

GBIF Work Programme 2014-2016

The updated Work Programme 2014-2016 was approved by the GBIF Governing Board at GB22 in Antananarivo, Madagascar, October 2015.

GBIF Work Programme 2014-2016 – Update 2015

Introduction	3
Executive Summary	3
Stream 1: Advance the Informatics – Persistence and validation	4
Key Deliverables	4
2015 Progress	5
2016 Plans	6
1.1 – Reengineer user interfaces of GBIF.org	6
1.2 – Research and develop services to support sample-based datasets	6
1.3 – Implement a mechanism to support long-term data preservation	7
1.4 – Revise GBIF taxonomic data services	7
1.5 – Operate cloud-based IPT deployments	8
1.6 – Provide publishers with list of papers citing data from their datasets	8
Stream 2: Advance the Engagement – Communication and implementation	10
Key Deliverables	10
2015 Progress	10
2016 Plans	12
2.1 – Biodiversity Information for Development (BID)	12
2.2 – Strategy and updates for GBIF.org	13
2.3 – GBIF regional activities	13
2.4 – GBIF capacity enhancement call	14
2.5 – Integrated training package	14
Milestones	14
Budget	15
Stream 3: Advance the Content – Relevance and fitness-for-use	15
Key Deliverables	15
2015 Progress	16
2016 Plans	16
3.1 – Improving data fitness for key user groups	16
3.2 – Strategy for accelerating discovery of biocollections	17
3.3 – Reference datasets	17
3.4 – Sample-based data promotion and support	17
3.5 – Foundations for a GBIF gap-filling strategy	18
Milestones.....	18
Budget	19
Summary Budget 2016	20

Introduction

At GB20 in 2013, and at GB21 in 2014, the GBIF Governing Board approved, with some modifications, the proposed Work Programme for 2014-2016. For each of the three streams, Informatics, Engagement and Content, this document outlines the projected key deliverables, fuller detail for continuing activity during 2016, and a Work Programme budget for 2016.

The budget secured for 2014 and for 2015 has been lower than that originally proposed and it is unlikely that funding will increase significantly in 2016. During 2015, receipt of deferred payments made it possible to expand the funding allocated to Work Programme activity in comparison with 2014. Nevertheless, missing contributions during 2015 mean that GBIF is expected to enter 2016 with close to the minimum level of reserve funding agreed by the Governing Board (€250K). As a result, this Work Programme for 2016 must limit expenditure to remain within the level of funding expected to be received during 2016.

However, at the start of 2015, GBIF received funding from the European Union for [Biodiversity Information for Development \(BID\)](#), a multi-year €3.9 million programme of content mobilization and capacity enhancement in sub-Saharan Africa, the Caribbean and the Pacific. The GBIF Secretariat is coordinating this activity, which will fund a significant number of projects in these regions, all contributing to GBIF's mission. These funds should accordingly be considered as a significant expansion of the GBIF Work Programme during 2016 and beyond, and the various streams of BID activity are accordingly included within this document. Although these activities do not directly support all GBIF regions, they will contribute to training and best practice materials of use to all. We would welcome opportunities to work with any Participant or external organization to seek complementary funds from their own sources to expand BID globally.

As with the Work Programme documents presented at GB20 and GB21, this document provides a priority order for expenditure within the budget tables. If the budget committed by GBIF Participants for 2016 is lower than €2.85M, lower priority activities (those listed lower in the tables) will be deferred. If the committed budget is higher, or if additional funds are received, some of the proposed additional activities may be undertaken.

GBIF is a global network and its activities encompass work by all Participants. We have sought in this Work Programme revision to identify some of the significant ways in which different Participants have advanced the work this year, and some of the opportunities for the GBIF Work Programme explicitly to be expanded through complementary activity by Participants. We encourage the whole GBIF community to consider ways to accelerate progress in achieving our shared mission.

Executive Summary

The activities and deliverables outlined in this update are organized according to the three strategic imperatives identified in the GBIF Strategic Plan 2012-2016: advancing the informatics, advancing the engagement, and advancing the content. The final year of the three-year Work Programme will coordinate work across all of these three streams to advance delivery of some key benefits to GBIF's users and stakeholders, placing the network in a good position to meet the challenges of the next Strategic Plan period 2017-2021. In summary, GBIF's work during 2016 will:

- Significantly improve the **functionality and usefulness of GBIF.org** and associated communication products to meet the needs of key audiences, both through technical enhancements and reorganization of content, including accommodation of selected information materials in multiple languages (1.1, 2.2)
- Offer new opportunities for sharing and discovery of **sample-based datasets** incorporating information such as sampling methodologies, species composition and abundances, through targeted mobilization and adaptations of GBIF.org to improve filtering and visualizations of data relating to sampling events (1.2, 3.4)
- Improve the **long-term persistence of data** through implementing a pilot repository, enabling for example the uploading of cleaned data from research projects to act as reference datasets for future discovery and re-use (1.3, 3.3)
- Enhance the **quality and fitness for use** of data accessed via GBIF.org through revision of taxonomic services and addressing recommendations of specialist user communities, including new services to filter data based on their fitness for particular uses (1.4, 3.1)
- Enhance the **capacity to mobilize and use digital biodiversity data** in developing countries and among diverse data-holding communities through support to capacity enhancement projects and development of cloud-based data hosting arrangements (1.5, 2.1, 2.4); as well as by developing an **integrated training package** to support development of skills in biodiversity informatics (2.5)
- Offer data publishers improved ability to **track use of mobilized data** through linking citations using Digital Object Identifiers (DOIs) to the source datasets to which they refer (1.6)
- Develop guidance and strategies for **documenting natural history collections through metadata**, to help prioritise further digitization of collection data (3.3)
- Build the foundations for a strategy to **fill gaps in the biodiversity data** available through GBIF.org, by evaluating potential methodologies, consulting expert communities and offering a community challenge to develop viable tools and models for this purpose (3.5)

