and the Caribbean Region”. Website snib.mx development include lists of species with distribution in Mexico for download. A total 102, 866 valid or accepted species and 8,543 valid or accepted infra species.
- **Naturalis Biodiversity Center:** Naturalis is working together with GBIF and NLBIF to coordinate and carry out work under the NLBIF-funded Catalogue of Life Plus (CoL+) project, led by Olaf Banki. The existing processes for constructing the monthly and annual Catalogue of Life checklists and for constructing the GBIF taxonomic backbone have been replaced with a single solution that delivers both products (together forming a “provisional checklist”). This new infrastructure is currently in beta testing and will be taken in production in 2019. The new infrastructure includes handling of existing GBIF Checklistbank capabilities (registry integration, images, descriptions and the infrastructure has be developed to have updates to the provisional checklist being reflected directly in the GBIF data index and vice versa. Naturalis is working with Catalogue of Life to develop and increase participation of a a responsive expert community for the names infrastructure.
- **Netherlands:** NLBIF is one of the main sponsors of CoL+ project.
- **Spain:** Publication of several checklists with national and regional scope.
- **Species 2000:** The Catalogue of Life + steering committee consists of the following parties: Barcode of Life Data Systems, Biodiversity Heritage Library, Encyclopedia of Life, Global Biodiversity Information Facility, and Species 2000 / ITIS / Catalogue of Life, and Naturalis Biodiversity Center. Recently the Distributed System of Scientific Collections joined the steering committee and discussion with Lifewatch to join are ongoing. The steering committee will become part of the Board of Directors of Species 2000 to jointly govern the new Catalogue of Life Infrastructure. The COL+ project team consists of representatives of Species 2000, GBIF Secretariat, Naturalis Biodiversity Center and the Illinois Natural History Survey. In 2019, the new tooling for the assembly of the Catalogue of Life will go in production by the 4th quarter. The monthly editions will be resumed in the new Catalogue of Life infrastructure and persistent, unique identifiers for names will be part of the new infrastructure. By the end of 2019 all functionality of the GBIF Backbone Taxonomy will be integrated in the new Catalogue of Life infrastructure. A new public portal will be in beta version by the end of 2019. The infrastructure will be hosted by GBIF. In 2019, the requirements for scientific names and taxonomic services of all COL+ consortium partners as well as other important stakeholders will be captured. Also the use of the COL in scientific publications will be analysed. GBIFS and Species 2000 had discussions with the European Environmental Agency concerning taxonomic backbone services, and in the context of the Synthesis+ project will continue these discussions together with DiSSCo. GBIF and Species 2000 are working together with the Museum für Naturkunde in Berlin to organize a workshop to gather the taxonomic expertise to build the best global resource for Lepidoptera taxonomic data. Together with the Conservatory and Botanical Garden of the City of Geneva and the World Flora Online Consortium a workshop is planned to gather the taxonomic expertise to build the best global resource for plant taxonomic data. Species 2000 and GBIF are organizing a symposium on getting towards a joint global names and taxonomic infrastructure at the Biodiversity_Next conference in October in Leiden, The Netherlands.

- **Sweden:** By synchronizing the national Swedish Dyntaxa taxonomic backbone with the Catalogue of Life, GBIF-Sweden contributes to the enhancement of the global names infrastructure.

2020 Work items

- Maintain and update processes for constructing the GBIF taxonomic backbone, 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 the Catalogue of Life.
- Implement a process enabling key checklists to be used in filtering occurrence data, such as Red Listed species and invasive alien species.
- Consult with relevant regulatory agencies, such as the European Environment Agency (EEA), for guidance on which legislative checklists should be incorporated to increase the relevance of COL+ to governments.
- Explore feasibility of supporting national taxonomies for exploring GBIF occurrence data to better enable national level reporting.
- Develop and pilot a process that allows qualified users to collaborate and edit sectors that contribute to the GBIF backbone taxonomy, aimed at reducing the delays before such edits appear on occurrence records from months to days.

2020 Participant plans

- **Argentina:** Keep working to publish more checklist at the nodes related on the CESP2018-011 and any other with interest.
- **Australia:** Further work to improve the currency of taxonomic information in the Atlas based on the Australian checklists.
- **Benin:** Continue previous work programme.
- **Biodiversity Heritage Library:** Continue to participate in GlobalNames Workshop and Catalogue of Life Plus meetings. Implementation of new Global Names services in BHL.
- **Canadensys:** We would strongly encourage CoL+ to continue taking into consideration the Canadian expertise for vascular plant taxa.
- **Germany:** Updated list of fungi and fungal-like organisms from Germany compiled by the German Mycological Society (DGfM) available via GBIF. Algal names from the PhycoBank algal registrations system available via GBIF. Application to General Nomenclature Committee to recognize PhycoBank as a global repository for algal names. Complete taxonomic backbone for Caryophyllales available, inter alia through World Flora Online.
- **Integrated Taxonomic Information System:** ITIS is developing and will deploy in 2020 an online taxonomic workbench that will allow for the development of taxonomies based on expert communities. This effort will support taxonomic sectors which currently lack adequate support and will improve alignment with other checklist efforts. This is part of the ITIS' commitment to the CoL+ (GBIF's names infrastructure.)
- **Japan:** Improvement of training data for Endangered Species to be revised.
- **Mexico:** About 8 new national checklist: Phengodidae, Lycidae, phytoplankton (Pacific Ocean),

ants, amphibians and reptiles, Lamiaceae, and echinoderms. Comparing Catalogue of Mexican species vs. Catalogue of Life CoL 2018, only 33.5% of species and 18.42% of infra species in the CoL with distribution in Mexico.

- **Naturalis Biodiversity Center:** Work on CoL+ will be continued early 2020 to provide a end-user interface to the renewed Catalogue of Life infrastructure and to replace the CoL website with a new one to be hosted by GBIF.
- **Netherlands:** NLBIF continues the contribution to the CoL+ project and the development of the new CoL infrastructure to serve as taxonomic backbone for GBIF, DiSSCo and aligned projects.
- **Spain:** Intend to publish national list of invasive species and regional species lists from natural parks.
- **Species 2000:** A long term vision for Catalogue of Life+ as incubator project for the *alliance* for biodiversity knowledge will be developed. This will result in the scoping of a second phase of the Catalogue of Life Plus project for which funding will be sought. This second phase will likely focus on empowering the taxonomic community to make better use of the Catalogue of Life and implement taxon concept identifiers. The second phase will also encompass the needs for names and backbone services of the COL+ consortium partners and other key stakeholders as best as possible. Special attention will be made in linking DNA barcode information and the Catalogue of Life in discussion with the International Barcode of Life and GBIF amongst others.
- **Sweden:** More taxonomic names and concepts (esp. related to fungi and procaryotes) will be included in the set of services offered by GBIF-Sweden.
- **Switzerland:** Publication of national species checklists for red list groups and important invertebrate groups.

Rationale

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.

Approach

GBIF and many other partners have worked on this challenge and much progress has been made, but we are still far from a comprehensive shared solution. GBIF has been in discussion with

Catalogue of Life, EOL, BHL, BOLD Systems, nomenclators and others about pooling resources to deliver the best possible complete nomenclator and catalogue of all species, along with improved tools to enable the taxonomic community to own and maintain these resources more effectively. The challenges are not primarily informatics issues. The most important requirement is to understand the constraints and needs of existing content holders and the features that are required from an infrastructure that can be embraced by the majority of taxonomists. The solution must build on existing initiatives and give sufficient credit and benefit back to those who have invested in developing data. It must be flexible enough to accommodate existing well-managed datasets without disrupting their activity and to accommodate more open mechanisms to support wide community input for taxa which need more work. In the longer term, it should support evolution towards ownership of curation responsibilities by international taxonomic societies or other bodies recognized by researchers for each group. The infrastructure should include processes to review and interpret unrecognized name strings found by GBIF and others in aggregated data. Once these requirements have been resolved, implementation must rapidly follow to offer these resources as open public datasets for use by all.

The Netherlands has coordinated a significant commitment for 2017 and 2018, led by NLBIF and including resources both from Species 2000 and Naturalis. This funding will enable GBIF and partners to direct significant effort to this area over the period.

Activity 2c: Catalogue collections

Tasks

- i. Partner with other biodiversity informatics initiatives to deliver a single, comprehensive catalogue of the world's natural history collections
- ii. Use collection metadata as a first stage in content mobilization from natural history collections, including identification or development and adoption of necessary data standards

2019 Progress

In 2019 the [Global Registry of Scientific Collections database](https://gbif.org/grscicoll) [https://gbif.org/grscicoll] (GRSciColl) was migrated to a public website managed by GBIF and editing capabilities provided through the enhanced, and open, registry administration console. The migration and resurrection of GRSciColl identified that there are several other catalogues (CETAF, IndexHerbariorum) that need consideration; GRSciColl needs to evolve to synchronize with authoritative catalogues such as Index Herbarium and collection management systems directly. Therefore, GBIF continues involvement in the TDWG task group on the development of a collections description data standard (TDWG CD). This work continues with community consultation under SYNTHESYS+ and the alliance for biodiversity knowledge to plan future work.

Identifiers have been used over the years by the community and are referenced in 42 million records. These identifiers now resolve again, e.g. <http://biocol.org/urn:lsid:biocol.org:col:34994>.

2019 Participant contributions

- **Argentina:** Since 2003 Argentina have a catalogue of institutions and collections and keeps growing.

