GBIF Global Biodiversity Information Facility

GBIF and the *alliance for biodiversity knowledge*

Joe Miller, GBIF Executive Secretary, jmiller@gbif.org

June 121, 2019

GBIF

GBIF—the Global Biodiversity Information Facility—is an international network and research infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth.

GBIF

- 1999 OECD Megscience Forum recommendation

GBIF

- 1999 OECD Megascience Forum recommendation
- 2001 Secretariat in Copenhagen

GBIF

- 1999 OECD Megscience Forum recommendation
- 2001 Secretariat in Copenhagen
- Governing Board, Executive Committee
  - Science Committee
  - Budget Committee
  - Nodes Steering Group

GBIF

GBIF SECRETARIAT

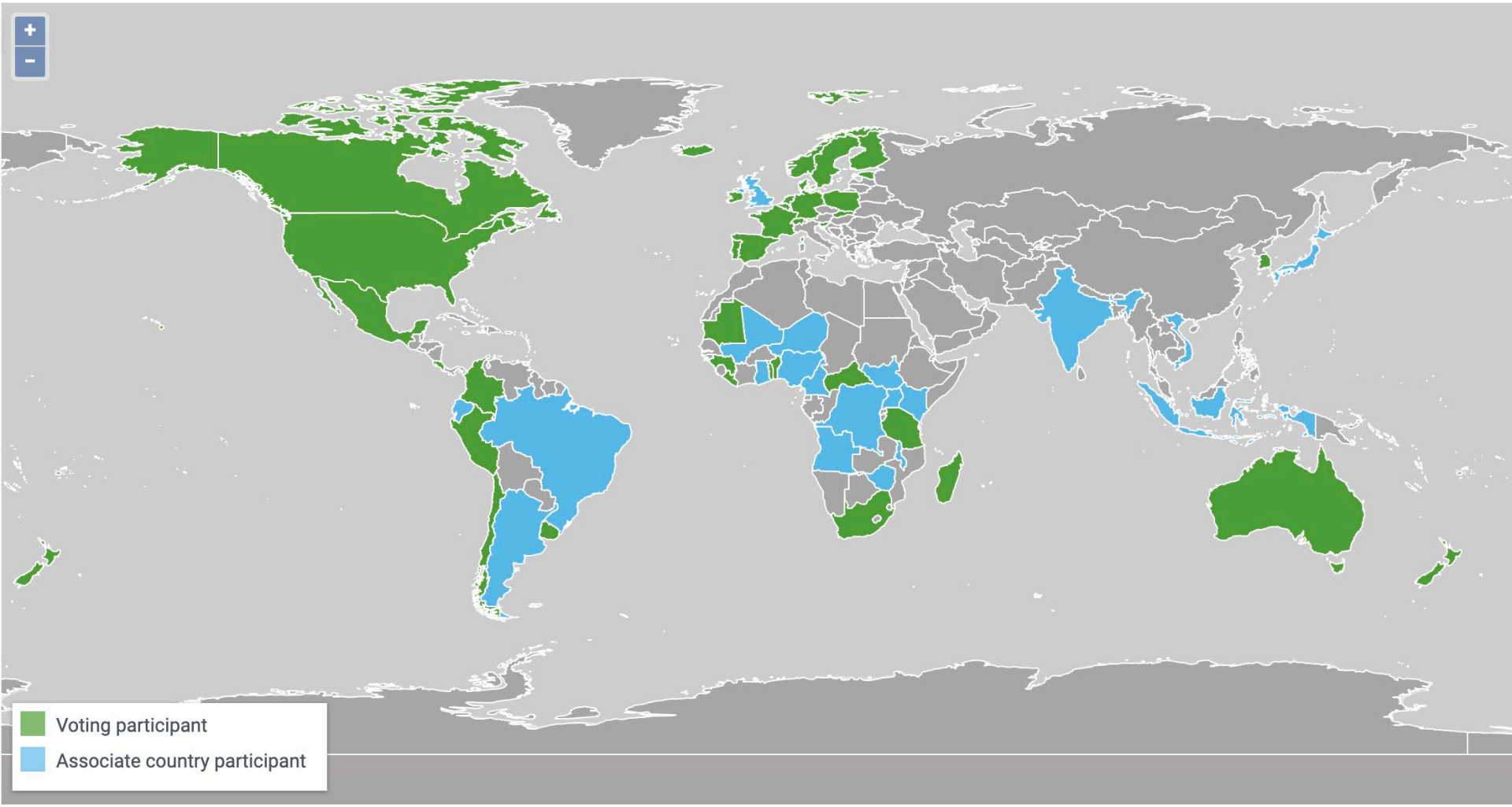
Participation and Engagement is responsible for operating the network of Participants and publishers, recruiting new members and enhancing the capacity of current ones.

Data Products is responsible for the quality and scientific value of the integrated data products produced by the GBIF network.

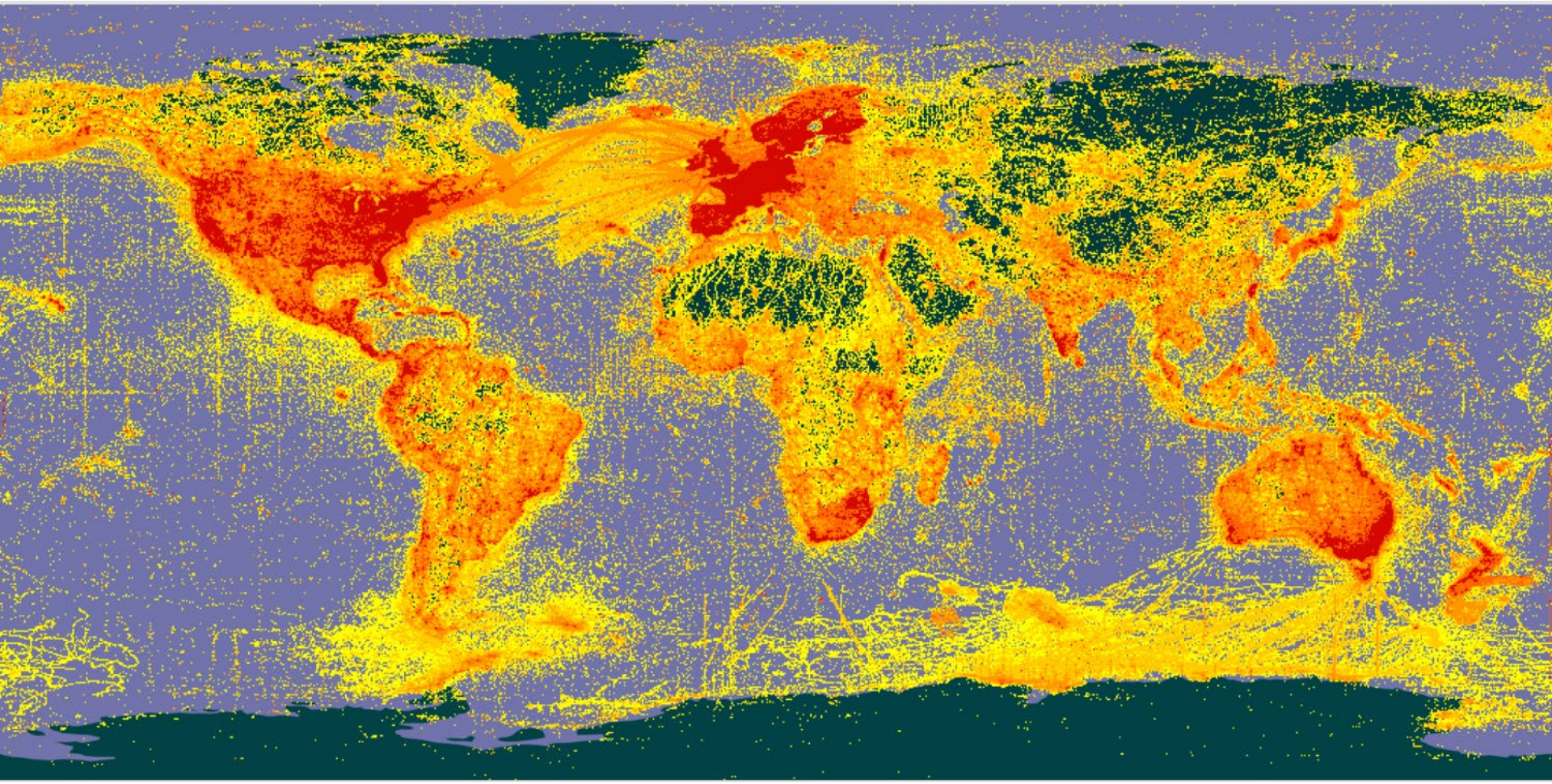
Informatics is responsible for data management, software development and the overall operation of the GBIF infrastructure.

Administration is responsible for maintaining both the network and the Secretariat's underlying operations and processes.