Stream 1: Advance the Informatics – Persistence and validation

Key Deliverables

From GBIF Work Programme 2014-2016, as approved at GB20:

- Stable operation of new **GBIF.org** and services as a global infrastructure
- Processes to allow all data sets and data records to have **stable identifiers**
- Tools to support **data citation** needs for data publishers and for data users
- Clear **machine-readable data licences** and guidelines on usage for all GBIF-mobilised data
- Guidelines and supporting tools to assess and improve **metadata completeness** for all data
- GBIF.org upgrades to report **data quality and fitness-for-use** for each data set and species
- Proposal for use of repositories to improve **long-term persistence** of GBIF-mobilised data
- Proposal for expanding GBIF model for **endorsement of data sets**
- Demonstration of mobilization of **sample-based data** through the GBIF network

- Criteria and process for **evaluating software compliance** with GBIF recommendations

2015 Progress

The following progress can be reported against Work Programme deliverables:

- **GBIF.org**
 - [Resource library](#) updated and [data use cases](#) integrated in GBIF.org
 - Near real-time indexing of [checklists](#) in place
 - [‘Simple CSV’ downloads](#) and download support for R client introduced
 - GBIF [taxonomic backbone reworked](#) (expected ready for review by end 2015)
 - Maps generated for some occurrence search results (expected by end 2015)
- **Stable identifiers**
 - Integrated Publishing Toolkit (IPT) [enables data publishers to allocate Digital Object Identifiers \(DOIs\)](#) during the publication process using an existing DataCite or EZID account, or to request allocation of a DOI through GBIF through its agreement with DataCite.
- **Data citation**
 - GBIF.org now [includes DOIs for all datasets and user downloads](#)
 - Recommended citations adjusted to take account of new DOI facility, enabling users to cite data downloads with a single DOI that resolve to a page listing all contributing datasets.
- **Machine-readable data licences**
 - IPT [requires compulsory selection of CC0, CC-BY or CC-BY-NC licence](#) for all occurrence datasets
 - Licence declarations for all existing datasets in GBIF.org interpreted for alignment to accepted CC licences, and provisional licence allocation proposed in each case
 - National nodes contacted with proposed licences for all existing datasets from their endorsed publishers
 - All publishers invited to respond to proposed CC licence allocation (expected Nov-Dec)
 - Revised data use and data publishing agreements drawn up taking account of new citation recommendations and new licensing arrangements (expected by October)
- **Metadata completeness**
 - Not yet started
- **Data quality and fitness-for-use**
 - Implementation awaiting recommendations of fitness-for-use task groups
- **Long-term persistence**
 - [Memorandum of Cooperation signed with DataONE](#) to enable implementation of a repository for upload of cleaned datasets with GBIF-assigned DOIs
- **Endorsement of data sets**
 - New procedure implemented for registering and endorsing GBIF data publishers [via online form](#), with [associated guidance](#) for nodes on evaluating endorsement requests
 - Procedure in place for [community-based endorsement](#) of publishers from non-Participant countries, through the Nodes Committee
- **Sample-based data**

- Darwin Core extension for sample-based data [ratified by Biodiversity Information Standards \(TDWG\)](#)
- IPT supports publication of sample-based datasets
- Targeted mobilization of sample-based datasets initiated
- Manual published and associated training organized on publication of sample-based data (by GB22)
- **Evaluating software compliance**
 - Not yet started

Contributions to Stream 1 from GBIF Participants during 2015 include:

- National portals based on the Atlas of Living Australia software have been adapted and deployed in Spain and France, with a prototype in the United Kingdom (Atlas of Living Scotland). This was made possible due to considerable support from Australia. Several other nodes are exploring a similar approach.

2016 Plans

In addition to providing ongoing operational support to nodes, data publishers, users of GBIF.org and the community adapting and adopting the Atlas of Living Australia (ALA) software, GBIF aims to evolve the GBIF infrastructure with the following enhancements during 2016.

1.1 – Reengineer user interfaces of GBIF.org

Note: this activity provides the informatics components required for Engagement activity 2.2: strategy and updates for GBIF.org.