- **Benin:** Three metadata have been published in the framework of our joint CESP project with GBIF France
 1. Inventaire et dénombrement des oiseaux du Parc Naturel Communautaire de la Vallée du Sitatunga (Sud Bénin) <https://www.gbif.org/dataset/3194e21c-447a-410d-bb09-31398482de1f>
 2. List of plant species of forest formation in Benin <https://www.gbif.org/dataset/d882e391-10f9-43f6-b0f9-29636396ba9e>
 3. Inventory of plant species produced in CERF nurseries <https://www.gbif.org/dataset/f7598557-adb5-4827-9c28-8025da802dc4>
- **Canada:** CBIF staff contribute to the Collections Description standard.
- **Canadensys:** Metadata on Canadian collections are relatively up to date on the Canadensys website (through the IPT and the ALA framework)
- **iDigBio:** Several iDigBio staff are involved with the Natural Collections Description (NCD) group. iDigBio maintains a comprehensive list of known biodiversity collections in the United States.
- **Japan:** Preliminary questionnaire provided to data providers for future stock of data.
- **Mexico:** “Colecciones Biológicas científicas Mexicanas” <http://www.biodiversidad.gob.mx/especies/colecciones/> is a web portal with 747 metadata of scientific Mexican collections in 237 national institutions. The collection metadata was developed with standard international Natural Collections Description (NCD).
- **Naturalis Biodiversity Center:** Naturalis led the development of the DiSSCo Prepare proposal which includes preparations towards a catalogue of the European collections. As part of a project financed by the Dutch research council (NWO, ALWIN.005), Naturalis together with NLBIF, and the foundation for Dutch Natural History Collections has created a collections dashboard for the Netherlands. This dashboard is the first online representation of the natural history collections in the Netherlands that are digitized as well as not yet digitized. The dashboard will be published in the country page of DiSSCo.eu and NLBIF.nl.
- **Netherlands:** NLBIF is part of TDWG NCD that furthers the CD (Collection Description) standard <https://www.tdwg.org/community/cd/> to describe natural history collections. A preliminary standard has been implemented in a dashboard that shows the Dutch national collection as part of the NLBIF activities under DiSSCo-NL.
- **Spain:** We maintain, improve and keep updated the Spanish registry of biodiversity collections and databases that currently hosts 230 institutions and 471 collections/databases. This metadata repository is a first step for data mobilization and it is available at <https://www.gbif.es/registro-instituciones/>.
- **Sweden:** In collaboration between The Swedish Museum of Natural History GBIF-Sweden runs the development of a web-based collection management system (DINA web).
- **Switzerland:** Contribution to the report of the Swiss Academy of Sciences: National significance of natural history collections in Switzerland.
- **United States:** Continued significant investment and contribution from ABCD, iDigBio including hosting an annual conference and working directly with collection managers.

2020 Work items

- Based on community consultation, build mechanism to synchronize Global Registry of Scientific Collections (GRSciColl) with other catalogues.
- Improve linkages between collections, institutions and occurrences (or specimen) objects indexed by GBIF.
- Develop the user interfaces and services necessary to support a collection catalogue system.
- Work with the community to ensure the content is fit for use, and promote community editing of the registered content.

2020 Participant plans

- **Argentina:** Keep working with our catalogue and collaborate with the GRBio / GRSciColl initiative presented at GBIF.
- **Canada:** CBIF will continue to contribute to the Collections Description standard.
- **Canadensys:** If needed, we can help review and update information about Canadian Collections on the Global Register of Biodiversity Collections (GRBio)/Global Register of Scientific Collections (GRSciColl) dataset. We will maintain the collections' metadata up-to-date on Canadensys.
- **iDigBio:** iDigBio will continue to support the Natural Collections Description (NCD) group. iDigBio will continue maintaining its list of U.S. collections. Ultimately, iDigBio's goal is to integrate this list with the result from the GBIF/GRBio/GRSciColl collaboration.
- **Japan:** Evaluation of questionnaire to know/analyses the status.
- **Naturalis Biodiversity Center:** Naturalis is planning to establish a proto-version of the future DiSSCo collections catalogue and is investigating to use GBIF as one of the authoritative sources of data. For the catalogue a minimum information specification for collections (MICS) needs to be developed in conjunction with work on the TDWG CD standard and based on the work already done in ICEDIG.
- **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.
- **Sweden:** Expanding DINA web to cover not only biological objects (including sequence data) but also palaeontology and geology will continue in 2019.
- **Switzerland:** Update of Swiss GRSciColl records. Promotion and support of collection metadata acquisition.
- **United States:** ABCD, iDigBio to continue work in advancing and facilitating sharing of catalogue collections.

Rationale

Natural history collections are the largest source of data on biodiversity outside nature itself. Collectively the worlds 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.

Approach

GBIF will collaborate with existing biodiversity collection registries such as GRBio to develop a rich collections catalogue to facilitate the discovery and use of the world's collections.

Collections will be guided and facilitated to publish metadata-only datasets by drawing on the recommendations of the GBIF task force on accelerating the discovery of bio-collections data. The generation of a collection catalogue through metadata assessments will help establish roadmaps for further digitization of the collections.

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

Tasks

- i. Organize and integrate requirements identified by fitness-for-use groups
- ii. Establish open mechanism for researchers and users to document specific data needs
- iii. Develop assessments and visualizations of strengths and weaknesses of GBIF data in key dimensions (taxonomic, geographic, environmental, temporal)
- iv. Maintain well-publicized GBIF priorities for data publishing, incorporating needs of IPBES and other networks
- v. Provide Governing Board with annual review and progress assessment for digitization of major natural history collections
- vi. Provide GBIF Participants with actionable priorities and targets for content mobilization

2019 Progress

During 2019 the Secretariat is developing actionable guidance for nodes, publishers and funders on addressing priority gaps through mobilization targets and strategies based on spatial, temporal, taxonomic and thematic dimensions of biodiversity data. This guidance document will indicate best practices on how to prioritize mobilization to develop national priorities and how to integrate those priorities into a mobilization strategy. The guide content undergoes assessments that can be updated upon knowledge gained for the GBIF communities in activities such as the global node

events.

The Secretariat performed [data analyses](https://data-blog.gbif.org) [https://data-blog.gbif.org] building on the input from the 2016 Ebbe-Nielsen Challenge. It further examined possible approaches of ‘hunger-mapping’ to the prioritization of data gaps, as well as analysing the availability and coverage of regional and national checklists of taxa that could support this approach, and the possibility of making use of ES50 maps as a measure of species richness.

2019 Participant contributions

- **Argentina:** We started working with the environment secretary to publish the national biodiversity inventory and national species lists. On the other hand, it is expected to feed the national portal with these national listings.
- **Australia:** Recommend that the community works collaboratively to define targets around coverage and extent of our data products (e.g. temporal/ spatial and taxa). This will help define the level of effort required for data publishers to met these gaps. Also define level of use from Occurrence to Abundance to True Absence.
- **Benin:** Many data gaps are being identified and will be filled in the framework of the master program in biodiversity informatics
- **Colombia:** Santander BIO and Boyacá BIO. two projects to promote knowledge, conservation, management and sustainable use of biodiversity in these territories. With the joint work of several research institutes and organizations these projects made a large amount of data available through the SiB Colombia in regions of the this departments previously difficult to access. In this way it has been possible to fill data gaps in Colombia through these regional initiatives. Santander BIO and Boyacá BIO Datasets: <https://www.gbif.org/dataset/search?q=Santander%20BIO> <https://www.gbif.org/dataset/search?q=Boyac%C3%A1%20BIO>
- **iDigBio:** iDigBio hosted an Invertebrate Collections Digitization workshop in February 2019. Participants discussed best practices and considered ways to improve on existing methods to improve efficiency and speed of data mobilization. iDigBio hosted a Digital Future of Entomology symposium at the 2019 Eastern Branch EntSoc meeting. Attendees learned exactly what digitization means and discussed current and future prospects of digitization in entomology.
- **Integrated Taxonomic Information System:** ITIS has developed numerous taxonomies to close know taxonomic gaps and has provided them to the global community as part of ITIS' commitment to names infrastructure.
- **Japan:** Research program on gap analysis for endangered species to be continued.
- **Netherlands:** The NLBIF call 2018/2019 has funded a project to share data from the Ukrainian UkrBin database with GBIF.
- **Norway:** GBIF Norway was during 2019 an active partner of Nordic crop wild relative projects and international agrobiodiversity processes to represent and promote implementation of GBIF infrastructure technology and solutions (see also activity 3a, 4a, and 5b).
- **Spain:** We have mobilized data from public administrations and marine data.
- **Sweden:** By developing standards for new data types GBIF-Sweden has been able to reach out to previously less well covered data providers.

2020 Work items

- Continue work on items initiated in late 2019, with an emphasis on developing actionable guidance for data publishers and nodes, integration of user needs into prioritization for data mobilization, and data search analysis.
- Continue to improve visualizations of GBIF-mediated data that identify gaps by engaging in interactive community consultations. This work is a candidate for curated discussion through the alliance for biodiversity knowledge prior to implementation on GBIF.org.

2020 Participant plans

- **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.
- **Benin:** Many data gaps are being identified and will be filled in the framework of the master program in biodiversity informatics.
- **Canadensys:** In coordination with CBIF, the official Canadian node, it would be great to start a hunger-mapping project in order to identify the gaps in Canada and to find solutions to resolve those gaps.
- **iDigBio:** iDigBio will continue its work to expand knowledge of what is contained within collections so that it is easier to determine the remaining volume of work.
- **International Centre for Integrated Mountain Development:** Advancing Biodiversity Informatics Capacities in the HKH- Regional Workshop (define data gap map for Asia): explore possibilities of collaboration with GBIF, relevant institutions in HKH regional member countries, and Asian Node.
- **Japan:** Gap analyses to be carried out in some model collections.
- **Naturalis Biodiversity Center:** Work on large scale digitization plans will continue in DiSSCo-linked projects.
- **Norway:** GBIF Norway and GBIF Portugal intend to represent GBIF in the agrobiodiversity community during 2020 with several new project proposals under development.
- **Spain:** We will start mobilizing data from private sector in collaboration with Nodes from Norway, Portugal and Colombia. Define priorities for content mobilization in coordination with GBIF Secretariat and Ministry of Environment.
- **Sweden:** Among the collaborating institutions of the Swedish Biodiversity Data Infrastructure work will continue to develop and adapt systems for accepting and presenting new data types.
- **United States:** Work with the Global Ocean Observing System and GEO BON to identify priority datasets related to Essential Ocean and Essential Biodiversity Variables.
- **Zimbabwe:** Citizen science will be a focus.

Rationale

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. The Ebbe Nielsen Challenge for 2016 focuses specifically on tools

and algorithms to identify significant gaps. These gaps may relate to different facets of the data, including geography, taxonomy, time periods, and coverage of particular ecosystems or land units. Addressing these 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.

Approach

The GBIF Secretariat will harmonize and document data mobilization demands from different sources. Simple tools are required to support needs capture, including informative documentation and justification for such demands. Automated assessment and reporting of gaps will be included where this proves possible and valuable. This combined information can support transparent decision making and target setting for gap-filling efforts, allowing all interested actors to step in at appropriate levels. GBIF will coordinate with efforts through the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) to identify and address significant knowledge and data gaps, including outreach and funding strategies for gap-filling. A thorough, regularly updated overview of data coverage in gbif.org both makes it easier to identify gaps, and to monitor progress and efficiency of mobilization efforts over time. GBIF should offer (e.g. annually) a brief report of significant gaps which need to be addressed. Such a report may be valuable to Participants and funding bodies to stimulate and evaluate digitization and mobilization options.