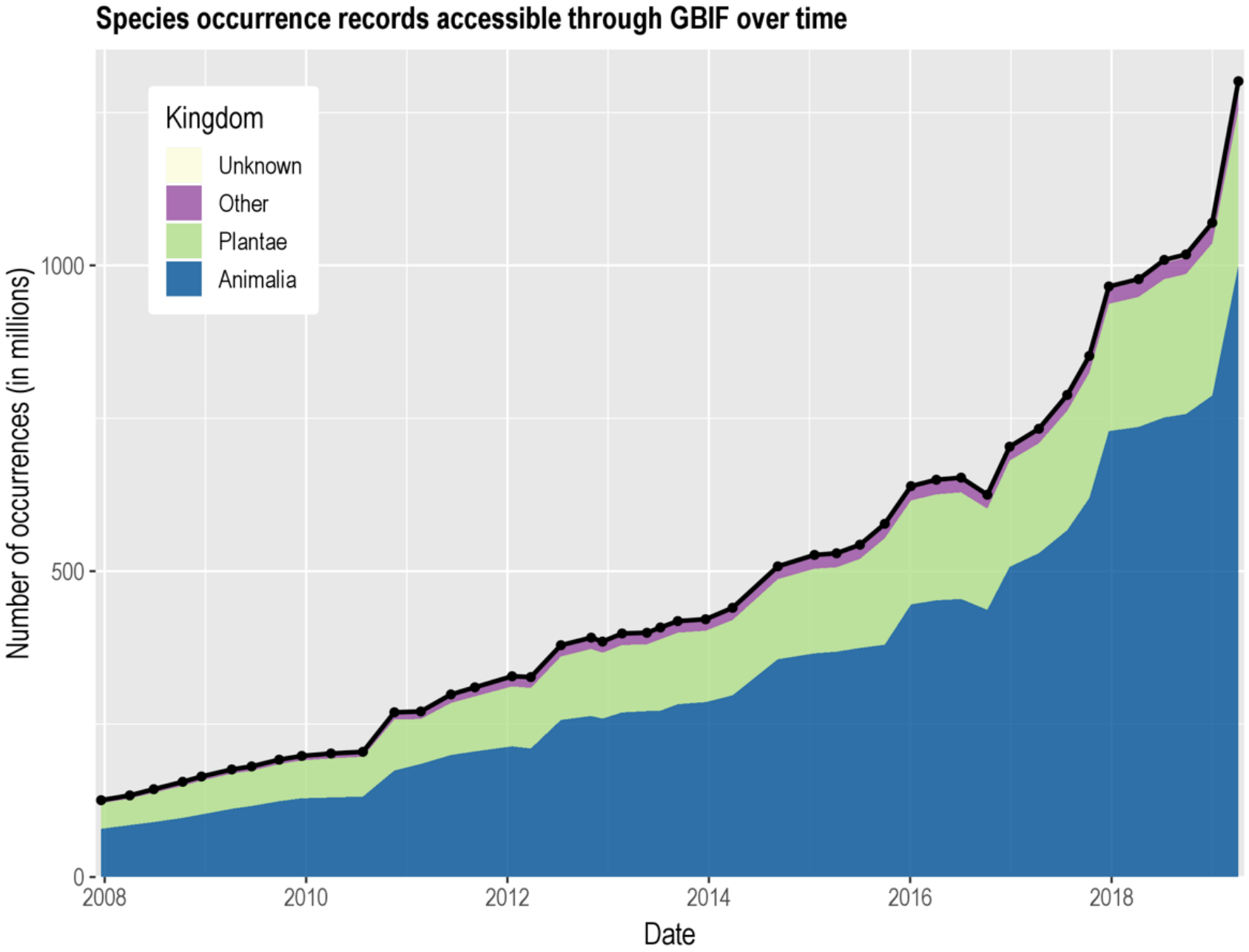
GBIF NODE NETWORK



39 VOTING PARTICIPANTS | 20 ASSOCIATE COUNTRY PARTICIPANTS | 38 OTHER ASSOCIATE PARTICIPANTS | 1,408 PUBLISHERS



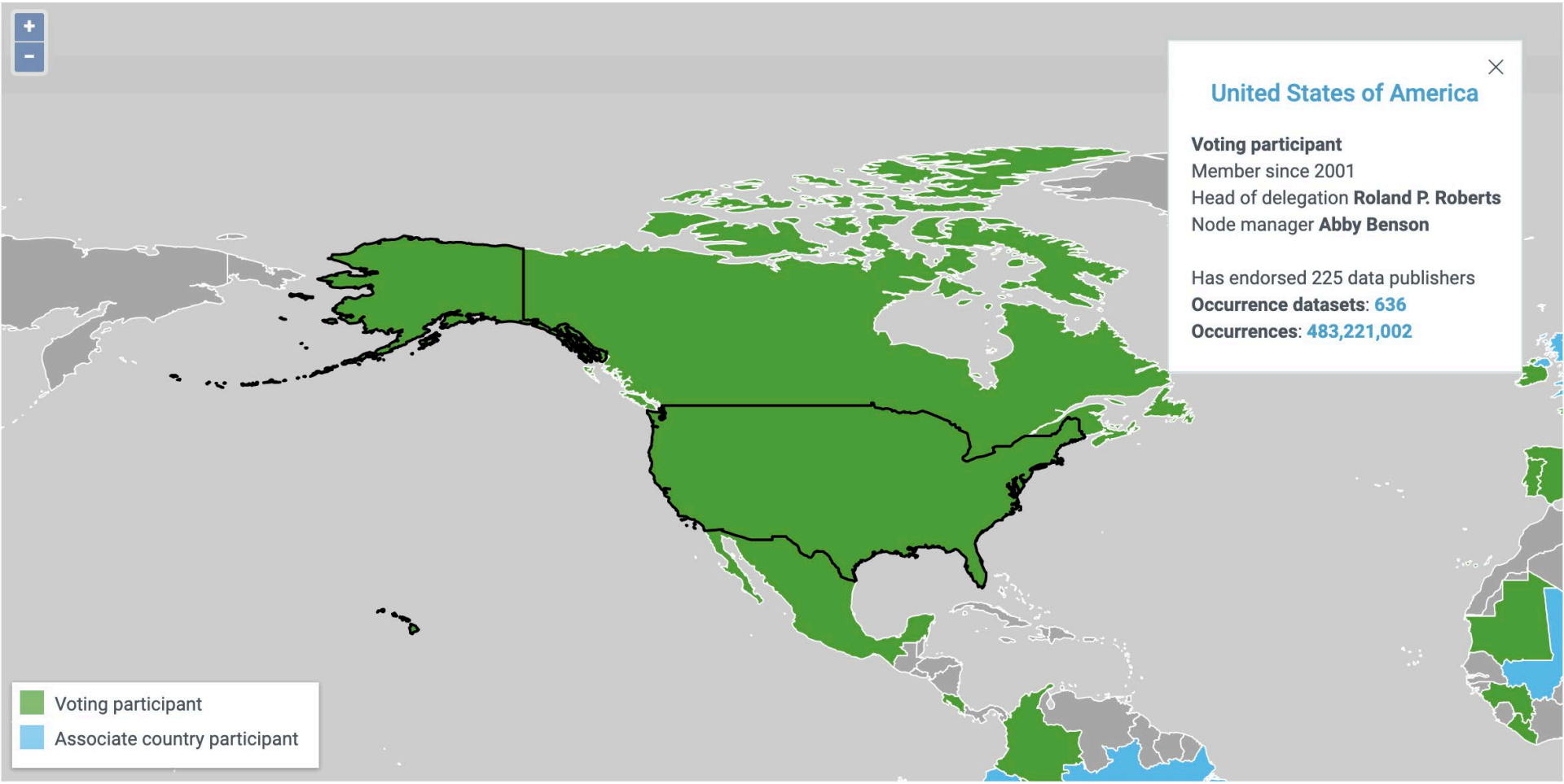
GBIF



175 million specimen

Citizen science and metagenomics

GBIF


United States of America
Voting participant
Member since 2001
Head of delegation Roland P. Roberts
Node manager Abby Benson
Has endorsed 225 data publishers
Occurrence datasets: 636
Occurrences: 483,221,002
Voting participant
Associate country participant


+
-
Canada
Voting participant
Member since 2001
Head of delegation Benoît Girard
Node manager James Macklin
Has endorsed 29 data publishers
Occurrence datasets: 186
Occurrences: 64,887,842
Voting participant
Associate country participant

GBIF IS THE COMMUNITY

- GBIF data is your data

GBIF IS THE COMMUNITY

- GBIF data is your data
- The GBIF Secretariat cannot unilaterally clean the data

GBIF IS THE COMMUNITY

- GBIF data is your data

- The GBIF Secretariat cannot unilaterally clean the data
- To make our data more usable we have to work together

GBIF IS THE COMMUNITY

- GBIF data is your data

- The GBIF Secretariat cannot unilaterally clean the data
- To make our data more usable we have to work together
- Incoming and outgoing pipelines
    - Recipocal annotation

GBIF SECRETARIAT

- Goal of annual work programme:

GBIF SECRETARIAT

- Goal of annual work programme:

Do things that help your work

GBIF SECRETARIAT

- Goal of annual work programme:

Do things that help your work

- Value proposition

- Fitness for use

DEVELOPING OUR VALUE PROPOSITION

- Showing impact through citations

CITATIONS AND USE

Use of data

- 155K user sessions per month
- 36 billion records downloaded per month
- 3,583 peer-reviewed articles citing use of GBIF-mediated data
- Topics: climate change, invasive species, human health, etc.

GETTING DATA

Typically, users request data on specific taxa, periods in time and/or geography

- Bluewhales (69 datasets)
- Birds in Japan (78 datasets)
- Reptiles in Africa between 1950 and 1980 (88 datasets)

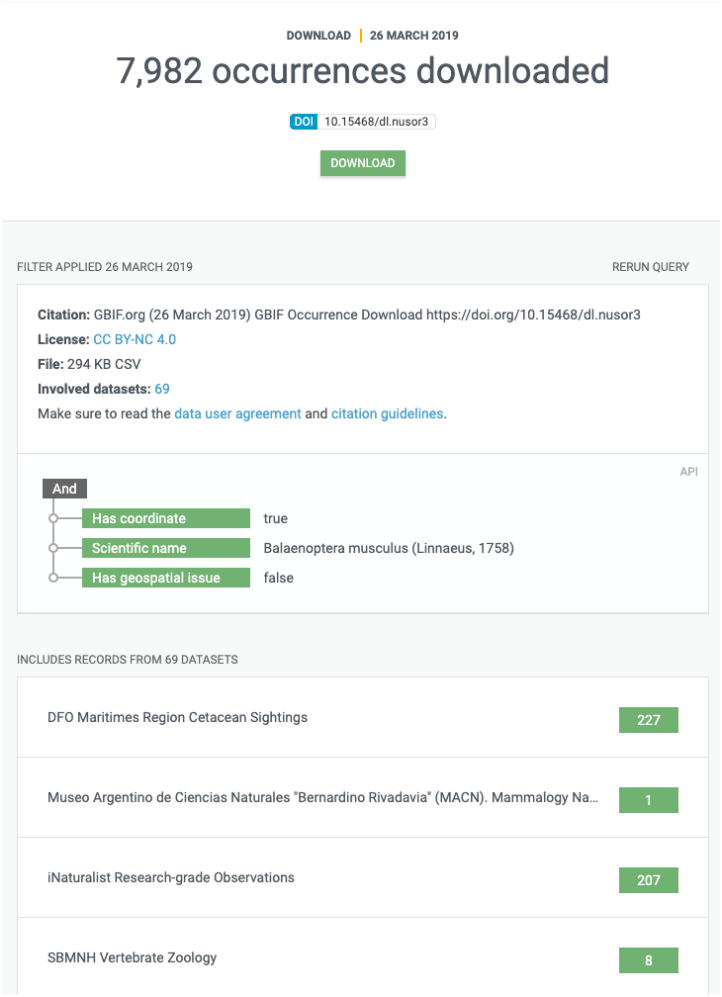
- How to cite use of data spanning 88 datasets?