Revise the content structure and architecture of the user interface of GBIF.org with the aims of

- Accommodating the extensive feedback received since launch of the new GBIF.org platform in 2013, and captured in more than 1000 issues
- Improving the user experience, particularly around a site-wide search function and content navigation
- Providing better support for mobile devices (e.g. responsive width)
- Accommodating support for multilingual interface
- Improving performance by providing architecture for a cloud-based deployment
- Providing architecture for software to allow for rapid enhancement and continuous deployment with new features

Within this implementation, GBIF will enhance the public GBIF API to expose content such as news and resources.

1.2 – Research and develop services to support sample-based datasets

Note: this activity provides the informatics components to complement the Content activity 3.4: sample-based data promotion and support

With the release of the IPT v2.3 and supporting data standards in 2015, GBIF has the ability to mobilize sample-based datasets and register them in GBIF.org. It is expected that with targeted data mobilization effort and promotion, GBIF will mobilize significant amounts of content (see activity 3.4). During 2016, GBIF will research and progressively incorporate services to enable discovery, visualization and access to such data. Subject to feedback during the mobilization process, such developments may include:

- Services to monitor trends in the number and types of sample based datasets mobilized
- Visualizations showing the temporal and geographic coverage of sample based datasets
- Filters to query and access individual records documenting sampling events
- Studies into the feasibility of offering visualizations of the relative abundance of taxa at sites under consistent monitoring

1.3 – Implement a mechanism to support long-term data preservation

Note: this activity provides informatics support for Content activity 3.3: reference datasets.

GBIF is an effective network for publishing datasets for reuse, but lacks a consistent long-term data preservation mechanism, which leaves data at risk of being lost. To address this, GBIF seeks a mechanism to offer redundant storage across multiple data centres and reduce the risk of data loss under technical (or other) failure scenarios. The [DataONE \(Data Observation Network for Earth\)](#) infrastructure is an attractive option to provide this as common data standards are in use by both GBIF and DataONE. GBIF will develop the technical connections so that any dataset published through GBIF can be archived in the DataONE network to provide redundancy, and will increase the visibility of the data to wider audiences. It is anticipated that the connection will also help in making the monitoring data already archived in DataONE more readily accessible to users of the GBIF network.

Once the workflow for secure storage is operational, GBIF will offer a repository for users to upload and archive unstructured and semi-structured datasets – such as datasets representing cleaned versions of downloaded GBIF records that are used in publications and the ‘reference datasets’ to be scoped in activity 3.3 below.

1.4 – Revise GBIF taxonomic data services

Note: this activity provides some of the informatics support for Content activity 3.1: improving data fitness for key user groups.

The [GBIF backbone taxonomy](#) is one of the most encompassing taxonomies available. It is algorithmically assembled using the checklists mobilized through GBIF and drives the services that organize all content accessible through GBIF.org. During 2016 GBIF will enhance this in the following ways:

- Revise the API service to make a clear distinction between names and taxa
- Revise the underlying data models to link names to the literature documenting the species
- Work with key stakeholders, such as maintainers of the large nomenclature databases, to find a model of credit that would satisfy integration of their databases

- Provide a rapid means of correction issues within the taxonomy
- Publish regular versioned editions of the taxonomy, reorganizing occurrence content with each new version

Such changes will improve the quality of all data accessible through GBIF.org for all users, and thus complement activity 3.1 on improving data fitness for key use groups.

Funds were allocated during the 2015 Work Programme for a workshop to promote delivery of a shared taxonomic framework with other biodiversity informatics initiatives. During 2015, GBIF has been in discussion particularly with the Catalogue of Life, Encyclopedia of Life, Biodiversity Heritage Library and Barcode of Life Database. These partners share a vision for collaboration to accelerate delivery of a common digital taxonomy or checklist including all species and mechanisms for community curation and correction. The workshop funds will be used late in 2015 or early 2016 to progress this agreement as an implementation plan.

1.5 – Operate cloud-based IPT deployments

Note: this activity provides informatics support to Engagement activity 2.1: Biodiversity Informatics for Development

Installations of the GBIF IPT enable hundreds of institutions to publish datasets through the GBIF network. Increasingly data publishing institutions are sharing installations of the tool, making use of the capabilities to clearly separate attribution of data ownership from the data hosting service, and benefiting from the reduction in maintenance costs. Building on this trend, the GBIF Secretariat will offer cloud-based deployments of the IPT specifically for institutions involved in data mobilization projects, but without the technical capacity to run installations robustly. This will apply to GBIF Participant nodes wishing to make use of such a service, and will be especially useful as a means to support data mobilization through the Biodiversity Information for Development programme (activity 2.1).

The perceived benefit of this model is that investment in data mobilization can be spent exclusively in digitization, advocacy, training and data management activities. Users of the cloud services will receive robust data hosting solutions at no cost, with the option of migrating installations in the future as technical capacity increases.

A small budget has been allocated from BID funding for Helpdesk services to support hosting costs and some technical support for this work. Most of these costs represent the work of Secretariat informatics staff to support deployment and any necessary adaptation of the IPT software. The Secretariat will explore opportunities to use appropriate network and hardware resources offered by Participants in preference to commercial solutions.