Activity 3b: Expand data streams

Tasks

- i. Promote use of sampling event data model for ecological and monitoring datasets
- ii. Partner with BHL and others to support integration of species occurrence records based on literature
- iii. Work with bioinformatics initiatives and databases to form robust bidirectional linkages with molecular data
- iv. Explore opportunities to integrate species-level data from remote sensing

2019 Progress

2019 saw the continued growth of occurrence data on GBIF.org. Major new data types available are occurrence records derived from single sequence and metagenomic datasets of fungi and bacteria. GBIF is also ingesting occurrence data arising from barcode sequences deposited in the Barcode of Life Database (BOLD). GBIF presented at the Living Norway symposium targeting ecological datasets and is investigating protocols for sampling event data. GBIF is proactively investigating new data streams through outreach seminars at major events.

2019 Participant contributions

- **Argentina:** With the work with the environmental secretary we're going to promote use of sampling event data model for ecological and monitoring datasets.
- **Australia:** The Atlas went through a process of business analysis on species level and specimen level trait data, assessing availability and the use of standards and ontologies. The analysis also documented potential use cases for the use of traits in Atlas.
- **Biodiversity Heritage Library:** Implemented the addition of already transcribed materials to BHL.
- **Colombia:** Training Workshop with the National Institute of Health to publishing data relating to vectors and hosts of human diseases. The workshop took place on 20 September 2018 at National Institute of Health, Bogotá. A total of 10 participants attended the workshop. <https://sibcolombia.net/taller-sirap/> <https://sibcolombia.net/taller-ins/>
- **Germany:** As part of GFBio (German Federation for Biological Data, www.gfbio.org): establishment of workflows and software infrastructure for archiving and publishing data sets from national biodiversity research projects. GFBio includes the provision of observation and collection data to GBIF via the established GBIF-D data centres and BioCAsE/ABCD interfaces. The German Botanical Node hosts the technical secretariat of the Global Genome Biodiversity Network. 90 Biobanks provide more than 1.6 million standardized DNA and 370,000 tissue sample data. Most of the linked specimen data records are also published in GBIF. MetBaN (<https://github.com/sproft/MetBaN/tree/v0.1.0>) is a bioinformatic pipeline which implements a modular and flexible species delimitation approach by streamlining metabarcoding and phylogenetic software packages. MetBaN is developed at the Botanical Node.
- **iDigBio:** iDigBio sponsored a Phenology Deep Learning workshop in January 2019. This workshop focused on exploring such technologies as deep convolutional neural networks (CNN), data protocols and standards for linking trait data to specimen records, tracking phenological synchronization across phyla, and tools for trait measurement and analysis. iDigBio sponsored an Imaging and Digitization Workshop for Amateur Paleontologists workshop in July 2019.
- **International Centre for Integrated Mountain Development:** Provided technical support to Department of Plant Resources to publish occurrence data of endemic flora of Nepal via HKH-BIF platform (http://rds.icimod.org:8080/hkh-bif/resource?r=endemic_flora_of_nepal_kath)
- **Naturalis Biodiversity Center:** DiSSCo is discussing with other RIs, Plazi and in RDA how to serve expanded collections data with improved provenance and all data derived from specimen linked back to the specimen.
- **Norway:** Based on the recommendations and guidelines from the GBIF CESP Bireme report (developed by the GBIF Nodes of Belgium, France, Ireland, Norway, Portugal and Species2000), GBIF Norway has engaged with the Agency of Environment in Norway to promote the use of the GBIF data streams in EU-directive biodiversity information reporting. The Norwegian Environment Agency are positive to engage with the European Environment Agency (EEA) to explore Bireme recommendations further. The need for adjustments to the data standards and data models (see Activity 2a) should be expected.
- **South Africa:** Citizen Science community and records catalysed and available through iNaturalist.

- **Sweden:** Increasing the exactness of metadata and sample-based data is highly prioritized with GBIF-Sweden, as is that of integrating procaryote and sequence data.

2020 Work items

- Enhance the data exchange standards for sampling-event data, collaborating with partners that generate data to provide sources for filling current gaps. This work aims to establish partnerships with long-term monitoring communities.
- Improve linkages between records originating from museums and BOLD in order to link information that is currently treated as two occurrences.
- Carrying over the proposed 2019 work item, mobilize data on vectors and hosts of human diseases. Establish an expert group (€25,000) to identify priority needs for biodiversity data supporting disease research, critical gaps in availability of such data in GBIF.org, and potential sources of data to fill these gaps. The campaign will use this analysis to engage directly with relevant data holders, support data publication through GBIF and inform data mobilization priorities for use by nodes, publishers and funders (see Activity 3a).
- Continue linking and integration of sequence-based data streams.

2020 Participant plans

- **Andorra:** We plan to review the literature, especially old contributions, on biodiversity related with Andorra in order to digitize data useful to add to the GBIF database.
- **Argentina:** Continue with the promotion of the use of sampling event data model for ecological and monitoring datasets.
- **Biodiversity Heritage Library:** Review options for transcribing from within BHL.
- **Germany:** “GFBio activities continue with the aim of ensuring the sustainability of the infrastructure. MetBan will be fully operable. In consultation with GGBN network partners, advance the integration of GGBN and GBIF with respect to DNA and tissue samples.
- **Japan:** Seek data to be used in education (high school biology).
- **Naturalis Biodiversity Center:** DiSSCo will provide a demonstrator how to deliver expanded collections data using a Digital Object infrastructure.
- **Norway:** Dependent on continued stable research infrastructure funding for GBIF activities in Norway, GBIF Norway will mobilize and implement new data streams for genome, environmental DNA, and ecological data from Norwegian data publishers in GBIF (see also Activity 2a).
- **South Africa:** iNaturalist platform being put in place for Southern Africa by SANBI-GBIF. Also a Node of IBOL is being developed, by the South African community, with a catalytic meeting at the annual BIM-FBIP Forum. This will facilitate the mobilization of molecular data.
- **Sweden:** GBIF-Sweden seeks to expand activities within the above mentioned fields of work in 2020.
- **United States:** BISON project will focus on occurrence datasets for human and wildlife disease vector/host species as well as update native insect datasets. BISON also working to publish forest inventory and bird banding data. Additionally, GBIF node manager through workshops

and at national conferences will work to expand marine data accessibility and broad scale ecosystem datasets like NEON and LTER. Will also focus on eDNA.

Rationale

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.

Approach

Existing GBIF models include support for occurrence records and for sampling-event datasets which organize occurrence records as sets of observations deriving from a single field sample (which make provision for GBIF to accommodate “absence data” from surveys which did not record a particular species despite searching). These approaches are core to all potential streams of data to be added. GBIF therefore needs to ensure that existing tools and documentation are clear and usable for relevant research communities and that GBIF sufficiently understands existing data management by these communities to avoid proposing unnecessary additional work. During 2016, GBIF is coordinating a consultation which builds on past engagements with genomics activities such as the Global Genome Biodiversity Network. Recommendations from this consultation are expected to guide improvements in GBIF tools, documentation and communications to support publishing of molecular data in formats which can be integrated within GBIF. Several projects are working on automated or human mining of data records from literature. GBIF needs to learn from these initiatives and ensure that its tools support integration in a simple way. GBIF should also seek exemplar projects for bringing occurrence records from remote sensing into the network.

Activity 3c: Engage data holders

Tasks

- i. Engage with natural history collections not yet publishing to GBIF
- ii. Work with national citizen science groups
- iii. Promote national policies and platforms to enable publishing of environmental impact and monitoring data
- iv. Promote importance of data sharing to funding bodies, licensing authorities and industry

bodies

v. Address the tropical data gap

2019 Progress

The Secretariat has signed a Memorandum of Cooperation (MoC) with IUCN to further collaborations. The purpose of the MoC is to strengthen technical and institutional collaboration between GBIF and IUCN with a view to improving the visibility, timeliness and usefulness of data and information exchanged between the two networks, thus adding value to their respective products and services. The immediate objectives are to build upon past close work on red-listed and invasive species in data sharing, visualization and indicator development.

Contributions of the citizen science community to the [publication](https://data-blog.gbif.org/post/gbif-citizen-science/) [https://data-blog.gbif.org/post/gbif-citizen-science/] of primary occurrence data are analysed through an algorithm-assisted process, tagging datasets for further processing and metrics generation. An update to the citizen science study from 2016 is given in the [GBIF data blog](https://data-blog.gbif.org/post/gbif-citizen-science-data) [https://data-blog.gbif.org/post/gbif-citizen-science-data].

2019 Participant contributions

- **Andorra:** One of the main tasks of GBIF-Andorra is to make to know to our Andorran partners the sense and the need of GBIF network for promoting the use of ALA-Andorra. At the same time, we are encouraging people to participate in citizen science projects related to biodiversity.
- **Argentina:** Argentina promotes the realization of the City Nature Challenge at the national level and the iNaturalist application to be used in different projects. In addition, we participate as a founding member of the Iberian American Network of Citizen and Participatory Science (RICAP). Also we're promoting the CC with workshops, talks and activities for scientists and public in general.
- **Australia:** Uptake of the BioCollect platform has continued to grow and consolidate across Australia in citizen science data collection as well as professional ecology and state & local government sectors. WA state government is using BioCollect as a project-based mechanism for sharing proponent data from planning and development applications. The Atlas is commencing a project to parse and harvest biodiversity occurrence data from those projects. An instance of BioCollect has been established in Queensland to support project-based state government program delivery and data collection and sharing for environmental intervention activities within the terrestrial watershed to the Great Barrier Reef.
- **Benin:** 17 data publishers are registered on GBIF site and will be more and more active in publishing their data.
- **Canadensys:** Not directly in link with the items in the work program, but the workshops developed during the CESP program were intended to reach and engage data holders from Canada and the United States. Participation during more general events in Montréal have also helped engage citizens.
- **Colombia:** "Work with the Colombian community of iNaturalist supporting the City Natua Challenge 2019 in Bogotá and Medellín. <https://sibcolombia.net/reto-naturalista-urbano-2019/> <https://www.inaturalist.org/blog/25956-colombia-inaturalist-world> To June, 119 national organizations as data publishers through GBIF Colombia.
- **Consortium of European Taxonomic Facilities:** The helpdesk offered by the BioCAsE

secretariat (which also functions as CETAF's participant node for GBIF) is consulting and assisting potential new data providers in the process of sharing their collection and observation data with GBIF. This explicitly includes providers from countries not yet participating in GBIF, such as Italy and Albania.