GBIF

DOWNLOAD DOIS

Each download request is stored and is assigned a DOI resolving to a landing page

- Date and size of download
- Filters used in query
- Link to re-download
- Details of contributing datasets
- Relationships modeled in metadata



EXAMPLE

This study:

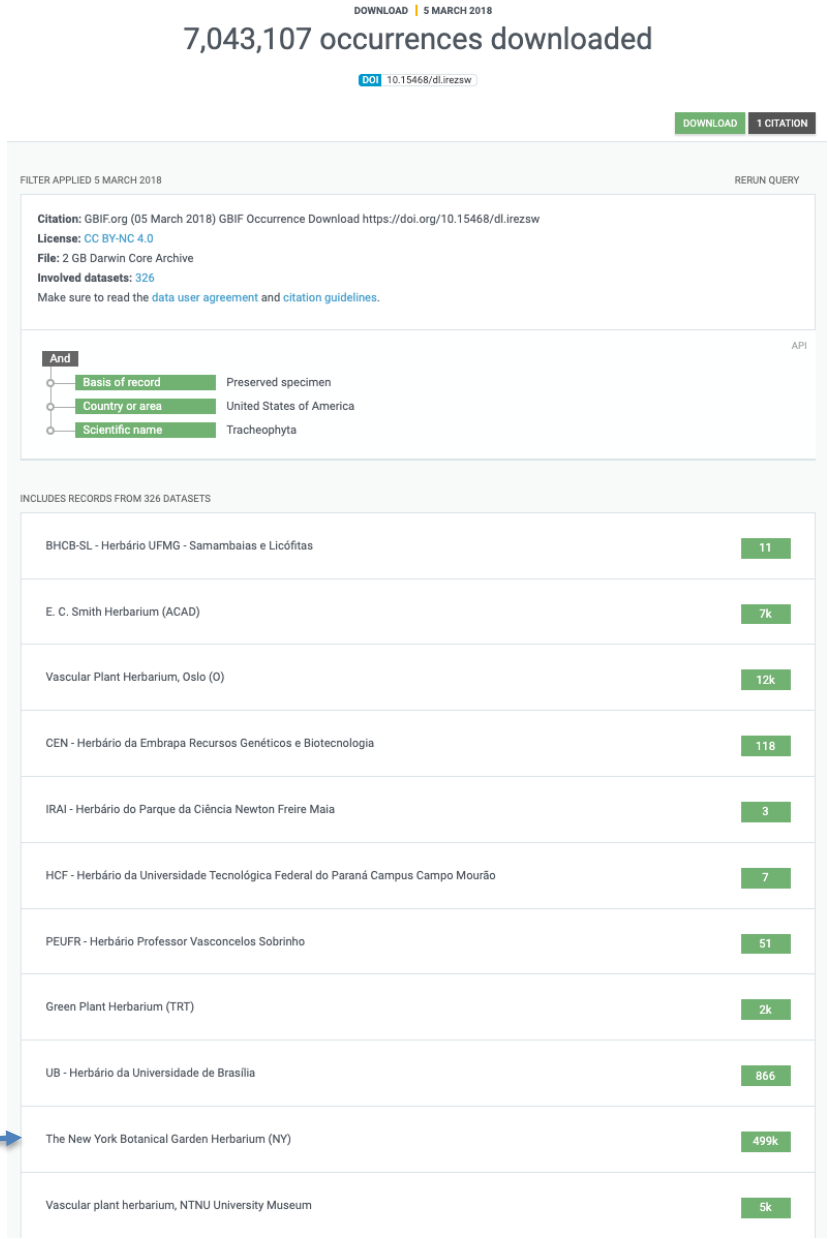
Monarch butterfly and milkweed declines substantially predate the use of genetically modified crops, J. H. Boyle, H. J. Dalgleish, J. R. Puzey

Proceedings of the National Academy of Sciences Feb 2019, 116 (8) 3006-3011; https://doi.org/10.1073/pnas.1811437116 (see https://www.pnas.org/content/116/8/3006#ref-38)

Used this data:

https://doi.org/10.15468/dl.irezsw

7,043,107 vascular plant occurrences, including almost 500,000 records from NYBG



OCCURRENCE DATASET | REGISTERED MARCH 13, 2009

Botany Division, Yale Peabody Museum

Published by Yale University Peabody Museum

Larry Gall • ✉ Patrick Sweeney

DATASET METRICS ACTIVITY ⬇ DOWNLOAD

151,097 OCCURRENCES 93 CITATIONS

Founded in 1864 by Daniel Cady Eaton from his personal library and plant collection, the Yale Herbarium is an internationally recognized repository with holdings of approximately 350,000 specimens from throughout the world. There are an estimated 3,000 type specimens. The collection is particularly rich in ferns, bryophytes and grasses, as well as in historically important materials from early botanical collectors. In addition, it was the herbarium of record for the flora of southern New England from 1864 until 1955, when that function passed to the University of Connecticut at Storrs.

Metadata last modified: June 8, 2019

Data last changed: May 24, 2019

Hosted by: Yale University Peabody Museum

License: CC0 1.0

How to cite DOI 10.15468/hrztgn

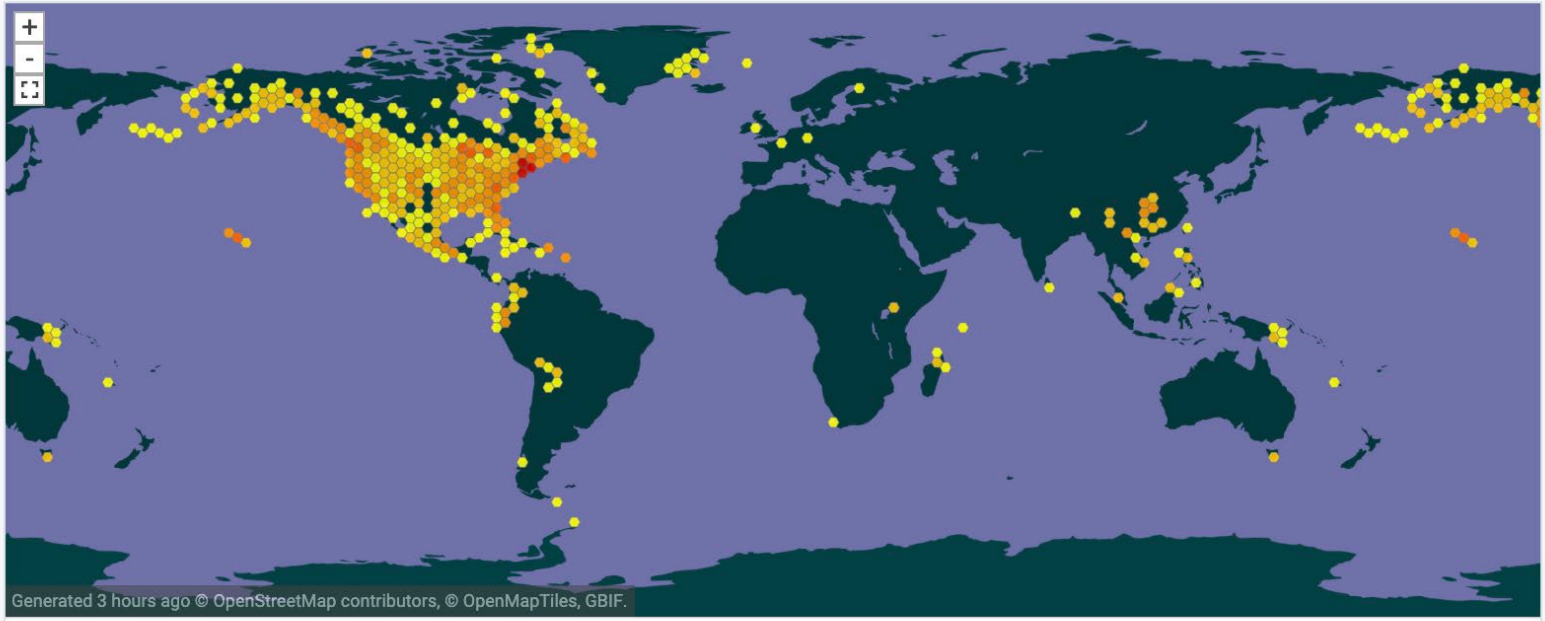
151,097 Occurrences

99% With taxon match

59% With coordinates

65% With year

89,542 GEOREFERENCED RECORDS



Read more about literature, how it's discovered and linked to GBIF-mediated data.