1.6 – Provide publishers with list of papers citing data from their datasets

During 2014 and 2015, GBIF has established and deployed clear mechanisms for data citations built around Digital Object Identifiers (DOIs). All datasets published through GBIF.org and all data downloads receive a DOI. Users downloading data receive a recommended citation including a DOI that resolves to a page listing each source dataset. Guidelines clarifying recommended practice for citing data used in scientific publications are being incorporated into the revised GBIF data use agreement (see 2015 progress above).

Working with [DataCite](#) and scholarly publishers during 2016, GBIF will start mining the publications to link cited use of data to the source datasets. A service will be deployed that allows GBIF publishers to discover which papers have cited use of their data. The intention is to help strengthen the perceived importance of investing in open data infrastructure and digitization projects, by clearly demonstrating the scientific use such activities enable.

Milestones

Item	Due
GBIF becomes an operational member node of the DataONE network (with GBIF data discoverable through DataONE)	Mar 16
First cloud-based IPT deployment	Apr 16
GBIF operates a repository for datasets backed by DataONE storage	Jun 16
Revised taxonomic data services (separation of names and taxa, links to publications, etc.)	Sep 16
Implementation of GBIF DataONE member node for data archival	Sep 16
Newly deployed (re-engineered) GBIF.org site	Oct 16
GBIF.org includes initial visualisations and services for sample-based data	Oct 16
Discovery service for published papers known to have used particular datasets	Dec 16

Budget

Item	Costs (K€)
Stream 1 (Informatics) 2016 ACTIVITIES	
Staff salary <i>Includes supplementary funding (reducing costs for core funds to €992K):</i> <ul style="list-style-type: none"> · EU BON €25K · EMODnet 2 €10K 	1027
Operating Expenses Cloud-based IPT repository <i>Fees for trial operation of a cloud-based service for an IPT repository (from BID funds)</i>	5
PROPOSED ADDITIONAL ACTIVITIES (if funds increase)	
Additional software developer <i>Extra capability to deliver outstanding enhancements</i>	90

Stream 2: Advance the Engagement – Communication and implementation

Key Deliverables

From GBIF Work Programme 2014-2016, as approved at GB20:

- Comprehensive communication strategy
- Essential documents and information resources for GBIF stakeholder groups
- Guidance and planning for national biodiversity information facilities
- Improved collaboration with **organizational participants**
- 2nd **Global Biodiversity Informatics Conference** to advance GBIF-compliant data and infrastructure
- Investigation of model for building **Supplementary Funds** for GBIF-related activity
- Regular **update reporting** on content and research value for GBIF national Participants

2015 Progress

The following progress can be reported against Work Programme deliverables:

- **Communication strategy**
 - [Strategy document](#) revised following community consultation and in use for guiding development of GBIF.org revision and associated communication products
- **Essential documents and information resources**
 - GBIF [information brochure](#) updated and available in six UN languages (by GB22)
 - GBIF [data publishing documentation](#) updated and consolidated (by end 2015)
 - 'How to' guide on publishing sample-based data available in English, French, Spanish (by GB22)
- **National biodiversity information facilities**
 - Capacity self-assessment guidelines developed and tested by selected nodes (release by GB22)
 - [Guidance document](#) on establishing an effective GBIF Participant node available in English, French, Portuguese and Spanish
- **Organizational participants**
 - Memorandum of Cooperation signed with [Bioversity International](#) on data mobilization for agrobiodiversity
 - Plan agreed with [OBIS \(UNESCO/IOC\)](#) for integration of updated marine datasets into GBIF via regional OBIS nodes
- **Global Biodiversity Informatics Conference**
 - Deferred following discussion at GB20
- **Supplementary Funds**
 - [Biodiversity Information for Development \(BID\)](#) programme approved and, communicated, structures set up within GBIF Secretariat to manage programme

- Survey conducted on biodiversity data needs in Africa, to support first phase of BID
- [Africa Rising conference](#) involving 100 participants from 21 African countries, launches declaration and action plan on biodiversity information for sustainable development in
- [First call for project proposals in Africa](#) prepared and launched (expected September)
- [Biodiversity Information Fund for Asia \(BIFA\)](#) launched with funds provided by Government of Japan, final project selection expected by November
- [Eye on Earth](#) project funded by Abu Dhabi Environment Agency engages regulators and consultants in Arabian Gulf region to publish datasets associated with environmental impact assessment through [workshop in Oman](#) (September)
- **Update Reporting**
 - Following consultations on a prototype national report, automated national reports with six-monthly updates implemented for all countries (expected delivery for GB22)

In addition, the GBIF Secretariat made progress in the following areas:

- **Capacity enhancement**
 - New [GBIF capacity enhancement framework](#) released after public review
 - Six collaborative projects co-funded through [2015 support call](#)
 - Past GBIF capacity enhancement actions evaluated (report for GB22)
 - Draft GBIF regional strategy developed by NSG for discussion at GB22
- **Regional node meetings**
 - Seventh European regional nodes meeting in Paris (hosted by GBIF France, Participant funded)
 - Sixth Asian regional nodes meeting in Taiwan (funded by BIFA and TaiBIF)
- **Global nodes events**
 - [Nodes training event](#) and global nodes meeting held in Madagascar associated with GB22
- **Strategic Plan development**
 - Major strategic priorities for GBIF 2017-2021 established through [community consultation](#)
 - Draft GBIF Strategic Plan 2017-2021 presented to GB22 for approval
- **External engagement**
 - Support for [IPBES Task Force on Knowledge and Data](#) in developing guidelines on handling of data in IPBES assessments
 - Memoranda of Cooperation (MoCs) developed with IUCN and Conservation of Arctic Flora and Fauna (CAFF) – signature of both anticipated before end 2015