- **France:** “Data publishing support to TOTAL Company Contribution to two CESP projects:
 - “OpenPSD - Engage and promote the private sector in open biodiversity data publication”, lead by GBIF Spain
 - “To use the CBD's CHM infrastructure and network in order to strengthen biodiversity data acquisition and data sharing”, lead by the French Head of CHM focal point (Denis Duclos, DREI, MNHN, Paris).
- **Germany:** “A Workshop for German herbaria focusing on digitization and data infrastructure was organized by the German Botanical Node. The Mycological Node organized training workshops for users of the Diversity Workbench software.
- **iDigBio:** Data mobilization is a core activity of iDigBio, including reaching out to collections.
- **Japan:** Expanded through hosted meetings.
- **Mexico:** Increase data publishing to be continued, approximately 500, 000 more occurrences records.
- **Naturalis Biodiversity Center:** DiSSCo applied to be come an other associate participant so it can engage its data holders to supply data to GBIF As part of a project financed by the Dutch research council (NWO, ALWIN.005), Naturalis together with NLBIF, and the Foundation for Dutch Natural History Collections have facilitated the signing of a declaration to work together on a Netherlands biodiversity information infrastructure. This declaration is signed by all members of the DiSSCo/NL consortium and the Foundation for Dutch Natural History Collections. The DiSSCo/NL consortium will be coordinated by NLBIF.
- **Netherlands:** NLBIF, as DiSSCo-NL NTF lead has engaged with all Natural History Museums in the Netherlands to develop a National Collection Dashboard. In a follow up DiSSCo-NL partners will be supported to contribute their digitized collections to GBIF.
- **South Africa:** SANBI-GBIF and the Systematics Division will implement the 2019 National Biodiversity Information Management and Foundational Biodiversity Information Management Forum (BIMF-FBIP Forum).SANBI-GBIF 2019 Training event will be conducted alongside the BIMF-FBIP Forum. SANBI-GBIF will conduct a national training event, which will also include regional representation, as well as trainers which form part of the GBIF mentors network. This will support the development of communities of practice in data management nationally and regionally and aligns with the objective to develop the Centre for Biodiversity Information Management.
- **Sweden:** GBIF-Sweden keeps searching for and approaching new potential data providers.
- **Switzerland:** Collaboration with the Laboratory of Applied Microbiology, University of Applied Sciences of Southern Switzerland (SUPSI) and the Vector Biology and Control Group, Swiss Tropical and Public Health Institute (SwissTPH) in order to mobilize and share the data of cantonal and federal surveillance programs on invasive alien mosquitoes.
- **Zimbabwe:** First meeting held in January 2019.

2020 Work items

- Continue work with DiSSCo project team to maximize opportunities for mobilizing collections data from European institutions, including in countries not yet participating in GBIF.
- Work with iNaturalist, iDigBio and nodes community to maximize opportunities for public engagement in GBIF data mobilization.
- Develop private-sector data mobilization guidance and training. Promote revised guidance on mobilizing EIA data and run training programme for private sector consultants at IAIA conference in Spain 2020, and incorporate guidance from CESP project in nodes guidance package.

2020 Participant plans

- **Andorra:** In the near future, we will work to add the biodiversity information from citizen science in the GBIF database. The data from the Ornitho project is an example we are working at this moment (<https://www.ornitho.ad/>)
- **Argentina:** Keep promoting the CNC and participate at the ECSA in Trieste with a researcher and member of the RICAP. Keep promoting the CC.
- **Australia:** Continuing work to promote and support use of the BioCollect platform across multiple communities. Continuing to work with the international citizen science and standards communities on a global standard for project-based data exchange. This work is progressing steadily and ALA is playing a key role in it. Preparing for a major UI refresh of BioCollect. Improving APIs for data and metadata exchange as well as external platforms interaction with BioCollect such as external mobile apps for data collection and external data consumer platforms. The Atlas's engagement with the federal Department of Environment and Energy (DoEE) via the MERIT application is yielding increasingly hi-resolution and accurate rich data about Australian government investments in environmental interventions and outcomes. Working with DoEE it is planned for biodiversity data embedded in these data sources to be mined to augment existing occurrence data streams.
- **Canadensys:** We will continue to reach new data holders and work with the one that already have published data on Canadensys or other platform.
- **Consortium of European Taxonomic Facilities:** BioCASE helpdesk will continue to assist data providers with data publication.
- **France:** Promotion of data publishing and activities of the CESP projects to be continued.
- **iDigBio:** iDigBio will continue its data mobilization efforts, with a vision of mobilizing ALL collections data in the United States.
- **International Centre for Integrated Mountain Development:** Publish ICIMOD programme based biodiversity data from different landscape initiatives Proposal development support to partners in the HKH regional member countries- BIFA/ GBIF-YRA Explore funds to support students in the Central Department of Botany to mobilize university herbarium based dataset
**Subject to fund availability.
- **Japan:** To be expanded through meetings.
- **Naturalis Biodiversity Center:** DiSSCo will engage with natural history collections in DiSSCo not yet publishing to GBIF to supply data to GBIF.

- **Norway:** As part of the GBIF CESP OpenPSD project (developed by GBIF Nodes in Spain, Portugal, Norway, Colombia and France) we will engage new data holders in the private sector (see also Activity 1g).
- **South Africa:** SANBI-GBIF and the Systematics Division will implement the 2020 National Biodiversity Information Management and Foundational Biodiversity Information Management Forum (BIMF-FBIP Forum). SANBI-GBIF 2019 Training event will be conducted alongside the BIMF-FBIP Forum.
- **Spain:** We plan to replicate the LichenCity project in 2020 in different Spanish cities. Organize workshops on data mobilization and data use for specific stakeholders such as private sector, and public administrations. We plan to mobilize data from the private sector.
- **Sweden:** By approaching citizen science groups, researchers and management practitioners substantially more data and data types are going to be published in the coming years.
- **Switzerland:** Ensure data publishing for all major collection-holding institutions of Switzerland. Continuation of efforts towards a partnership with collection holding and research institutions active in DNA sequencing (linkage of sequence data, DNA-samples and reference specimens). Collaboration with national biodiversity data centres in order to capture and use species trait data.
- **Zimbabwe:** Capacity enhancement workshops e.g. Darwin core standards, data quality tools and IPT.

Rationale

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.

Approach

The primary need is for national Participants to engage with the broadest possible spread of data holders within their countries, for organizational Participants to share relevant data and for the whole GBIF network to promote the importance and value of sharing data. The GBIF Secretariat will focus on enhancement to documentation and tools and on highlighting priorities for complementing existing data and addressing gaps. Data holders should be encouraged and assisted in sharing data in the richest form appropriate for the data in question (sampling event data where relevant elements are available, occurrence data for other spatially explicit data, checklists otherwise) and with the most open data licences possible. As well as natural history collections, Participants should identify opportunities to build partnerships with citizen science groups and promote the value of open access to data from environmental impact assessments and monitoring. GBIF should argue the case for open data as part of the policy for funding agencies, research councils, industry bodies, licensing authorities, development banks and other stakeholder groups.

Activity 3d: Rescue datasets

Tasks

- i. Develop tools for reporting potential data sources for integration into GBIF
- ii. Develop support materials (including accreditation) for collaborative data preparation and mapping datasets in GBIF
- iii. Develop site and support mechanisms for users to adopt and map datasets
- iv. Review and update definitions of data publisher within GBIF to reflect collaborative data publishing
- v. Develop partnerships with data journals to support data papers for rescued datasets

2019 Progress

Explored metadata-only dataset publication of the datasets proposed through <https://www.gbif.org/suggest-dataset>. Currently developing categorization of datasets to improve implementation of the ‘Suggest a dataset’ tool in GitHub.

2019 Participant contributions

- **Canadensys:** We have accepted to take care of orphan datasets from Canadian institutions currently published through outdated protocols (DiGIR, TaPIR) by CBIF. We are waiting for CBIF approval to start the ingestion process.
- **Colombia:** Rescue of data dataset from Biota Colombiana Journal. Dataset before the implementation of data paper model. All datasets published through <https://ipt.biodiversidad.co/iavh/>
- **Integrated Taxonomic Information System:** ITIS' taxonomic workbench development will provide the biodiversity community with a tool that provides a stable and long-term solution for updating and maintaining taxonomic information.
- **International Centre for Integrated Mountain Development:** Migrated 14 checklist data to HKH-BIF. Our IPT registered earlier in 2012 became dormant so had to be reinstalled and re-registered. So we had to move our earlier published datasets into this current HKH-BIF platform.
- **Sweden:** As in Work Programme item 3c.
- **United States:** Several projects—including iDigBio, VertNET, BISON—working with researchers to rescue, QA/QC and provide access to U.S.-originating datasets.

2020 Work items

- Continue to implement workflow for prioritizing and drawing upon potential data sources reported through the ‘dataset catcher’ tool, including involvement of nodes, mentors and crowdsourced solutions.
- Roll out a workflow for ‘Suggest a dataset’ processing.

2020 Participant plans

- **Canadensys:** We will temporarily host the CBIF orphaned datasets until CBIF is ready to host them again.
- **Integrated Taxonomic Information System:** In 2020 ITIS plans to release its taxonomic workbench that will provide the biodiversity community with a tool that provides a stable and long-term solution for updating and maintaining taxonomic information.
- **Netherlands:** Once the NLBIF website is renewed it will be used to further promote the vision and mission of GBIF and NLBIF, including the rescue of datasets.
- **Sweden:** As in Work Programme item 3c.
- **United States:** BISON will focus on rescuing amphibian data.

Rationale

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.

Approach

The GBIF Secretariat, or an interested Participant, should develop a test environment to explore this model. The model should support identification of basic information on datasets which may be rescued, including details of ownership, etc. Volunteers may be required to undergo some training or demonstrate some knowledge of GBIF data publishing and the taxa concerned prior to adopting a dataset for mobilization. Mobilization should include consultation or partnership with the owner and should deliver quality metadata and a valid mapping of the original information. Opportunities should be explored for publication of resulting datasets as data papers as an incentive to all parties.

Activity 3e: Liaise with journals

Tasks

- i. Develop scalable approach to support research journals and data journals in publishing to GBIF network
- ii. Produce relevant support materials to justify benefits and explain processes to publish primary data
- iii. Integrate support for data journals into hosted IPT infrastructures and data rescue processes

2019 Progress

In 2019, the focus is on the development of standard workflows and simple recommendations that support and eventually mandate the process of depositing supplementary primary biodiversity data, both to aid submitting authors and publishing houses. By the end of 2019, a first version of an information page on GBIF.org will be available for journal publishers to reference when recommending GBIF as a data repository, outlining the data publication process for authors, and pointing at simple data spreadsheet templates. The option of offering hosted IPT installations for journal publishers for this purpose is under evaluation.