The grass subfamily Pooideae: Cretaceous–Palaeocene origin and climate-driven Cenozoic diversification

Literature

Schubert, M. Marcussen, T. Meseguer, A. Fjellheim, S. (2019) Global Ecology and Biogeography

Aim: Frost is among the most dramatic stresses a plant can experience, and complex physiological adaptations are needed to endure long periods of sub-zero temperatures. Owing to the need to evolve these complex adaptations, transitioning from tropical to temperate climates is regarded as difficult. …

Poaceae • Pooideae • diversification • evolution • grasses • niche transition

Journal article | Peer-reviewed

Data used in study DOI 10.15468/dl.h6fdbn DOI 10.15468/dl.kyz82l DOI 10.15468/dl.obmdq7 DOI 10.15468/dl.vzezkq DOI 10.15468/dl.zwbz44

Rates of niche and phenotype evolution lag behind diversification in a temperate radiation

Literature

Folk, R. Stubbs, R. Mort, M. Cellinese, N. Allen, J. Soltis, P. ... - (2019) Proceedings of the National Academy of Sciences

Environmental change can create opportunities for increased rates of lineage diversification, but continued species accumulation has been hypothesized to lead to slowdowns via competitive exclusion and niche partitioning. Such density-dependent models imply tight linkages between diversification and…

angiosperms • diversification • niche • phenotype • radiation

Journal article | Open access | Peer-reviewed

Data used in study DOI 10.15468/dl.9iecjg DOI 10.15468/dl.a9hz6l DOI 10.15468/dl.aucbrd DOI 10.15468/dl.bkcnxp DOI 10.15468/dl.ccuq1o DOI 10.15468/dl.cfz0nf DOI 10.15468/dl.egxdlt DOI 10.15468/dl.ftlsyi DOI 10.15468/dl.lbz680 DOI 10.15468/dl.lvk0wy DOI 10.15468/dl.pwvfdm DOI 10.15468/dl.t4rzb8 DOI 10.15468/dl.xdzqm6 DOI 10.15468/dl.ximwod DOI 10.15468/dl.xuuitg

Global distribution patterns of mycoheterotrophy

Literature

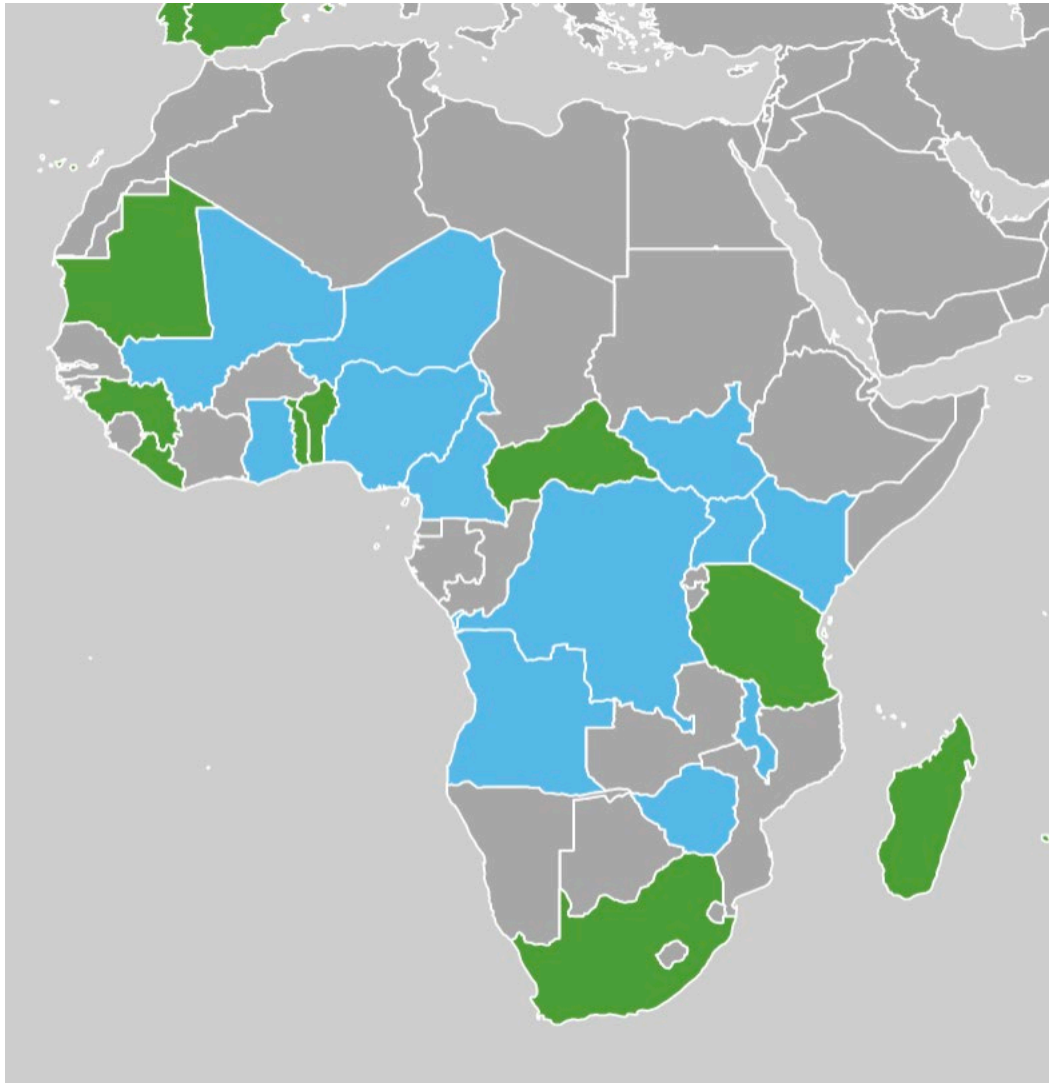
Gomes, S. van Bodegom, P. Merckx, V. Soudzilovskaia, N. (2019) Global Ecology and Biogeography

Aim: Mycoheterotrophy is a mode of life where plants cheat the mycorrhizal symbiosis, receiving carbon via their fungal partners. Despite being widespread, mycoheterotrophic plants are locally rare, hampering the understanding of their global environmental drivers. Here, we explore global environmen…

arbuscular mycorrhiza • cheaters • ectomycorrhiza • forest type • global distribution • mycoheterotrophic plants

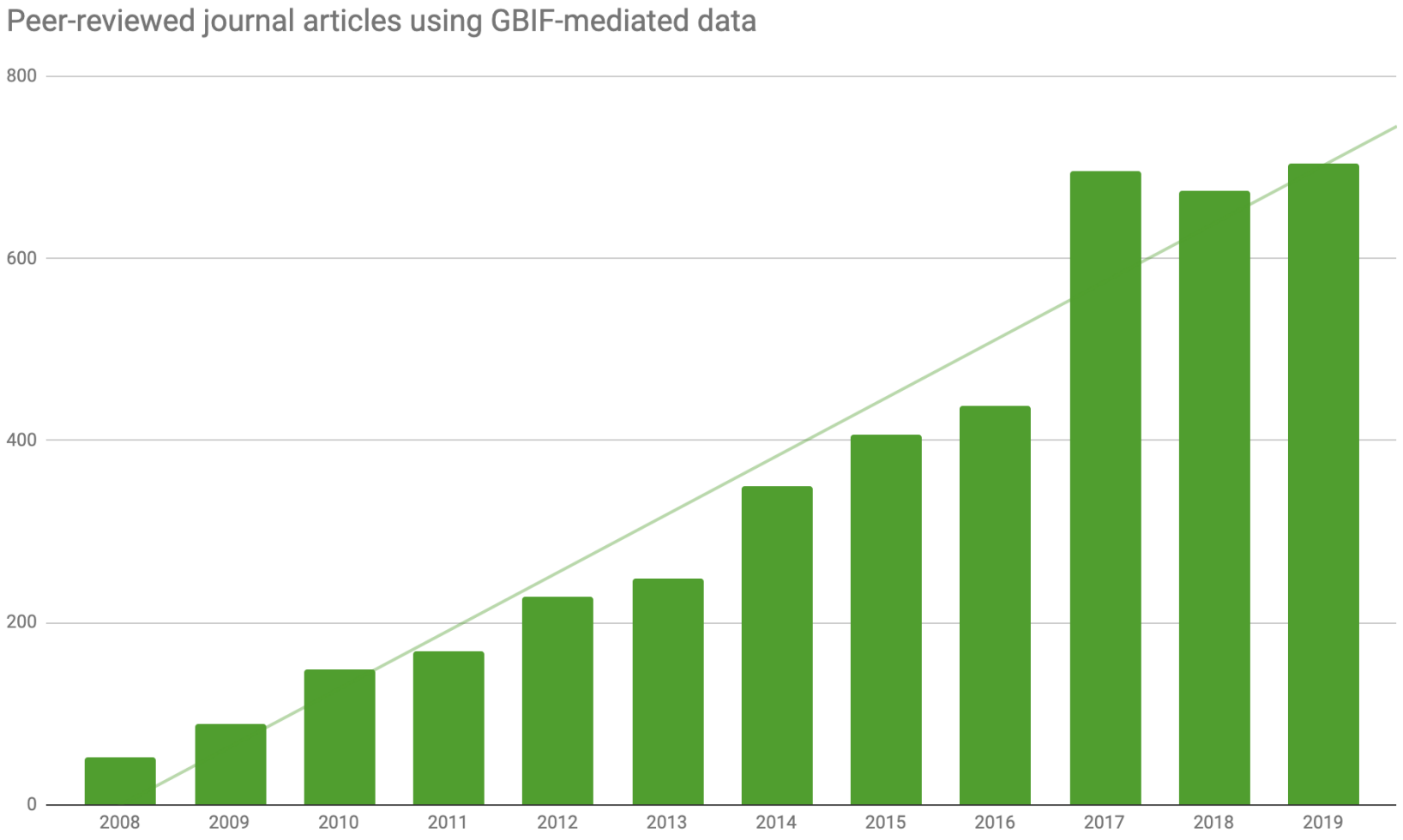
Journal article | Open access | Peer-reviewed

Data used in study DOI 10.15468/dl.0ldwjh DOI 10.15468/dl.4w1p9o DOI 10.15468/dl.6c3sgs DOI 10.15468/dl.8uhdjd DOI 10.15468/dl.8yia6c DOI 10.15468/dl.ae6rud DOI 10.15468/dl.bacvsb DOI 10.15468/dl.cijta2 DOI 10.15468/dl.ipjoqf DOI 10.15468/dl.m6inih DOI 10.15468/dl.qzm0kl DOI 10.15468/dl.suo1qb DOI 10.15468/dl.tomqwn DOI 10.15468/dl.zmadsa