Contributions to Stream 2 from GBIF Participants during 2014 include:

- Translation of documentation by GBIF Portugal and SiBBr (PT), GBIF France and BeBIF (FR), TaiBIF (CN) and JBIF (JP), review of Arabic translation by GBIF Mauritania
- Funding of European regional nodes meeting by GBIF France
- Co-funding of Asian regional nodes meeting by TaiBIF
- Funding for Biodiversity Information Fund for Asia by Government of Japan
- Co-organisation of Africa Rising conference by SANBI

2016 Plans

In addition to providing ongoing operational support to GBIF Participants and nodes, and participating in international fora such as the CBD, IPBES and GEO BON, GBIF will carry out the following activities during 2016.

Only limited funds are available for use in 2016 to support regional meetings and the GBIF capacity enhancement call. The Secretariat will prioritise these areas for allocation of any additional funds or underexpenditure in other areas. It should however be noted that the Biodiversity Information for Development (BID) activities also significantly address these needs within Africa, and the funding from Japan for the Biodiversity Information Fund for Asia (BIFA) contribute to such needs in Asia.

2.1 – Biodiversity Information for Development (BID)

Note: This activity includes dependency on Informatics activity 1.5: operate cloud-based IPT deployments; it also represents a significant investment in Stream 3 (Content) relating to data mobilization.

The [BID programme](#) will issue the first call for proposals for Africa in 2015, with a potential total of €900,000 in funding assigned to this call, and the selected projects are expected to start their activities in early 2016. This call for proposals will favour projects that:

- Mobilize biodiversity data related to protected areas, threatened species and invasive alien species
- Generate best practices in digitizing natural history collections and/or in mobilizing sample-based biodiversity data, species occurrence data, checklist data, or dataset descriptions
- Engage users of biodiversity data in producing guiding examples of the integration of biodiversity data into decision-making and research, and ensure the mobilized data are suited to the user's needs
- Develop lasting national, regional or thematic networks to support ongoing data sharing and reuse

Further calls for proposals will be issued to the Caribbean and Pacific regions in 2016. The specific themes and grant types to be included in the calls for proposals for the Caribbean and Pacific will be informed by commissioned studies and regional meetings, but will be broadly based on the model for the call for proposals issued for Africa.

The first BID capacity enhancement workshop in Africa will also be scheduled during 2016, focussing on equipping funded projects under BID with the skills needed to publish biodiversity data.

These activities will require coordination and management from within the Secretariat and an additional administrative officer will be recruited using BID overhead funds to ensure the smooth operation of these and other activities.

2.2 – Strategy and updates for GBIF.org

This activity complements Informatics activity 1.1 – reengineer user interfaces of GBIF.org

During 2016, the Secretariat will develop and implement a content strategy to support the revisions to the structure and organization of GBIF.org. This data-driven, audience-focused approach will improve access, usefulness and discovery of targeted, relevant content on GBIF.org for the audiences identified in the [GBIF Communications Strategy](#). Priority will be given to materials that promote and enhance the value of GBIF, both as a research infrastructure and a global network, equipping node managers to support their networks, guiding data publishers and enabling optimal use of data accessed through GBIF.

Related changes will improve production and publishing workflows, increasing the efficiency of compiling, distributing, repurposing and curating news, featured datasets, and examples of data use in research and policy. These enhancements will also make it easier for GBIF network members to access, redistribute and translate regular Secretariat publications like the GBits newsletter, the Annual Report and the Science Review.

With a consultation on multilingual priorities planned for late 2015, the changes accommodate support for multiple languages, providing a systematic and coherent means of displaying translations of priority resources, both those commissioned by the Secretariat and created by the network's active language communities.

Funds have been reserved to support key translation needs during 2016, partly from GBIF core funds and partly from BID funding (reflecting the need to support BID activities with timely documents in relevant languages). Where possible, translation needs will be met using skills within the GBIF network. When contracted work is required to complete translations, priority will again be given to use of expertise within the network.

2.3 – GBIF regional activities

During 2016, GBIF will build on recommendations made in the review of the network's regional strategy carried out by the Nodes Steering Group during 2015, and further discussed by Participants in a dedication session within GB22.

GBIF has limited funds to provide partial support for regional meetings during the year. The Secretariat welcomes offers from Participants to organize or provide further funding to support meetings in their regions, and encourages Participants to consider closely what strategy and deliverables may be most relevant for their region to target.

Given the limited funds, priority will be given to the most cost-effective options for organizing meetings alongside other regional events. Proposals for any adjustments to governance arrangements relating to regional activity will be presented to GB23.