2019 Participant contributions

- **Canadensys:** Canadensys publishes on its IPT data, such as specimens consulted for taxonomic revisions and georeferenced observations used in other research, using DOIs that are included in peer reviewed publications.
- **Colombia:** Continued support to Biota Colombiana with the data paper model, a national scientific journal. <http://revistas.humboldt.org.co/index.php/biota>
- **iDigBio:** iDigBio began developing comprehensive citation guidance for data obtained from its portal. The draft citation guidance is available on the Data menu in the Beta Portal.
- **South Africa:** SANBI-GBIF will integrate data paper publication within Bothalia - African Biodiversity and Conservation Journal

2020 Work items

- Work with journals to make it easier for article authors to deposit supplementary data in formats suitable for GBIF publication.
- Establish model guidelines for data deposition and citation and develop compliance criteria for distribution to publishers. Build on recent developments of COPDESS and Research Data Alliance with GBIF-specific guidance. Develop additional communication materials to describe the benefits of DOI-citation best practices and work with journals on implementation.

2020 Participant plans

- **Canadensys:** We will continue to use DOIs to link datasets and to keep track with the publications using data published on Canadensys.
- **iDigBio:** iDigBio will continue to refine its citation guidance in response to feedback.

Rationale

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.

Approach

GBIF Secretariat will engage with relevant journals in developing and promoting best practices and publishing workflows that reduce the effort required to publish data to GBIF. The aim will be to publish data once but be able to cite it as supporting data for a journal paper and where appropriate produce a data paper as an additional product. The Secretariat will use its experience with Pensoft Publishing where a workflow and a dedicated data paper publishing tool is established.

Priority 4: Improve Data Quality

Activity 4a: Ensure data persistence

Tasks

1. Identify and verify datasets within GBIF network without current owners
2. Publish reference instances of these datasets within hosted IPTs
3. Develop processes and mechanisms for adoption of orphaned datasets by suitable agencies or experts

2019 Progress

The exploration of necessary steps to achieve CoreTrustSeal data repository certification is starting in Q4 2019. This includes the data management services within GBIF.org, but also seeks to identify a set of trusted repositories for publishing datasets within the GBIF network.

Documentation for all stages in the GBIF data ingestion process to start Q3/Q4 2019, following deployment of the informatics infrastructure upgrade (pipelines project).

2019 Participant contributions

- **Biodiversity Heritage Library:** Make progress on adding BHL Europe data to BHL. Minimal progress has been made. Developed a plan for a rapid ingest of BHL Europe materials.
- **iDigBio:** iDigBio partnered with the Society of Herbarium Curators to conduct a Strategic Planning for Herbaria short course. The goal was to produce a strategic plan for each represented herbarium, including vision, mission, stakeholders, strategies, goals, objectives, evaluation, and sustainability, among other things. iDigBio has been working with the community on alternative data storage solutions, such as CyVerse.
- **Mexico:** In process, mentoring and collaboration with UNIBIO-Instituto de Biología UNAM publisher for reactivate data publish and rescue orphaned datasets.
- **Norway:** GBIF Norway contributed to best practices guidelines and implementation of specimen-level DOIs in collaboration with the UN Food and Agriculture Organization in coordination with the GBIFS (see also activity 3a, 4b, and 5b).

2020 Work items

- Continue revision and documentation of flagging routines used in GBIF data ingestion pipelines.

2020 Participant plans

- **Biodiversity Heritage Library:** Implement the plan for rapid ingest of BHL Europe materials.
- **Canadensys:** We will follow the recommendations from GBIF in order to ensure data persistence.
- **iDigBio:** iDigBio is working to improve its data mobilization efforts and workflows, including moving towards IPT as the preferred publishing mechanism. iDigBio will continue to work with the community on alternative data storage solutions and strategies.
- **Naturalis Biodiversity Center:** Persistent identifiers for specimen related data will be implemented in ELViS.

Rationale

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.

Approach

The task is to ensure that all datasets have a primary version available on the internet which acts as the source for GBIF.org to index. Orphaned datasets will be identified, extracted from the GBIF.org index and loaded into the most suitable data repository supporting versioning: either run by a GBIF participant or a central cloud installation of an IPT. As issues are identified anyone will be able to volunteer to correct the source data, upload a new version into the data repository, document the changes applied and follow editor guidelines. Once republished GBIF.org will reflect the updated data, and the provenance of changes will be traceable through the repository versioning system. Policies for editors, including attribution and the settlement process for disputes will be documented. This entire activity could be led and implemented by a GBIF Participant.

Activity 4b: Assess data quality

Tasks

- i. Develop extensible data validation tools framework in partnership with ALA, TDWG and other networks (e.g. Symbiota, iDigBio, VertNet)
- ii. Integrate consistent data validation tools in GBIF.org, national/regional portals, IPT and elsewhere
- iii. Improve presentation and reporting of data validation results
- iv. Develop regular data set reports for data publishers and nodes

2019 Progress

Performed data analysis on implicit geographic data accuracy/patchiness within datasets (gridded data, centroids), documented in the GBIF data blog, estimating the size of impact on further analyses (EBV, species distribution). Tagging of indexed datasets to support further automated data processing.

Exploring of suitable measures to indicate metadata completeness and appropriate content, focused on the contribution from the BID data mobilization.

2019 Participant contributions

- **Argentina:** Training workshops in different areas of the country for data providers. Permanent help desk in the publication and data quality.
- **Australia:** ALA and GBIF ran a combined workshop in Canberra in February to analyse the potential areas of collaboration and alignment of tools and code bases.
- **Benin:** The node of GBIF Benin is more and more capacitated and assure adequate quality to data through cleaning process before data publication.
- **Canada:** CBIF staff contributed to the work of the TDWG Biodiversity Data Quality Interest Group on data quality standards and controlled vocabularies.
- **Canadensys:** All datasets published on Canadensys are verified prior to publication using the GBIF Validator Tool. We also invite data holders to use that tool before submitting dataset to Canadensys. For already published dataset, we invite data holders to use the metrics linked to their resources in order to enhance data quality for their next publication.
- **Colombia:** Open Refine Scripts SiB Colombia is working in a set of scripts in OpenRefine for data quality management of primary biodiversity data. Some of these scripts are being used now by the Node Staff in the process of accompaniment that is done with the publishers, assessing the quality and generating quality reports for each dataset before them will be published through the IPT. The Colombian Node is making improvements on these scripts and developing new scripts to consolidate a toolkit in OpenRefine available for all the GBIF and informatics biodiversity community. An article about this work is also being built to be published in 2018–2019. OpenRefine scripts Repository: <https://github.com/SIB-Colombia/data-quality-open-refine>.
- **France:** GBIF France contributes to harmonies and put in place a better national data workflow.
- **iDigBio:** iDigBio staff are participating in the TDWG Data Quality Interest Group. iDigBio hosted a Collecting Metrics of Success symposium at SPNHC 2019. This symposium discussed the current metrics being collected, what metrics are needed, and who needs them. In addition, the symposium began a conversation about metrics the community could amass collectively, across collections and across the world.
- **International Centre for Integrated Mountain Development:** Supported TUTH – data publisher from Nepal to validate occurrence datasets for publishing compiled under BIFA project – “Mobilizing occurrence data of alien and endemic plant species of Nepal” (Data published through TaiBIF- http://ipt.taibif.tw/manage/resource.do?r=bifa3_25).
- **Japan:** As a result of assessment, some amount of data duplication was detected.
- **Norway:** GBIF Norway helpdesk assists Norwegian data publishers with assessment of data

quality before and after data publication. And a very active network of enthusiasts and professionals in Norway provide rapid feedback and suggestions to fixing data quality issues on Norwegian datasets (see also Activity 1d).

- **Spain:** GBIF Spain organized online data quality workshop for 28 people including international participation. We received nearly 100 inscription forms. We run data quality tests before publishing datasets through GBIF. Working on including new data validations in our tool “Darwin Test” to check more Darwin Core fields.
- **Sweden:** Employing data validation tools developed nationally at one partner institution of the Swedish Biodiversity data Infrastructure, GBIF-Sweden takes responsibility for data provided by Swedish owners.
- **Switzerland:** Participation in a project to establish a registry of Swiss collectors for validation purposes.

2020 Work items

- Review, consolidate and update existing documentation for data publishers. In particular, provide clear guidance on minimum requirements for published data.
- Develop metrics to track the completeness of core data elements and the degree to which supplied content is appropriate.
- Supply clear indicator measures for the completeness and usability of data as part of GBIF.org dataset pages, based on examples such as the GEOLabel data branding model.
- Extend data-quality assessment to include aspects only detectable above the level of individual records.
- Assess the patchiness of indexed data (geographical clustering, misleading accuracy or precision of coordinates), including evaluation of the apparent causes of data patchiness and include measures of data patchiness in the data index, at both dataset and record level in the data index.
- Ensure that users of data are able to identify datasets or records that do not fulfil their criteria for geo-accuracy, whether they are accessing data through facets in the GBIF.org, via the API or in downloads.

2020 Participant plans

- **Argentina:** Training workshops in different areas of the country for data providers. Permanent help desk in the publication and data quality. Develop regular data set reports for data publishers and nodes.
- **Australia:** GBIF work plan participation response GBIF work plan participation response 100% 10 SDG 2 of 2 Context: AA5 Continuing to collaborate with the SDG community and implement platform-based solutions to improve the value, applicability and accessibility of citizen science data for SDG's. Further review work is scheduled for this year as part of ALA infrastructure refresh., Focus areas will be data registries and occurrence data processing and data quality.Screen reader support enabled. Further review work is scheduled for this year as part of ALA infrastructure refresh. Focus areas will be data registries and occurrence data processing and data quality. Anonymous Wombat has joined the document.

- **Canada:** CBIF staff will continue to contribute to the work of the TDWG Biodiversity Data Quality Interest Group on data quality standards and controlled vocabularies.
- **Canadensys:** We will continue with the same protocol.
- **iDigBio:** iDigBio has a vision of integrating with other data validation tools. iDigBio will continue its participation with TDWG.
- **International Centre for Integrated Mountain Development:** Continue providing technical services to data publishers in the HKH countries as required- in relation to data publishing in HKH-BIF.
- **Japan:** Continued assessment of data duplication and missing data.
- **Mexico:** To be continued Mexican node assessment of data duplication and development tools for review data of occurrences.
- **Netherlands:** NLBIF will start looking into implementation of the Living Atlas platform and within that scope also look into data quality.
- **Spain:** Enable Darwin Test software to validate checklist datasets and sampling event data datasets.
- **Sweden:** Additional metadata will be covered in the above mentioned system for data validation.
- **Switzerland:** Engage collection holding institutions to complete the Swiss collector's registry.

Rationale

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.