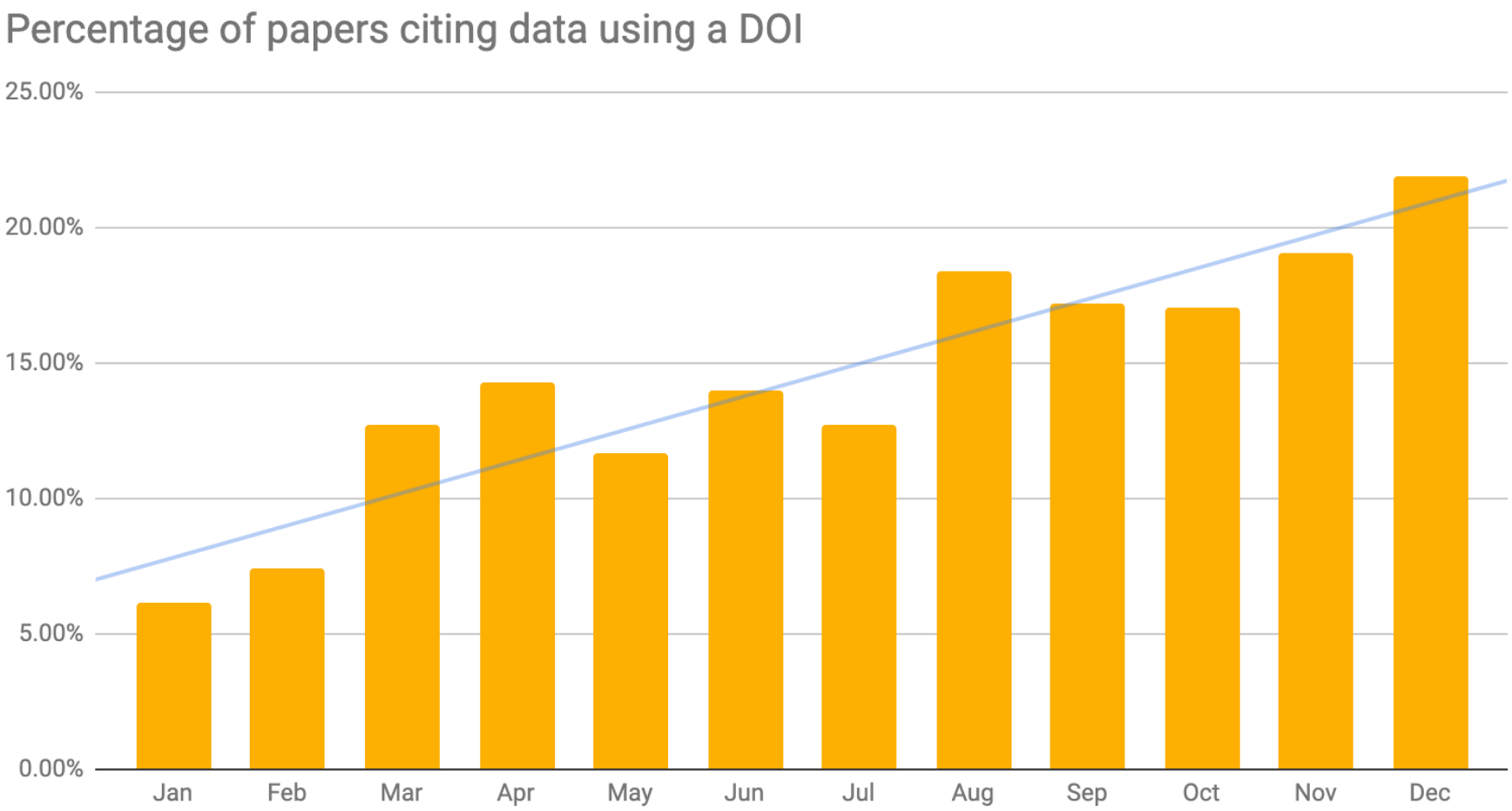



+
-
Tanzania, United Republic of
Voting participant
Member since 2002
Head of delegation Pantaleo KT Munishi
Node manager Hulda Gideon
Has endorsed 15 data publishers
Occurrence datasets: 23
Occurrences: 532,759
Voting participant
Associate country participant

GROWTH IN DATA USE

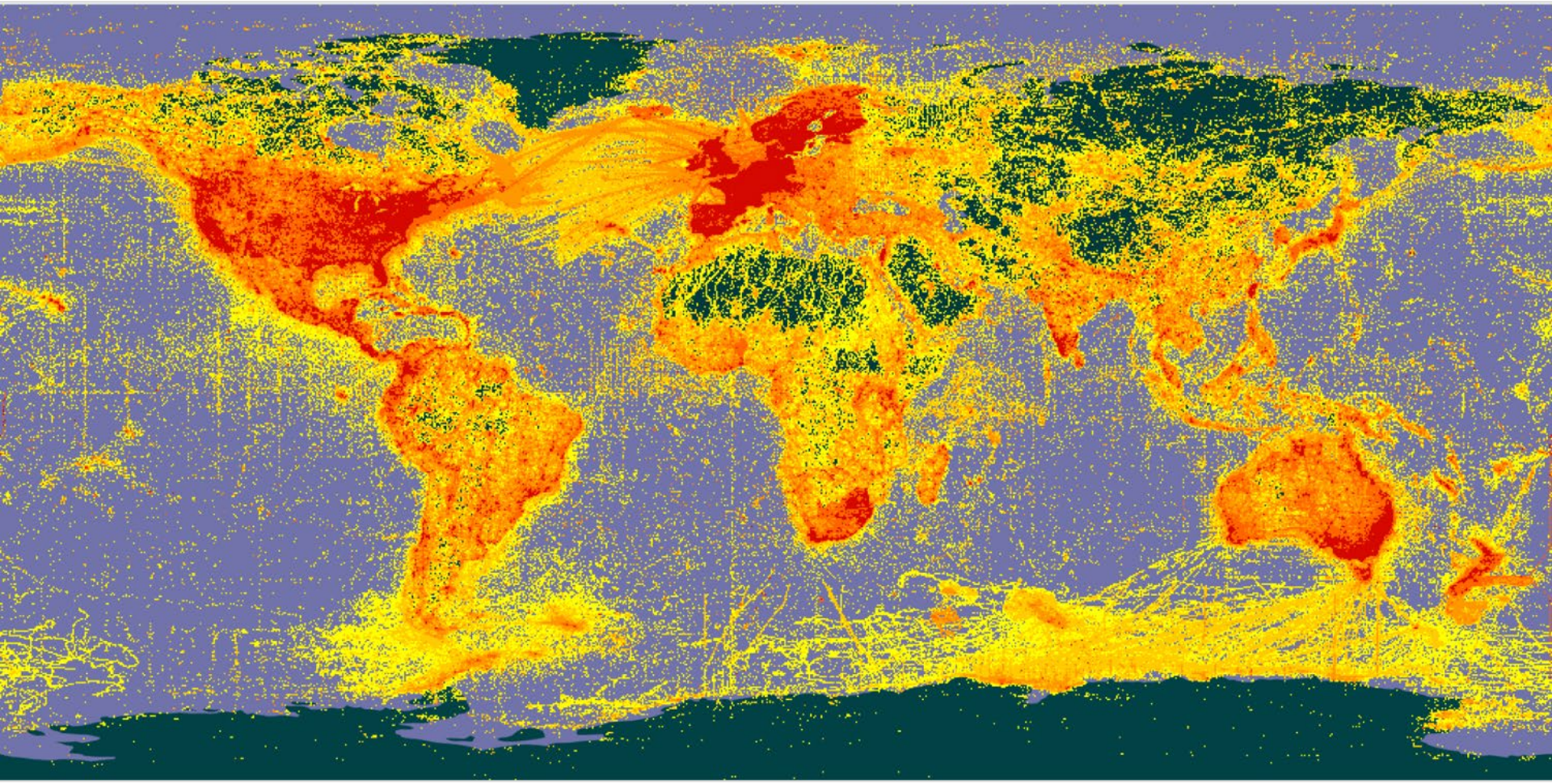


GROWTH IN DOI CITATIONS (2018)



IPBES GLOBAL ASSESSMENT SUMMARY FOR POLICYMAKERS (PDF)

BACK TO TOP



GBIF

GBIF DEVELOPMENT PROGRAMS

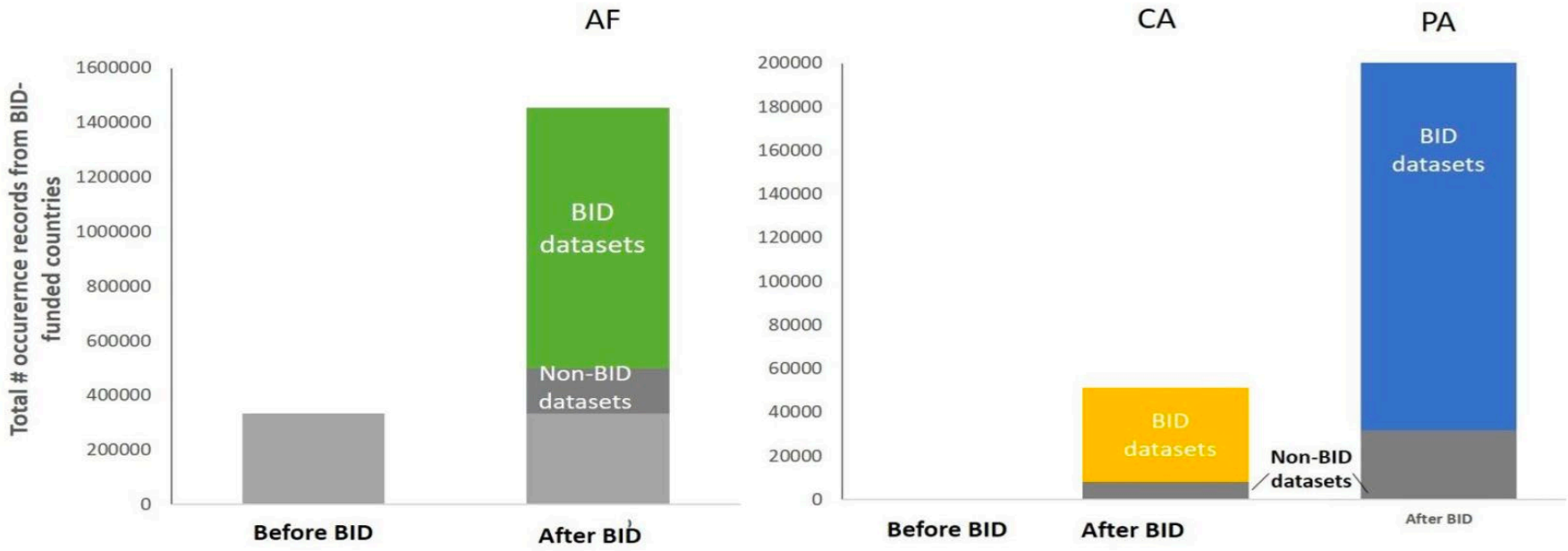
- Capacity Enhancement Support Programme
  - Mentoring activities
  - Support for regional events
  - GBIF advocacy actions
  - Documentation
  - Promotion of data use
- BIFA: Biodiversity Information Fund for Asia
- BID: Biodiversity Information for Development