2.4 – GBIF capacity enhancement call

The GBIF Capacity Support Programme has supported nine projects during its first two years of existence. GBIF.org is being updated to include new project pages highlighting the achievements, products and best practices identified during the implementation of each project. These pages will be rolled out for GB22 and further developed during 2016.

The 2016 capacity enhancement call has limited funds but will continue this programme, taking account of separate opportunities for project funding in specific regions under the BID and BIFA programmes. The selection of projects will take place earlier in the year than in previous calls, to enable better alignment with Participants' annual plans.

2.5 – Integrated training package

Many groups have defined curricula for the young discipline of biodiversity informatics. GBIF has been among those groups, recognizing early in its development that reinforcing capacity in this field was vital to achieve its mission.

Recent efforts to coordinate these parallel initiatives have crystalized in the TDWG interest group '[Biodiversity Informatics Curriculum](#)', which is already supported by many members of the GBIF community. The GBIF Secretariat is co-leading this group and in 2016 plans to strengthen GBIF's current capacity enhancement approach through an integrated training package, making sure it is aligned with similar efforts worldwide.

As part of this work, the Secretariat would like to establish a certification schema based on open standards (Mozilla Open Badges) that can be used inside and outside the GBIF community. A formal certification schema will allow public recognition of skills and knowledge across boundaries, help to identify talent and enlist qualified specialists to GBIF-related activities, and enable better ways to measure the real impact of the investments in training initiatives, among other benefits.

Milestones

Item	Due
Start of projects selected under 2016 capacity enhancement call	Feb 16
First BID funded projects start in Africa	Apr 16
Caribbean and Pacific regional meetings for BID	By Jun 16
BID call for project proposals in Caribbean and Pacific	Jun 16
BID capacity enhancement workshop in Africa	Aug 16
GBIF.org includes basic information about GBIF in at least the six UN languages	Sep 16
Presentation of proposed GBIF biodiversity informatics curriculum at TDWG 2016	Oct 16

Budget

Item	Costs (K€)
Stream 2 (Engagement) 2016 ACTIVITIES	
Staff salary – combined with Stream 3 ¹ <i>Includes supplementary funding estimated at €50K funding (reducing costs for core funds to €711K)</i>	761
Operating Expenses BID activities (funded projects, training, meetings) <i>Covered from BID supplementary funds</i>	990
Regional Nodes meetings <i>Costs to support organising and planning regional meetings</i>	35
Capacity enhancement call <i>Activities to be prioritised in line with the Capacity Enhancement Strategy</i>	35
Authoring/translating key documents <i>Contracts to GBIF Participants or other experts to assist with delivery or translation of priority documents (includes €10K from BID funds for supporting relevant translations)</i>	20
Administrative support for BID and other funded activities <i>Additional support within Secretariat to support BID activities (from BID overheads)</i>	60
PROPOSED ADDITIONAL ACTIVITIES (if funds increase)	
Additional funding for regional meetings <i>Expanded funding to support additional projects</i>	25
Supplementary funding: Capacity enhancement <i>Any additional funds specifically allocated to facilitate capacity enhancement</i>	Open

Stream 3: Advance the Content – Relevance and fitness-for-use

Key Deliverables

- Engagement of expert communities to form **fitness-for-use working groups**
- **Metrics and indicators** for assessing relevance of GBIF-mobilised data
- Identification of content mobilization priorities
- Evaluation of models for creating and curating **reference data sets** within GBIF network
- Quality and fitness-for-use reporting through GBIF.org and services

¹ Salary costs for Stream 2 (Engagement) and Stream 3 (Content) staff salary have been combined here for reasons of staff privacy with smaller numbers in Stream 3 – see *Summary Budget 2016* below.

- Support for publishing metadata for undigitized collections

2015 Progress

The following progress can be reported against Work Programme deliverables:

- **Fitness-for-use working groups**
 - Two [fitness for use task groups](#) established to report data needs of agrobiodiversity and distribution modelling communities
 - Data needs from each community captured through questionnaires
 - Interim progress reports with initial recommendations prepared for GB22
- **Metrics and indicators**
 - Historical [analytics of data publishing](#) at national level available and updated
- **Content mobilization priorities**
 - Implementation awaiting recommendations of fitness-for-use task groups and gap analysis work (see activity 3.5)
- **Reference datasets**
 - Activities deferred to 2016 (see activity 3.3)
- **Quality and fitness-for-use reporting**
 - Activities deferred pending recommendations of fitness-for-use task groups (see activity 3.1)
- **Metadata for undigitised collections**
 - [Task force established](#) on 'accelerating the discovery of bio-collections data'
 - Key issues around metadata capture of collections information analysed through global survey (by end 2015)
 - Strategy and action plan for metadata approach to data mobilization (by Nov)

In addition, the GBIF Secretariat made progress in the following areas:

- GBIF data gap analysis guidelines drafted and reviewed (expected by GB22) to support work on gap-filling strategy (see activity 3.5 below)
- Following the TDWG 2014 meeting in Jönköping, the GBIF/TDWG Biodiversity Data Quality Interest Group chaired by Arthur Chapman and Antonio Saraiva (<http://community.gbif.org/pg/groups/21292/>) started three Task Groups, all with GBIF involvement: TG1 Framework on Data Quality, TG2 Catalog of Tools for Data Quality, and TG3 Use Case Library for Data Quality. The Interest Group will report at TDWG 2015.