Approach

Consolidation requires an overview of existing data validation tools, their goals and application areas, building on existing community work to produce an annotated tools catalogue (including work by TDWG and the GEO BON “Bon in a Box”). To make best use of development resources, GBIF will support collaboration between networks to bring those developments together and harmonize efforts, so that further development can more efficiently concentrate on new priority areas. Consistent tests and reports will both inform users of the suitability of data for their use, provide feedback to publishers on their holdings, provide a measure for the overall state of the network, and help to prioritize improvement options. Ideally, the most common reporting measures and formats are agreed and unified to a degree that allows publishers an easy cross-walk between and integration of data quality reports supplied by different services and aggregators.

Activity 4c: Enable data curation

Tasks

- i. Develop network-wide approach to handling and processing annotations and feedback
- ii. Enhance IPT to offer dataset-level peer review and commentary mechanisms and record-level annotations
- iii. Develop mechanism and tools within GBIF.org for sharing cleaned and annotated datasets based on GBIF downloads (“reference datasets”)
- iv. Develop GBIF data workbench tool (within GBIF.org and possibly also as standalone) for cleaning and filtering network data (e.g. in red list assessments)
- v. Develop strategy and support mechanisms for expert communities to curate sections of GBIF data

2019 Progress

A pilot handle-based server was deployed demonstrating GBIF infrastructure is capable to participate in a handle network and issue record identifiers.

2019 Participant contributions

- **Canadensys:** We have started using UUIDs for new collections.
- **Colombia:** Data curation in more than 1 million of occurrences published through GBIF Colombia using work routines with Google Refine.
- **Germany:** In cooperation with CETAF member institutions: semantic enrichment of linked open (specimen) data published via CETAF IDs (see <https://doi.org/10.1038/546033d>). The Botanical Node continues to support the AnnoSys system for structured annotation of individual specimens, which is accessible via 13 biodiversity data portals.
- **iDigBio:** iDigBio continues to refine its GitHub workflow for maintaining its collections catalogue.
- **Mexico:** Tools for data cleaning and assessment taxonomic and geographic data.
- **Spain:** We released new version of Elysia, software application to manage natural history collections developed by GBIF.ES, with new features and improvements such as a more direct process to export collection's data from Elysia to the Darwin Core standard. Elysia is also available in English.

2020 Work items

- Continue to explore the use of the GBIF data index to support stable persistent resolvable identifiers for all specimens and occurrence records.
- Explore bidirectional data linking and synchronization with data management systems and publishers to achieve faster and more accurate mutual updates on data improvements and annotations (minimum €10,000).

2020 Participant plans

- **Argentina:** Develop strategy and support mechanisms for expert communities to curate sections of GBIF data
- **Canada:** Implementation of data annotation and attribution services in the DINA Collection Management System currently under development.
- **Canadensys:** We will continue to use UUIDs or other identifier using GBIF's recommendations.
- **Germany:** Further promotion of CETAF Identifiers and semantic enrichment activities. Continued support and promotion of AnnoSys.
- **iDigBio:** iDigBio will be continuing to streamline its collections catalogue and data mobilization workflows.
- **International Centre for Integrated Mountain Development:** Enhancing data quality of Herbarium Specimens and museum collections- regional training for HKH member countries.
** Need to explore possible collaborations with Chinese Academy of Science and other Asian Node member countries.
- **Naturalis Biodiversity Center:** Naturalis will lead on ELViS development that will enable data curation in collections for loans, visits and digitization on demand in DiSSCo.
- **Norway:** GBIF Norway aims to maintain, and further improve, a machine-readable persistent identifier (PID) resolver for as many entities as we have resources to cover within the Norwegian data streams to GBIF. Dependent of continued stable funding, we aim to provide machine actionable data annotation services from this service (see also activity 2a, 3b and 4b).
- **Spain:** Promoting the use of Elysia out of our national borders.

Rationale

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.

Approach

The main task is to provide tools and mechanisms that make it easy for users and experts to contribute knowledge to the available pool of data. Building on existing data filtering and data improvement workflows in the community, GBIF tools and mechanisms are to support the identification of relevant data, their cleaning and preparation for specific purposes, and the sharing of the results of such processes with the wider community. Input collected through existing feedback mechanisms is to be raised to a visibility level that supports and drives the usefulness of the published data.

Priority 5: Deliver Relevant Data

Activity 5a: Engage academia

Tasks

- i. Promote biodiversity informatics curriculum, support training of the users
- ii. Publicize GBIF as data tool via university libraries and faculties
- iii. Develop online collaboration through GBIF.org helpdesk to assist and support data users
- iv. Represent GBIF at the relevant science fora, esp. at the frontier directions such as ecology and molecular research

2019 Progress

The BioDATA project focused on developing skills in biodiversity data management and data publishing. Undergraduate and postgraduate students from Tajikistan, Belarus, Ukraine, and Armenia have an opportunity to take part in the intensive courses, gain practical skills and familiarize themselves with the Global Biodiversity Information Facility (GBIF), international data standards (Darwin Core), software on data improvement, data publishing with the Integrated Publishing Toolkit (IPT), and preparation of data papers. The project is run by the University of Oslo and GBIF Norway exceeding expectations, with more than 8 mentors and 40 students trained on the GBIF publishing and data skills in two events that took place in Belarus and in Tajikistan. The project will continue to train students in Armenia and in Ukraine in 2020, and plans for BioDATA II project are in place to further increase the geographic coverage in Northern and Central Asia and Southern Caucasus.

In Russia, Altay State University, supported by a GBIFS letter, received competitive funding to integrate GBIF and data education in regular biology training for the BSc and MSc programme, with 70 students expected to take part each year. As ASU is already a holder of a similar grant, implementation delayed until 2020, and links to TDWG curriculum group are encouraged.

2019 Participant contributions

- **Argentina:** Publicize GBIF as data tool via university libraries and faculties: this year the NM is going to participate in Argentine days of botanica and in the annual meeting of Argentine palaeontology, promoting the publication and reuse of the data present in GBIF and in the national portal.
- **Australia:** The ALA added some enhancements to Spatial Portal to allow for the exporting of species level trait information from ALA into the EcoCloud platform. The ALA has also commenced initial work on the Collaborative Species Distribution Modelling (CSDM) project which is a collaboration with EcoCloud that will result in functionalities with integration with EcoCloud.
- **Benin:** A functional Regional Master Program in Biodiversity Informatics opened its doors in October 2017 in Benin and students from many African countries are receiving in-depth capacity training to mobilize and use data to address the biodiversity conservation and

sustainable uses in the context of climate and global changes.

- **Canadensys:** We actively inform our users and data holders on how to use data from GBIF, and we use the GBIF Science Review to promote data-use and give examples. Most participants to our CESP workshops were coming from academia and are using data from GBIF for their research. Participants are usually becoming advocates to the use of data from GBIF during their classes, explaining to their students how to get and use this type of data.
- **France:** Contributions to modules in Biology and Systematics masters and doctoral modules at MNHN and Sorbonne University.
- **iDigBio:** iDigBio hosted the 2019 Digital Data in Biodiversity Research Conference at Yale University in June 2019. The conference provides a forum for biodiversity researchers to share and examine the uses of digitized data across all biodiversity disciplines, with special emphasis on using the wealth of digitized specimen data being generated worldwide. In March 2019, iDigBio hosted the 5th Life Discovery – Doing Science Biology Education Conference held in Gainesville. The theme was “microbiomes to ecosystems: evolution and biodiversity across scale, space, and time” which aligned well with iDigBio’s mission to promote the use of biodiversity data in both research and education. iDigBio maintains public libraries on Google Scholar and Mendeley. Recently, iDigBio used the Mendeley API to create a searchable library, including iDigBio’s tags such as grant number, portal used, etc. iDigBio facilitated guest blogs on its website to highlight data use, data cleaning, and other topics.
- **Japan:** Papers using the data from Japan Node increased.
- **Naturalis Biodiversity Center:** Naturalis and NLBIF are hosting the Biodiversity_Next conference, co-organized with GBIF, TDWG, iDigBio, DiSSCo, LifeWatch and CETAF. At Biodiversity_Next, Naturalis Biodiversity Center and NLBIF are organizing a symposium on digital biodiversity data as frontier for science as part of the Biodiversity Open Data Ambassador programme. As part of a project financed by the Dutch research council (NWO, ALWIN.005), Naturalis together with NLBIF, and other partners conducted a survey and in depth interviews amongst 100 scientists within the Netherlands on the need for biodiversity information infrastructures, particularly on natural science collections, to support all the steps of the research data management cycle.
- **Norway:** The Norwegian BioDATA project provide graduate and post-graduate level university training on biodiversity data management and mobilization (see activity 1b). During 2019, BioDATA provided academic training courses in Belarus (February 2019) and in Tajikistan (June 2019). GBIF Norway provides helpdesk services to support Norwegian students and researchers on accessing and using GBIF mediated data together with data from other sources.
- **South Africa:** Further engagement to seek funding and consider the BDI Research Chair appointment, to be conducted with DST/NRF, University Partners and other funders. SANBI-GBIF will look at mechanisms to operationalize the Centre for Biodiversity Information Management, through added strategic and technical assistance.
- **Sweden:** GBIF-Sweden participates, alongside consortium partners in the Swedish Biodiversity data Infrastructure, in a great number of activities directed towards educational and research institutes.
- **Zimbabwe:** Data access workshop held for academia

2020 Work items

- Conduct survey on how GBIF informatics is incorporated into relevant graduate and undergraduate curricula to identify gaps and opportunities and key entry points. Use survey findings to prepare a campaign that could roll out relevant resources through GBIF nodes in 2021.
- Engage with academic-based projects that use or could use better GBIF data in their pipelines and protocols. If funding allows, hold a workshop to stimulate use (€25,000).

2020 Participant plans

- **Argentina:** Keep continue attending scientific meetings to continue promoting the national biological data system and GBIF.
- **Australia:** Further work on CSDM and delivery of components and enhancements to support CSDM.
- **Canadensys:** We will continue to engage academia whenever possible.
- **France:** Participation to master and doctoral modules to be continued.
- **iDigBio:** “iDigBio will host the 2020 Digital Data Conference at Indiana University, Bloomington. iDigBio will continue to facilitate guest blogs on its website. iDigBio is planning a 2020 Biodiversity Summit, which will be a collaborative meeting of the ADBC community, GBIF governing board, TDWG, and the National Museum of Natural History (Smithsonian). The meeting will feature the evolution and accomplishments of specimen-based science and the impact of digitization. This international event will be an excellent opportunity to highlight to a global audience the excellent work that our ADBC community is doing on behalf of U.S. collections.
- **International Centre for Integrated Mountain Development:** Publication of Data Paper / Use of GBIF-mediated biodiversity data – with partners. (Birds diversity in the HKH)
- **Japan:** Continue to increase numbers of data from Japan.
- **Naturalis Biodiversity Center:** Naturalis will conduct a survey under the RDA-Europe Ambassador programme to identify the current needs of the international biodiversity and geodiversity data community for a discipline-specific group under RDA (currently the Biodiversity Data Integration IG)
- **Netherlands:** NLBIF plans to further promote the use of GBIF data for scientific analyses and introduce the value of GBIF data for biodiversity studies in university curricula.
- **Norway:** The emerging new Living Norway network (currently pending research infrastructure funding) has established a working group on how to teach open biodiversity science and FAIR data practices at university level in Norway.
- **South Africa:** Work will continue to appoint a BDI Research Chair. SANBI-GBIF is engaging the Department of Science and Innovation, previous Department of Science and Technology, to support a Priority Research Chair.
- **Spain:** Continue engaging academia. Make an inventory of masters and university degrees that promote GBIF data in classrooms.
- **Sweden:** GBIF-Sweden participates, alongside consortium partners in the Swedish Biodiversity

data Infrastructure, in a great number of activities directed towards educational and research institutes.