BIODIVERSITY INFORMATION FOR DEVELOPMENT (BID)

- 3.9 M euros
- Sub-Saharan Africa, Caribbean, and Pacific Islands
- 63 projects

BIODIVERSITY INFORMATION FOR DEVELOPMENT (BID)

- 3.9 M euros
- Sub-Saharan Africa, Caribbean, and Pacific Islands
- 63 projects



GBIF

RESULTS OF THE BID PROGRAM

- Most of the locations were new to GBIF
- 2,400 new scientific names to GBIF

RESULTS OF THE BID PROGRAM

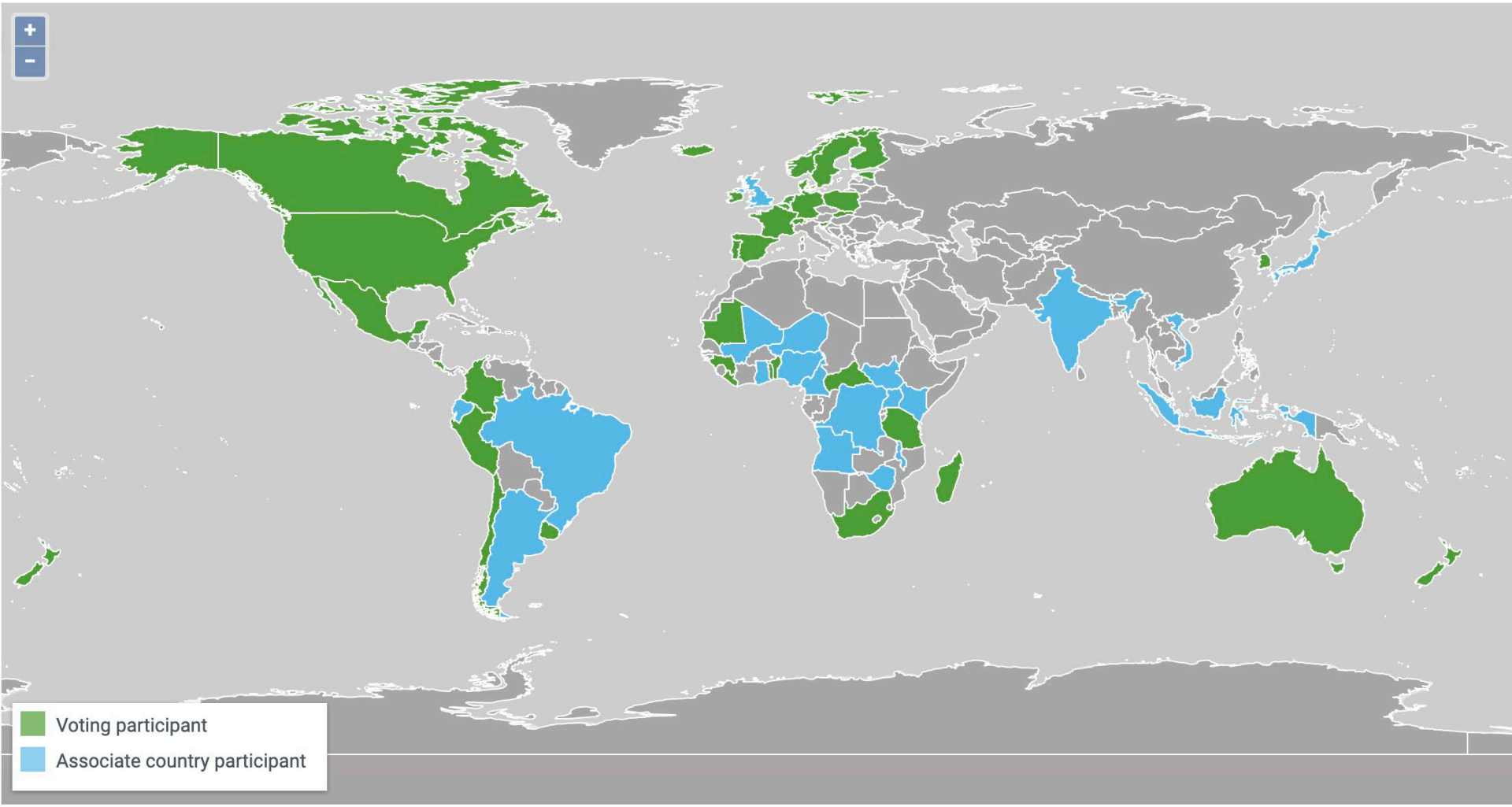
- Most of the locations were new to GBIF
- 2,400 new scientific names to GBIF
- 27% of all taxa included were first for the country
- 1,100 species from IUCN threatened categories
  - 22 of these taxa only known from BID
- Increased data downloads from BID countries

RESULTS OF THE BID PROGRAM

- Most of the locations were new to GBIF
- 2,400 new scientific names to GBIF
- 27% of all taxa included were first for the country
- 1,100 species from IUCN threatened categories
  - 22 of these taxa only known from BID
- Increased data downloads from BID countries
- Increased biodiversity informatics capacity
- 10 new country participants in Africa