2016 Plans

In addition to providing ongoing operational support to data publishers and to users of data through the GBIF network, and significant content mobilization through BID funding (see Engagement activity 2.1 – *Biodiversity Informatics for Development (BID)*) GBIF will carry out the following activities during 2016.

3.1 – Improving data fitness for key user groups

GBIF data are aggregated from many sources and are consequently heterogeneous, varying in fitness for various uses. During 2015, GBIF established two task groups on data fitness-for-use, in agrobiodiversity research and in distribution modelling respectively, to document

how these communities use GBIF data and to understand their data quality demands. The resulting reports will inform data mobilization, data processing and improvements to GBIF.org.

Based on the recommendations of the task groups, GBIF aims to formalize the needs of these communities as far as possible through 'data use profiles', defined as use-specific filters of data published through GBIF.org. These profiles may be used to facilitate rapid assessment and download of data exhibiting the preferred characteristics for a particular use or analysis, and can also form the basis for GBIF to evaluate progress in mobilizing new data or improving existing data to support key needs.

During 2016, GBIF will expand this work to incorporate more key areas of use, provisionally focussing on: 1) DNA evidence for names and occurrences, 2) invasive alien species and 3) marine biodiversity. Depending on resources, these groups will operate through a combination of face-to-face meetings and online or remote collaboration.

3.2 – Strategy for accelerating discovery of biocollections

The [GBIF Task Force on Accelerating the Discovery of Bio-collections Data](#) has been meeting during 2015 to develop recommendations on best practices around documentation of natural history collections via metadata, and the use of such metadata assessments to establish roadmaps for further digitization of the collections. The recommendations from this group will be publicized during 2016 and incorporated into GBIF tools, documentation and training materials for collection digitisation, thus contributing to Engagement activities 2.1, 2.4 and 2.5.

3.3 – Reference datasets

Note: Support for this activity will be provided by Informatics activity 1.3: implement a mechanism to support long-term data preservation.

As GBIF increases the use of Digital Object Identifiers (DOIs) to identify datasets and unique identifiers for individual records in each dataset, it will be possible for users to download and clean data from the GBIF network and to deliver a derived reference dataset representing the data used within a scientific manuscript. GBIF aims to facilitate deposition of such reference datasets in repositories to increase the transparent use of GBIF data in research (see activity 1.3), and to give appropriate credit to data publishers for their contribution. Evaluating the use of datasets and records in constructing such reference datasets will also offer an additional mechanism for assessing the value of particular data records for a given use and even to derive annotations and corrections for the original data.

During 2016, GBIF will include test archival of reference datasets with DOIs and explore options for identifying datasets of high-value in key areas such as IPBES assessments. This approach will then be developed to offer best practices for global and national data publishers and research funding bodies.

3.4 – Sample-based data promotion and support

Note: Support for this activity will be provided by Informatics activity 1.3: research and develop services to support sample-based datasets

The GBIF Secretariat will work with nodes and publishers to identify existing datasets that are better suited to be sample-based datasets and migrate them. This data management activity will help in perception of overall data quality, in particular by identifying and categorizing the so-called *gridded datasets*.

The following activity is planned to promote and support mobilization of sample-based data:

- Identify and engage with nodes and holders of sample data types representing selected ecosystems, geographic and taxonomic scope, and sampling methods; capturing experiences from early mobilization of these datasets to inform future standard development;
- Further document training needs and develop essential documentation for mobilizing sample-based data;
- Capture user needs for sample-based data in research and policy, including for example Essential Biodiversity Variables developed for GEO BON, indicators for the CBD and IPBES assessments, modelling.

3.5 – Foundations for a GBIF gap-filling strategy

The draft GBIF Strategic Plan 2017-2021 recognizes that a key requirement for long-term enhancement of the quality of the data products offered by the GBIF network is to understand the extent and significance of gaps in published data, including gaps in geospatial, temporal, taxonomic and other aspects of coverage which may be significant for analysis. From a different standpoint, this issue can be explored by assessing the extent of our ignorance of detail in biodiversity patterns.

In preparation for the next GBIF strategic planning period, work will be carried out in 2016 to begin evaluation of requirements for GBIF processes and models to assess gaps. Relevant recommendations are expected from the Fitness-for-use Working Group on Distribution Modeling and these will be incorporated in this assessment.

The 2016 Ebbe Nielsen Challenge will specifically focus on development of viable models and tools which could be integrated into GBIF workflows to identify and assess the significance of gaps in data coverage. This represents a development of the Challenge from the open-ended nature of the first call in 2015, focussing the available funds on seeking to solve a specific problem for GBIF in a timely fashion, engaging the wider community to contribute creative solutions.