- **Switzerland:** Promote GBIF data use and data sharing at Swiss universities and research institutions.
- **Zimbabwe:** Data access and use workshop to be held by March 2020.

Rationale

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.

Approach

GBIF Participants may be well-positioned to lead in engaging with taxonomic societies and other interested research groups, including assuming responsibility to serve as ambassadors for GBIF within particular communities. Particular focus should be given to opportunities to explore models for such communities to assist GBIF with data curation. Few university programs include clear guidance and promotion for open data sharing. GBIF should make use of its own documentation materials and of curriculum materials from throughout the GBIF community to engage with biological science faculties and university libraries to share information on GBIF tools and resources and practices.

Activity 5b: Document needs

Tasks

1. Prioritize areas for fitness-for-use assessments
2. Support fitness-for-use groups

2019 Progress

Continued consideration of recommendations from the Data Fitness for Use reports 2016/2017 (agrobiodiversity, invasive and alien species) in user interface and data processing changes. Some of the recommendations were implemented already, particularly around the incorporation of (to date) 122 GRIIS checklists on introduced and invasive species for countries and territories; other recommendations are still valid as possible future options, some pending structural changes (e.g. the taxonomic backbone offering more flexible ranks).

2019 Participant contributions

- **Canadensys:** Answering the needs from our users and data holders, we have developed documentation for the Canadensys Explorer based on ALA framework, and for data-cleaning. We have also worked on translation of the GBIF website and some documentation for both GBIF and ALA.
- **iDigBio:** iDigBio representatives participated in the Harnessing Natural History Collections Data workshop that led to the report released in April 2019. The report is entitled “Extending U.S. Biodiversity Collections to Promote Research and Education” and articulates a national vision to leverage digital data from biodiversity collections for novel and innovative uses. An iDigBio representative was appointed to the National Academies of Sciences, Engineering and Medicine committee investigating the value and future of biological collections. The committee will review how National Science Foundation-supported collections of organisms are used in research and education and how to maintain them so that they continue to benefit science and society in the future.
- **Japan:** Documents and administrative process reviewed and refined to increase user-friendliness and efficiency.
- **Norway:** GBIF Norway maintained follow-up activities (with focus on Nordic and European crop wild relative communities) based on the recommendations from the 2016 fitness for use task group reports (see also activity 3a).
- **United States:** Contributed feedback through GitHub and feedback mechanisms on the website to help identify needs of the communities we work with.

2020 Work items

- Explore the creation of lightweight, customized website landing pages to address thematic interests on GBIF.org. The pages will increase flexibility to include curated, more informative clustered information, ensuring that users have access to both broad search results and prioritized views of data and information.
- Continue to implement recommendations of past expert user groups. In particular, build on the incorporation of GRIIS checklists to support richer, more targeted information on invasive and alien species, e.g. by highlighting documented occurrences with a relevant IAS status by country. Explore best options to identify and alert users of new occurrences of potential invasive species. In addition, review and address the most feasible and valuable recommendations of the agrobiodiversity group, (see 5c below) with support of continued community involvement.

2020 Participant plans

- **Canadensys:** We will continue helping in the different translation project when needed, as well as creating documentation answering the needs of our users.
- **iDigBio:** iDigBio has several workshops, webinars, symposia, and other events planned.
- **Japan:** Periodical review of documents and administrative process.
- **Netherlands:** The NLBIF node manager as Open Biodiversity Data Ambassador will document user needs for GBIF data wherever relevant.

- **Norway:** Dependent on continued and stable national funding for 2020, the wider Norwegian GBIF community will capture, document and improve the GBIF fitness for use in academic and applied research on ecological monitoring using the sampling event data model (see also activity 2a).

Rationale

GBIF 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. During 2020, GBIF will expand this work to incorporate more key areas of use, provisionally focusing on: 1) Phylogenetic and (continuation from 2016 work) DNA evidence for names and occurrences, 2) human health and vectored diseases and 3) marine biodiversity. Depending on resources, these groups will operate through a combination of face-to-face meetings and online or remote collaboration.

Approach

The DNA work will continue through online meetings and e-mails. Human health and vectored disease fitness for use group will be formed and will operate using the basic model with two in-person meetings. Work on fitness for use in marine research will continue depending on the availability of marine data from OBIS. Participants are encouraged to organize working groups in additional areas to review the state of GBIF data relevant to particular domains and themes and to provide recommendations on critical gaps, improvements in recommended metadata or data elements, minimum criteria for usable data, etc. The GBIF Secretariat will compile these recommendations and, where possible, develop query profiles to support rapid access and monitor progress by the network in delivering data suited for the needs of these domains.

Activity 5c: Support biodiversity assessment

Tasks

- Enhance GBIF.org presentation of sampling event data
- Support development of species distribution/population abundance EBVs
- Showcase tools and practices for use of GBIF in Red Listing and other species assessment processes

2019 Progress

Development of SQL based downloads now allows that arbitrary data queries can be accommodated without the need to download all data. For further discussion of proposals for sampling event data interaction, user interface visuals are being developed in 2019. Meanwhile, a review of available data showed that structured information on sampling protocols is scarce, and will require further work in 2020 to enable visualization and analyses across study sources.

A contract was developed under the IUCN-GBIF Memorandum of Cooperation to allocate €75,000 from the 2019 budget to the IUCN Invasive Species Specialist Group, to fund completion of work on the Global Register of Introduced and Invasive Species (GRIIS) and its connection to GBIF.org. The work will be completed in early 2020.

Data analysis explored options for species density/species richness representation, documented in the data blog to open for further discussion.

2019 Participant contributions

- **Australia:** Collaboration with the SDG community and implement platform-based solutions to improve the value, applicability and accessibility of citizen science data for SDG's.
- **Colombia:** Support to consolidate the IPBES National Assessment, particularly in the chapter 2 - Status of Colombian Biodiversity <http://humboldt.org.co/es/actualidad/item/1257-convocatoria-evaluacion-ipbes-colombia> Status and trend of Colombian biodiversity: An annual publication from the data published. <https://cifras.biodiversidad.co/>.
- **iDigBio:** iDigBio plans to participate in the WeDigBio event in October 2019. Worldwide Engagement for Digitizing Biocollections, or WeDigBio, is a global data campaign, virtual science festival, and local outreach opportunity, all rolled into one.
- **South Africa:** SANBI-GBIF will engage with the IPBES Chair for data and knowledge, to reinforce a stronger IPBES perspective to the ACM work, as it shows the data needs and the potential. The IPBES Africa Assessment reinforces the need for investment in biodiversity information and the ACM.
- **Sweden:** GBIF-Sweden has contributed to the GEO-6 reports and will continue to do so at the level of co-authoring and evaluating contents.
- **United States:** Chairing session at Biodiversity_Next to better connect policy makers with relevant data streams.
- **Zimbabwe:** National biodiversity report for CBD

2020 Work items

- Explore [BIOSCAN](https://ibol.org/programs/bioscan) [https://ibol.org/programs/bioscan] as a pilot for visualizing sampling-event data. Pending advances in the revised data model, improve the representation of sampling-event data, particularly abundance measures. Identify sources to support a consistent, controlled list of sampling protocols.
- Continue work with GEO BON to establish needs to deliver relevant data in suitable formats to support ongoing community-agreed implementation of Essential Biodiversity Variables (EBV). Support further community scoping as and if appropriate within the alliance for biodiversity knowledge framework.
- If funds are available, the Secretariat will engage visiting scientists from the assessment/modelling community for part-time secondments to improve shared understanding of data-delivery needs for the EBV and associated communities.
- Develop and roll out guidance for use of GBIF-mediated data in species risk assessments. Continue work based on the IUCN memorandum of cooperation and proposed training event at the 2020 IUCN World Conservation Congress in Marseille.

- Host hackathon to align GBIF with post-2020 Biodiversity Framework (€40,000 if available). The event would develop tools from the CESP CHM project aimed at improve CBD links to GBIF-mediated data. As post-2020 framework/goal structure emerges, develop comprehensive guidance on the role of primary data to support targets and indicators.

2020 Participant plans

- **Australia:** Continuing to collaborate with the SDG community and implement platform-based solutions to improve the value, applicability and accessibility of citizen science data for SDG's.
- **Belgium:** Explore and prototype a tool, based on occurrences API, that will return a species list from a user-defined polygon.
- **Canadensys:** We would like to work with GBIF and other organizations to establish standardized links to leaf spectral data obtained on the ground and through imagery at various levels (linked to a pan-Canadian project called CABO, <http://www.caboscience.org/>)
- **iDigBio:** iDigBio plans to continue participation in the annual WeDigBio events each October.
- **Netherlands:** NLBIF will promote the contribution of long-term monitoring data to GBIF.
- **Spain:** Organizing a thematic workshop to promote the publication of sampling-event based data addressed to public administrations.
- **Sweden:** GBIF-Sweden has contributed to the GEO-6 reports and will continue to do so at the level of co-authoring and evaluating contents.
- **Switzerland:** Continuation of data digitization activities in relation with running programs of the Federal Office of Environment.

Rationale

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.

Approach

This activity depends on the activities defined for all goals, in order to improve the comprehensiveness, coverage and quality of data accessed through the GBIF network. An additional area for enhancement will be GBIF.org search interfaces and presentation (via maps and other visualizations) of sampling event data. Such interfaces should ensure that researchers can

properly evaluate and assess data suitable for biodiversity assessment activities. Additionally GBIF needs to be active within GEO BON to ensure that there is close alignment and evolution between the concepts and implementation of EBVs and the data and services offered by GBIF.