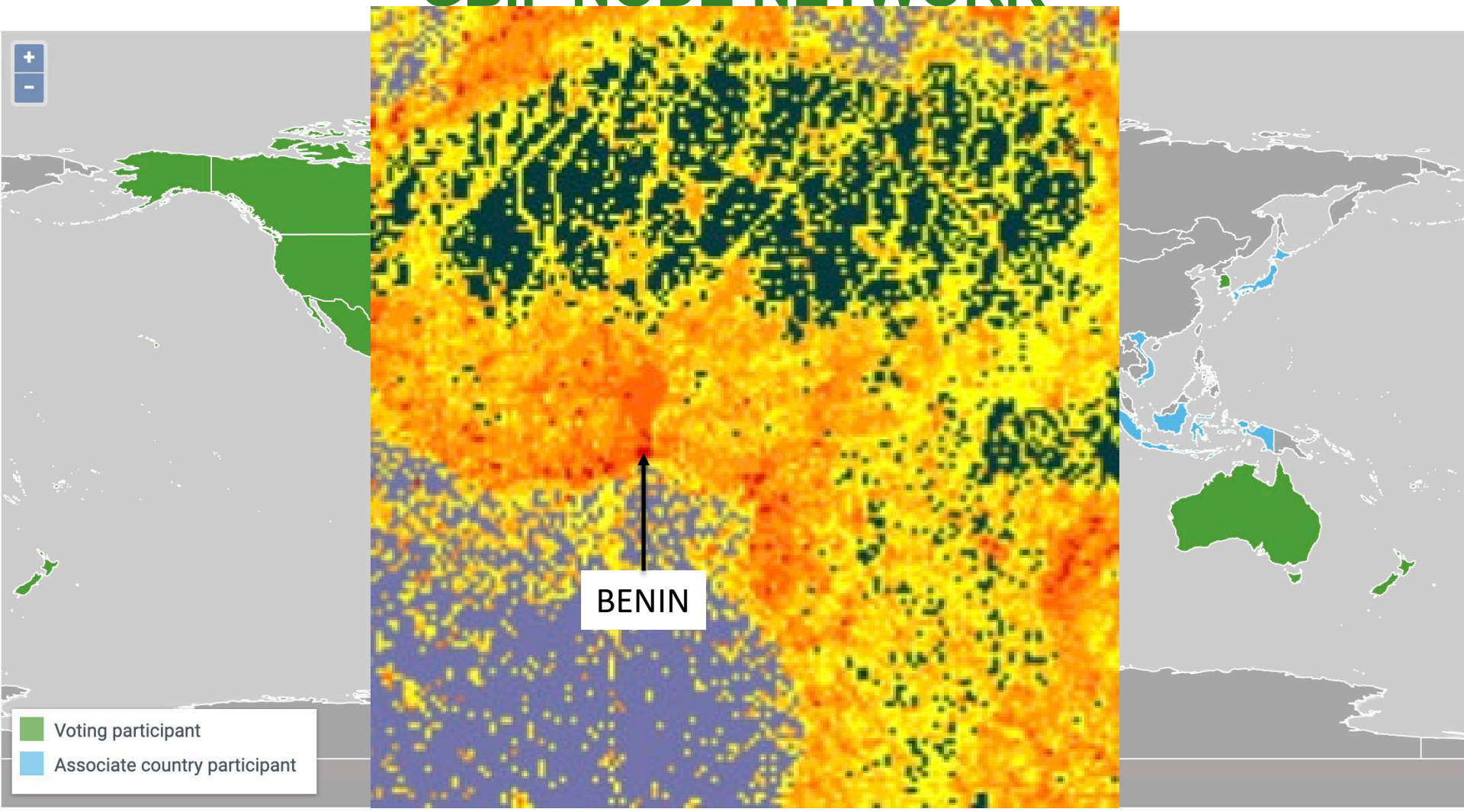
GBIF

GBIF NODE NETWORK



39 VOTING PARTICIPANTS | 20 ASSOCIATE COUNTRY PARTICIPANTS | 38 OTHER ASSOCIATE PARTICIPANTS | 1,408 PUBLISHERS

GBIF NODE NETWORK

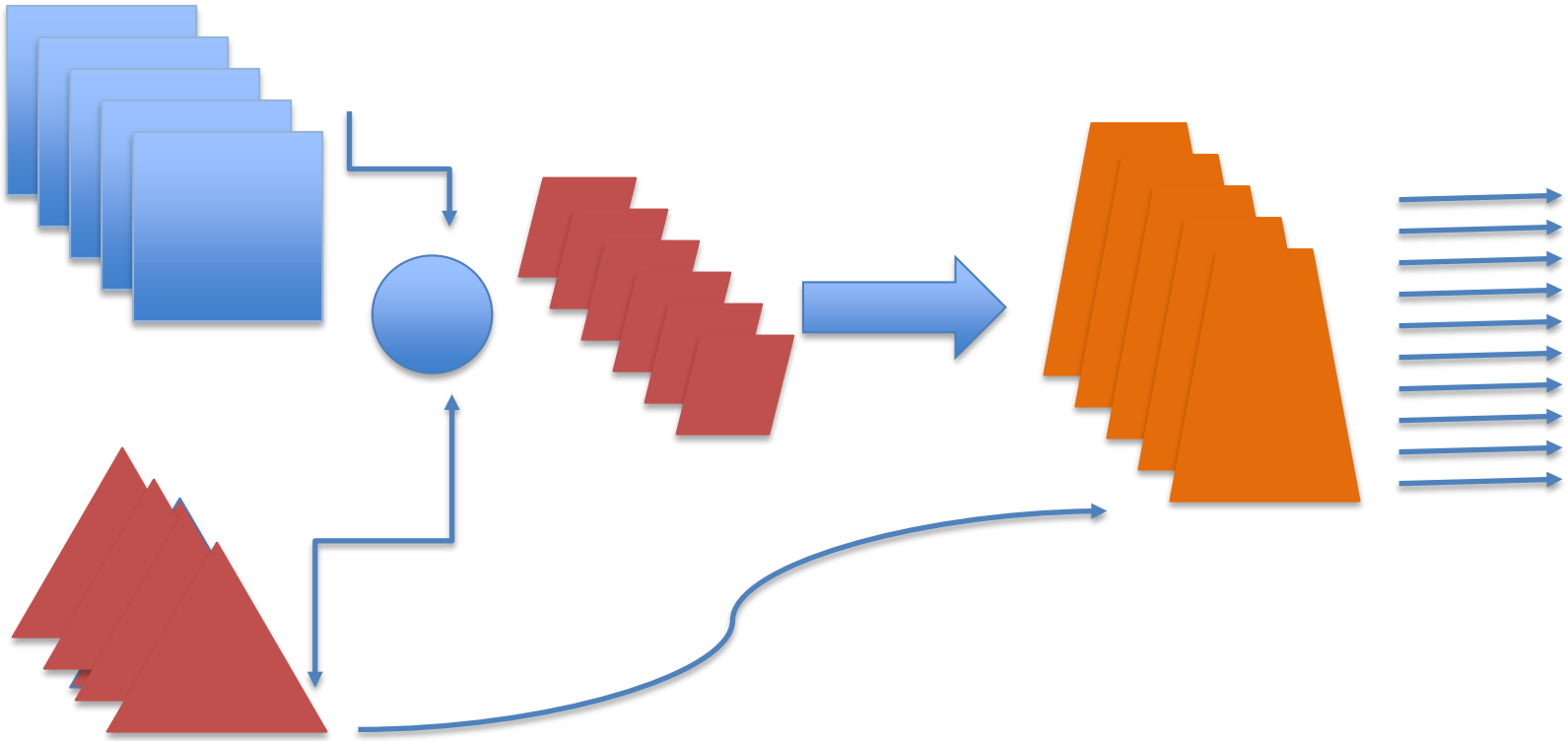


39 VOTING PARTICIPANTS | 20 ASSOCIATE COUNTRY PARTICIPANTS | 38 OTHER ASSOCIATE PARTICIPANTS | 1,408 PUBLISHERS

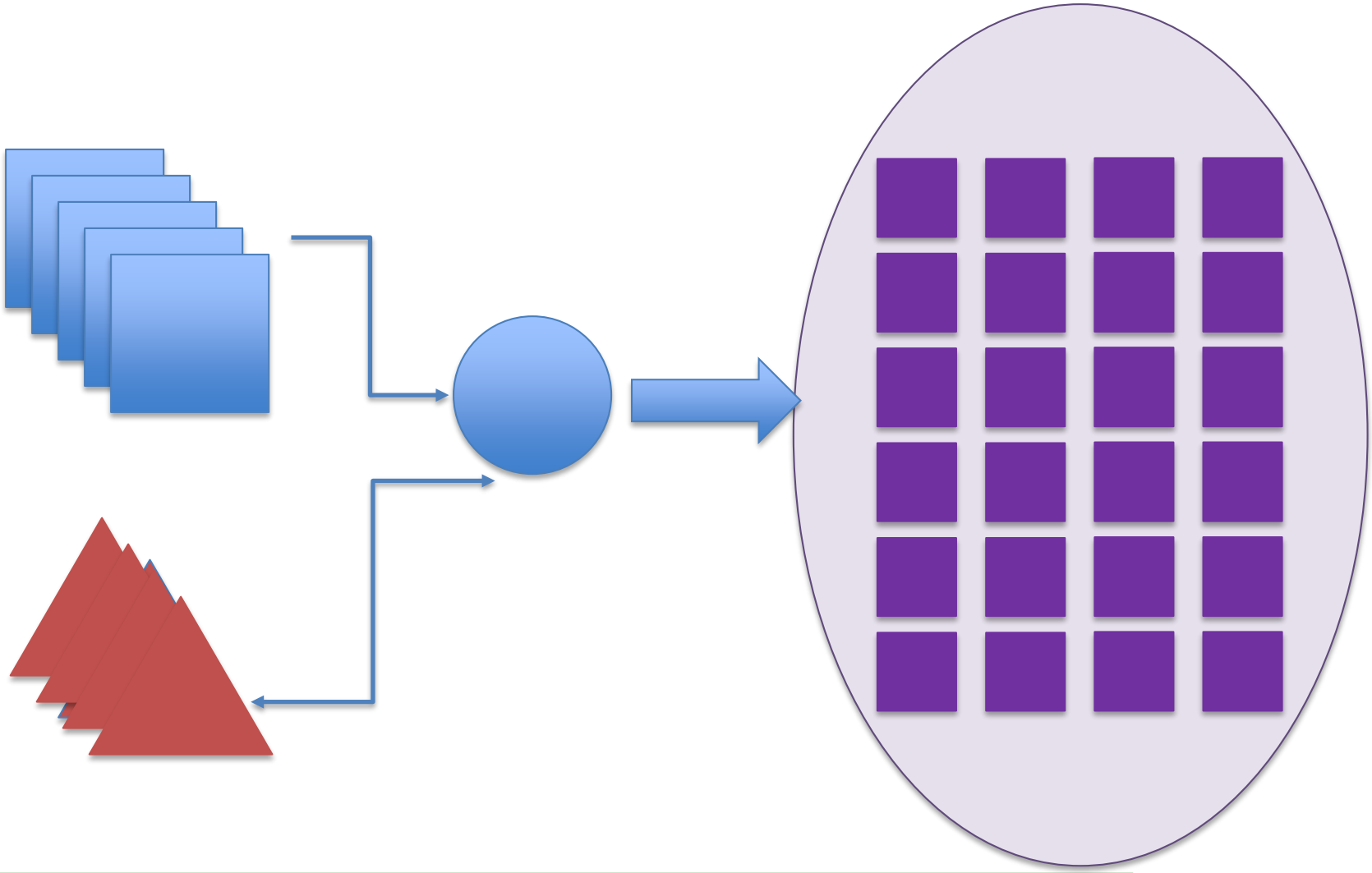
GBIF

2020 WORK PROGRAMME DIRECTIONS

CURRENT SYSTEM



FUTURE SYSTEM



GBIF FITNESS FOR USE

- 1.3 billion is a lot of data

GBIF FITNESS FOR USE

- 1.3 billion is a lot of data
- Data quality is important too
  - Provide ways for data providers to improve data submission
    - Annotation workflows
  - Provide data to users in formats needed
    - Filtered views
- I want your ideas

2020 WORK PROGRAMME DIRECTIONS

- Demonstrate our increased capabilities
  - With updated infrastructure we can implement new functionalities
  - Fitness for use
  - **Expectation management**
- Value proposition
  - Learn your stories
  - Tell our stories
  - Strengthen outreach to data users
    - Conservation planners, evolutionary biologists and ecologist

GBIF

GBIF AND THE ALLIANCE

- Global Biodiversity Information Facility Secretariat and Global Nodes Network

- The *alliance for biodiversity knowledge*

GBIF

GBIC2: *The Second* Global Biodiversity Informatics Conference

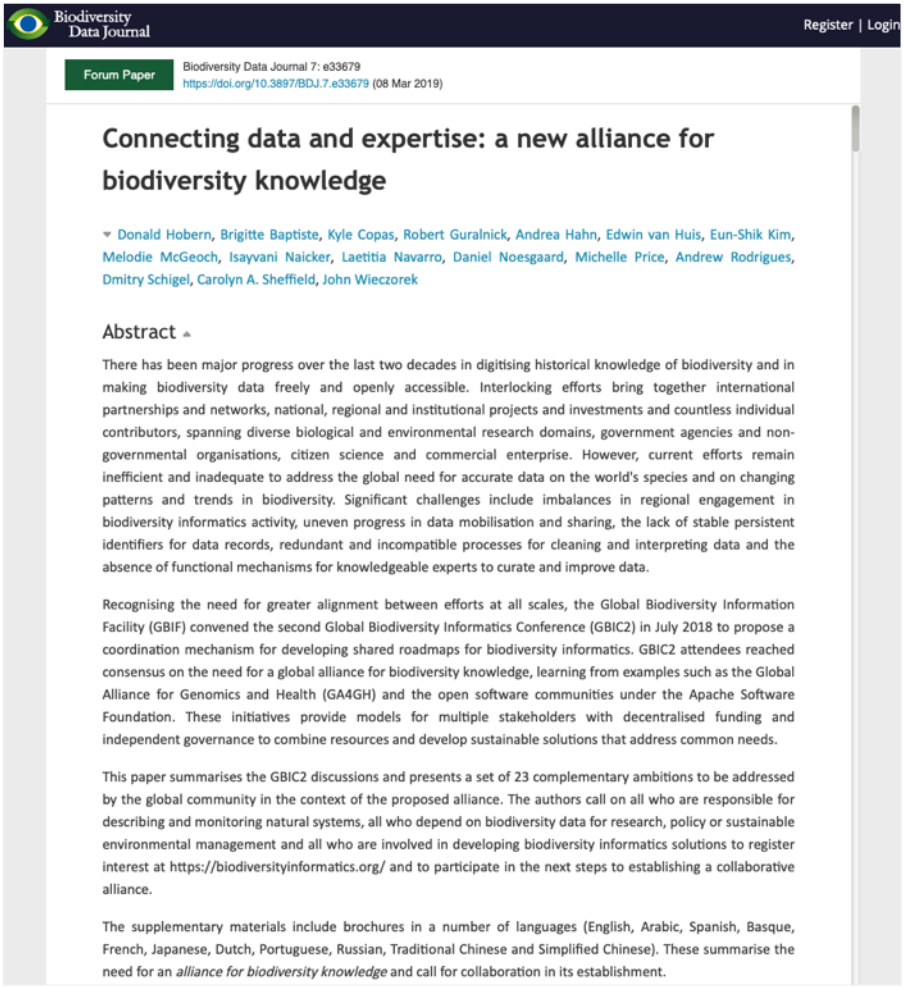
100+ global experts invited to three-day discussion about

- How an international 'coordination mechanism' for biodiversity informatics could operate
- How it could accelerate delivery of a linked and open data biodiversity infrastructure to the benefit of science and society

GBIF

call for an *alliance for biodiversity knowledge*

- GBIC2 meeting report outlines shared vision for connecting data and expertise
- Delegates tasked GBIF with "facilitating…**a global alliance to transform our understanding of biodiversity** by connecting all efforts to observe, measure and model the living planet."