Milestones

Item	Due
Workshop to identify priorities and goals for mapping gaps in GBIF data	Feb 16
2016 fitness-for-use working groups established	Mar 16
Best practice documentation and roadmaps for collection digitisation via metadata	May 16
Reference dataset processes included in GBIF.org	Sep 16

Ebbe Nielsen Challenge results around mapping gaps in GBIF data	Oct 16
Showcase range of sample-based datasets for testing through GBIF.org	Oct 16
First fitness-for-use data profiles implemented in GBIF.org	Dec 16

Budget

Item	Costs (K€)
Stream 3 (Content) 2016 ACTIVITIES	
Staff salary – combined with Stream 2 ² <i>Includes supplementary funding estimated at €50K funding (reducing costs for core funds to €711K)</i>	761
Operating Expenses Fitness-for-use working groups <i>Funds to support establishing work of fitness-for-use working groups</i>	15
Recommendations from Task Force on Accelerating Discovery of Bio-collections Data <i>Guidance on metadata for collections and roadmaps for collection digitisation</i>	10
PROPOSED ADDITIONAL ACTIVITIES (if funds increase)	
Additional fitness-for-use working groups <i>Funding or organisation for working groups</i> GBIF Participants may take leadership in such activities in areas of special interest to their research communities	Open
Supplementary funding: Content mobilisation <i>Any additional funds specifically allocated to content mobilisation</i>	Open

² Salary costs for Stream 2 (Engagement) and Stream 3 (Content) staff salary have been combined here for reasons of staff privacy with smaller numbers in Stream 3 – see *Summary Budget 2016* below.

Summary Budget 2016

The following table summarises the 2016 budget in the context of this Work Programme. Further information is presented in the *GBIF Budget Proposal 2016*. The following should be noted:

1. All figures representing costs and funding are in thousands of Euro.
2. Costs and information on funding for activities has been split between **2016 Activities** and **Proposed Additional Activities** (priority activities that would be funded if revenue is higher), following the breakdown used elsewhere in this document.
3. **2016 Activities** includes activities planned for 2016, including all three Streams and costs for the Governing Board and committees, for operating the global Secretariat and for management and support of GBIF staff. The **Governance** and **Secretariat** lines follow the *GBIF Budget Proposal 2016* – see that document for further explanation. All salary lines include on-costs (insurance, recruitment costs, etc.) which are included within the GBIF annual budget and in GBIF financial reports as part of the Management costs.
4. Salary costs for Stream 2 and Stream 3 (and hence total costs for these streams) are combined for reasons of staff privacy with smaller numbers in Stream 3.
5. Operating expenses (staff travel and other operating expenses) incurred internally within the Secretariat at this stage are in Management costs rather than a particular Stream. At present, the financial system is not designed to attribute them to Streams.
6. Within **2016 Activities**, the items are ordered according to recommended priority (most important first). If available funds are too low to undertake all of these activities, it is proposed that the last items in the list should be the first that are deferred to future years. If on the other hand additional funds become available, the first items in **Proposed Additional Activities** are recommended for uptake.
7. Each line in **2016 Activities** indicates the expected source(s) for funding the activity, either from **Core** GBIF funds or from **Supplementary** funding from other sources. At present supplementary funds are identified only for two of these lines. If additional supplementary funds are made available and can be applied to any of the listed activities, it is proposed to transfer corresponding core funds to delivering more of the **Proposed Additional Activities**.

Item	Stream	Costs (K€)	Funding (K€)	
			Core	Supplementary
2016 ACTIVITIES				
Staff Salaries (Incl. staff-related costs)				
Management and support staff	-	670	670	-
Stream 1 (Informatics) Supplementary funding from: § EU BON €25K § EMODnet 2 €10K	1	1026	991	35
Streams 2 & 3 (Engagement & Content) Supplementary funding from: § BID €30K § BIFA €10K	2/3	760	720	40
Staff-related costs ^[4]	-	199	199	-
Operating				
Governance Governing Board, Committees, Ebbe Nielsen Challenge, YRA	-	111	111	-
Secretariat Facilities and operating costs, including travel costs Note also €23K additional funding received annually from Copenhagen University for hardware support	-	185	185	-
Stream activities (in priority order)				
BID activities (funded projects, training, meetings)	2	990	-	990
Regional Nodes meetings	2	35	35	-
Capacity enhancement call	2	35	35	-
Administrative support for BID, etc.	2	60	-	60
Cloud-based IPT repository	1	5	-	5

Authoring/translating key documents	2	20	10	10
Fitness-for-use working groups	3	15	15	-
Task Force on Accelerating Access to Biocollections Data	3	10	10	-
TOTAL 2016 ACTIVITIES		4121	2981	1140
PROPOSED ADDITIONAL ACTIVITY (if funds increase)				

Additional software developer	1	90	-	-
Additional funding for regional meetings <i>Expanded funding to support additional projects</i>	2	25	-	-
Supplementary funding: Capacity enhancement <i>Any additional funds specifically allocated to facilitate capacity enhancement</i>	2	Open	-	-
Additional fitness-for-use working groups <i>Funding or organisation for working groups</i> GBIF Participants may take leadership in such activities in areas of special interest to their research communities	3	Open	-	-
Supplementary funding: Content mobilisation <i>Any additional funds specifically allocated to content mobilisation</i>	3	Open	-	-