Activity 5d: Assess impact

Tasks

- i. Develop sustainable approach to literature tracking
- ii. Automate detection and reporting of use of GBIF DOIs
- iii. Provide mechanisms to report and track uses of data and grey literature
- iv. Develop a valuation for the services delivered by the GBIF network

2019 Progress

It has now been 20 years since the OECD Megascience Forum recommended the establishment of a Global Biodiversity Information Facility and a systematic review of GBIF had not been undertaken since 2005. A contractor, CODATA, which led the first review of GBIF, was hired after GB25 to review GBIF's operation, services, governance and place within the broader landscape of biodiversity-related organizations. The review is near completion and will be presented to the GB26 in October.

2019 Participant contributions

- **Canadensys:** Anne Bruneau and Carole Sinou have been interviewed for the external review of GBIF. Did something really important.
- **Colombia:** Use of GBIF API to incorporate in the monthly reports citation of datasets by publishers. Example June: https://datastudio.google.com/u/0/reporting/1yqaSRnp8ANAJZ488Cag1Zlv_sAYQGPiS/page/Ge2V.
- **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.
- **South Africa:** The efforts to reinforce the data-science-policy value chain through the ABC projects and the Foundational Biodiversity Information Programme of SANBI, will continue into 2019.
- **Spain:** The Spanish Atlas has added DOI support for occurrence data downloads. It also downloads a file with providers that must be cited.
- **Sweden:** At the national level GBIF-Sweden continuously evaluates impact.

2020 Work items

- Implement recommendations of the 20-year review and use it as a basis for the development of the next five-year strategic plan for the period 2022–2026.
- Contract out study of use of GBIF-mediated data in the academic literature with a goal to identify areas of impact, links to research funding sources and networks of researchers. This information will be used to fine-tune GBIF value proposition message and align Work Programme to further impact (€20,000).

2020 Participant plans

- **iDigBio:** iDigBio will continue to conduct its annual community survey. iDigBio will continue to track and publish data use statistics.
- **South Africa:** The efforts to reinforce the data-science-policy value chain through the ABC projects and the Foundational Biodiversity Information Programme of SANBI, will continue into 2019.
- **Sweden:** At the national level GBIF-Sweden continuously evaluates impact.

Rationale

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.

Approach

The GBIF Secretariat will lead a consultation with Participants on the level of detail desired from literature tracking and possible approaches to maintain this as a shared task across the network. GBIF.org will be enhanced to present known uses of GBIF DOIs and to ensure that data publishers receive information on such uses. Within the BID programme, GBIF is already seeking reports from funded projects on non-research applications of data. A general model is required to enable Participants easily to contribute examples of such uses and for these to be integrated into GBIF's overall reporting on usage. Several countries have requested information on known costs and benefits from GBIF investment. The Secretariat wishes to identify one or more Participants interested in leading case study investigations to value GBIF benefits at national scale and to assist with developing general recommendations for a suitable way for GBIF to track and report this value.

GBIF Implementation Plan 2017–2021

The GBIF Strategic Plan 2017–2021 presents five broad overlapping priorities for the GBIF network. Each of these priorities sets a direction and guides activities for the coming period. This document here presents an Implementation Plan to enable the GBIF community as a whole to make significant progress in each priority area. It establishes the framework for GBIF's Annual Work Programmes over the period of the Strategic Plan.

The following notes provide context for the activities outlined here.

Delivery

Some activities in this plan will be led or coordinated by GBIF Secretariat staff, using core or supplementary funds. Other activities may be led or coordinated by GBIF Participants with particular interest or expertise in the areas in question.

Combining energies across the whole network within the framework of this plan will enable GBIF to progress the set of activities identified in this document more rapidly. Each Annual Work Programme in the period 2017 to 2021 will build on progress in past years and will focus available resources on remaining priorities identified in this document.

Structure of GBIF

GBIF operates on three primary levels:

- **Global** – GBIF is an international activity and works at the global level to increase standardization and adoption of best practices, to facilitate sharing of expertise and resources and to offer an integrated global dataset and associated informatics infrastructure.
- **Participant** – GBIF Participants, including the nodes that they establish, have primary responsibility for engaging stakeholder communities, mobilizing and curating data, providing support and training, and ensuring that national researchers and agencies gain maximum benefit from GBIF. GBIF regional networks provide a framework for national and organizational stakeholders to collaborate more closely, particularly when addressing challenges and opportunities they share.
- **Data Publisher** – Data-holding institutions, agencies and individual researchers are the foundations on which GBIF depends. GBIF has a responsibility to ensure that data publishers have access to the tools and support that they need and that they receive recognition for their work and contributions.

This Implementation Plan seeks to address needs at all three of these levels while aiming to reinforce the value of GBIF for stakeholders at each level. It is highly desirable to increase coordination and feedback between these levels. The 2019 version of the document includes reports from a few GBIF Participants on their 2018 activities and 2019 plans. It would be most beneficial to receive inputs from a wider range of Participants to give a clear picture of everything occurring in the network and to assist the Secretariat and all Participants with better alignment and reuse.

A particular requirement is for the GBIF network to maintain and coordinate software to support each of these levels and to ensure that the publication, integration, management and use of data are as efficient as possible.

At the global level, GBIF focuses on delivery of GBIF.org as a software platform to support both human users and tools (via web services). Many of GBIF's goals depend on being able to deliver a fully integrated view of all data from all sources in the best possible forms.

At the data publisher level, GBIF maintains the Integrated Publishing Toolkit (IPT) as its default tool for sharing GBIF-compatible data. Ongoing enhancements are required to increase both the usability of this tool and the value that its services provide to data publishers. At the same time, GBIF seeks to support GBIF-compatible data publishing via other tools such as collection management systems and citizen science platforms.

At the Participant level, several efforts have been made to deliver generic portal tools to assist node managers in their work and to deliver value at the national level. The most significant roles for such a platform would be to:

- Enable national- or regional-level integration and curation of datasets
- Enhance data records to reflect national administrative units and national species lists
- Provide tools that meet the requirements of national or regional researchers, policymakers and the general public

Over the last few years, a growing number of GBIF Participants have collaborated to leverage Australia's investment in the Atlas of Living Australia (ALA) and to customize these tools to deliver their own portals. The Implementation Plan responds to this with a focus on increasing community development and support to deliver a reusable and sustainable portal toolkit based on the ALA tools. This approach not only benefits Participants already committed to using these tools, but also provides a solution for rapid adoption by other Participants at national, regional, or even thematic levels. In addition, Belgium has indicated an interest in exploring the potential for an alternative lightweight solution based on repository tools such as CKAN.

The vision for the technical aspects underlying this Implementation Plan is that the components developed and offered by the GBIF network at each scale should combine to deliver a robust solution to support GBIF data management and that together they should continue to evolve to support a leading international data infrastructure. One aspect of this will be to ensure that all users can share and reuse common components such as data validation tests and that GBIF promotes and fosters wider activity around open projects.

GBIF is more than just a technical network, and this Implementation Plan also addresses the capacity enhancement, networking, communication and reporting activities necessary to advance GBIF's work in these areas. Since it is impossible for GBIF to support all aspects of its network using core funds, this plan focuses on resources which will support the work of GBIF Participants and data holders everywhere, coordinating the skills, expertise and investment of the whole network to maximize benefits to all stakeholders.

Timeline

None of the priorities, or the more specific goals identified for each priority, can be achieved in a single year; each will require coordinated effort across multiple annual work programmes. This document describes the set of activities currently recognized as important to deliver the Strategic Plan over the five-year period. It will be revised and expanded as required each year to address changes in the landscape in which GBIF operates.

This Implementation Plan was presented for approval by the Governing Board at GB23. It was accompanied by the Annual Work Programme for 2017. Both components were approved at the meeting.

Additional commitments from GBIF Participants or others to contribute or deliver particular elements will be welcomed throughout the five-year implementation period.

Resources

GBIF has a range of resources for use in delivering its programme of activity, including core funding from annual Participant contributions, supplementary funds (including BID and BIFA and funds received for activities within other projects), and investments by GBIF Participants.

These resources combine as follows:

Core Funding

- Supports GBIF Governance structures (Governing Board, committees)
- Supports staffing and operation of GBIF Secretariat to coordinate global activity, including:
 - Delivery of global-level IT infrastructure (GBIF.org)
 - Coordination of delivery of software components to support Participants and Data Publishers
 - Coordination of capacity enhancement activity and information materials for Participants, Data Publishers and Users
 - Coordination of supplementary-funded activity and Participant investments in shared GBIF activity
 - Monitoring and reporting usage and benefits
 - Administrative support for GBIF activities
 - Communication, outreach and partnership activities
- Limited funds to support workshops, capacity enhancement and developments to expand or enhance GBIF

Supplementary Funding

- Additional funds to support workshops, capacity enhancement and developments to expand or enhance GBIF

- Funding to support Participants and Data Publishers with content mobilization or development of national Biodiversity Information Facilities

Participant Investments

- Operating GBIF nodes
- Engaging with and supporting data publishers and users
- Curating data from Data Publishers
- Additional funds or staffing to support workshops, capacity enhancement and developments to expand or enhance GBIF

In past years, Annual Work Programmes have been developed to reflect the use of Core Funding, but this Implementation Plan is intended to offer a framework for GBIF Participants and other stakeholders to take leadership in delivering some of the identified activities and tasks. Where possible, Participants are encouraged to seek funds or allocate staff time to enable progress in areas beyond those for which Core Funding is adequate.

All such commitments reported by GBIF participants have been included within the 2017 Annual Work Programme and will be reported as part of GBIF's work during the year.

Additional Participant commitments will be most welcome at any stage in the implementation of the 2017 Annual Work Programme and for inclusion in subsequent Annual Work Programmes. Examples of such commitments include (but are not limited to):

- Organization of workshops to develop required new standards or best practice recommendations
- Capacity enhancement and training at regional or global levels, in particular for workshops and projects in regions which are not addressed by current Supplementary Funding sources
- Development of tools or software components to advance GBIF's work
- Representation of GBIF in international fora
- Recruitment of additional GBIF Participants

Please advise the Secretariat at the earliest possible opportunity of any such proposed commitments. A form is made available for this purpose.

Implementation Plan Structure

The Implementation Plan is structured around the five priorities identified in the GBIF Strategic Plan 2017–2021. These are presented here in reverse order from their sequence in the Strategic Plan, and numbered accordingly. This reordering allows the Implementation Plan to place its initial focus squarely on the global network of GBIF Participants and data publishers, building on these foundations to address in turn GBIF's informatics, content mobilization, curation and delivered products.

For each of the five priorities, the plan presents a series of numbered Activities. These are the major areas where GBIF needs to develop further in the coming period. A rationale and implementation

approach is presented for each Activity.

Based on the implementation approach, each Activity includes a series of Tasks. These Tasks are the items that the Annual Work Programmes need to address. They are presented here within each section describing the Activity.