Biodiversity Data Journal

Register | Login

Forum Paper

Biodiversity Data Journal 7: e33679
https://doi.org/10.3897/BDJ.7.e33679 (08 Mar 2019)

Connecting data and expertise: a new alliance for biodiversity knowledge

Donald Hobern, Brigitte Baptiste, Kyle Copas, Robert Guralnick, Andrea Hahn, Edwin van Huis, Eun-Shik Kim, Melodie McGeoch, Isayvani Naicker, Laetitia Navarro, Daniel Noesgaard, Michelle Price, Andrew Rodrigues, Dmitry Schigel, Carolyn A. Sheffield, John Wieczorek

Abstract

There has been major progress over the last two decades in digitising historical knowledge of biodiversity and in making biodiversity data freely and openly accessible. Interlocking efforts bring together international partnerships and networks, national, regional and institutional projects and investments and countless individual contributors, spanning diverse biological and environmental research domains, government agencies and non-governmental organisations, citizen science and commercial enterprise. However, current efforts remain inefficient and inadequate to address the global need for accurate data on the world's species and on changing patterns and trends in biodiversity. Significant challenges include imbalances in regional engagement in biodiversity informatics activity, uneven progress in data mobilisation and sharing, the lack of stable persistent identifiers for data records, redundant and incompatible processes for cleaning and interpreting data and the absence of functional mechanisms for knowledgeable experts to curate and improve data.

Recognising the need for greater alignment between efforts at all scales, the Global Biodiversity Information Facility (GBIF) convened the second Global Biodiversity Informatics Conference (GBIC2) in July 2018 to propose a coordination mechanism for developing shared roadmaps for biodiversity informatics. GBIC2 attendees reached consensus on the need for a global alliance for biodiversity knowledge, learning from examples such as the Global Alliance for Genomics and Health (GA4GH) and the open software communities under the Apache Software Foundation. These initiatives provide models for multiple stakeholders with decentralised funding and independent governance to combine resources and develop sustainable solutions that address common needs.

This paper summarises the GBIC2 discussions and presents a set of 23 complementary ambitions to be addressed by the global community in the context of the proposed alliance. The authors call on all who are responsible for describing and monitoring natural systems, all who depend on biodiversity data for research, policy or sustainable environmental management and all who are involved in developing biodiversity informatics solutions to register interest at https://biodiversityinformatics.org/ and to participate in the next steps to establishing a collaborative alliance.

The supplementary materials include brochures in a number of languages (English, Arabic, Spanish, Basque, French, Japanese, Dutch, Portuguese, Russian, Traditional Chinese and Simplified Chinese). These summarise the need for an *alliance for biodiversity knowledge* and call for collaboration in its establishment.

https://doi.org/10.3897/BDJ.7.e33679

GBIF

alliance **status update**

Status update

- Paper published
- Reconstituted Steering Committee
- Preparing support materials
- GBIF is developing a new web-based discussion platform
- Virtual International workshops

alliance international workshops

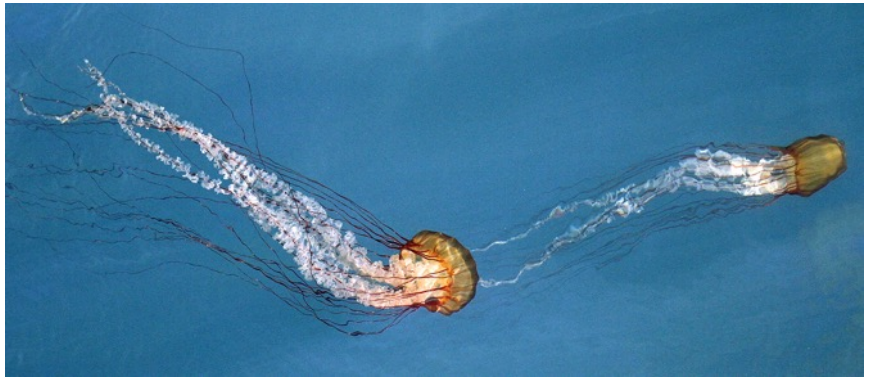
- Any alliance based topic
  - Community input
  - Governance
  - Stakeholder mapping

- GBIF responsible for organizing three collections-focused events over the next three years funded by EU Synthesis Plus project
  - Collections and specimens: building a global collections catalogue
  - Integrated occurrence data | Data standards: documenting a more inclusive information model
  - **Other suggestions welcome**

THANK YOU!

Joe Miller GBIF Secretariat | jmiller@gbif.org

- Sign on and subscribe to the *alliance for biodiversity knowledge*
- Propose our defining questions
- Shape a **biodiversity knowledge commons**



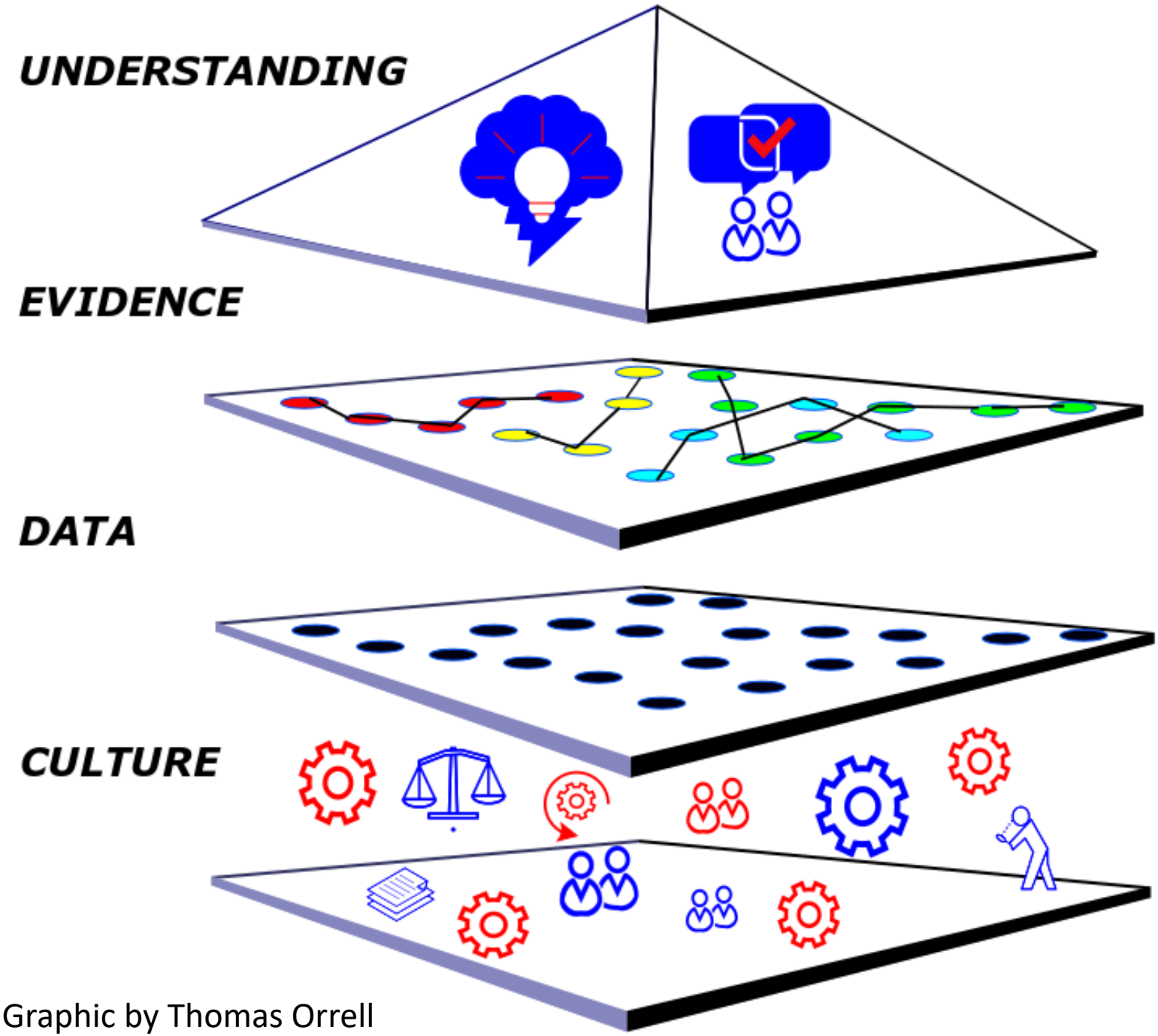
Pacific sea nettle (Chrysaora fuscescens) by Robin Agarwal via iNaturalist. Photo licensed under CC BY-NC 4.0.

biodiversityinformatics.org

GBIF incentive-prize competition on biodiversity data solutions

- **€34,000** in prizes
- Open to individuals or teams
- Deadline: **1 September 2019**

bit.ly/ebbe2019



Graphic by Thomas Orrell

GBIF

TOOLS AND RULES

TOOLS AND RULES

Tools: many are needed to progress our work

TOOLS AND RULES

Tools: many are needed to progress our work

- Databases
- Collections
- Standards

TOOLS AND RULES

Tools: many are needed to progress our work

- Databases
- Collections
- Standards

Rules

TOOLS AND RULES

Tools: many are needed to progress our work

- Databases
- Collections
- Standards

Rules

- Who does what
- How we collaborate

TOOLS AND RULES

Tools: many are needed to progress our work

- Databases
- Collections
- Standards

Rules

- Who does what
- How we collaborate
- Are rules needed?