

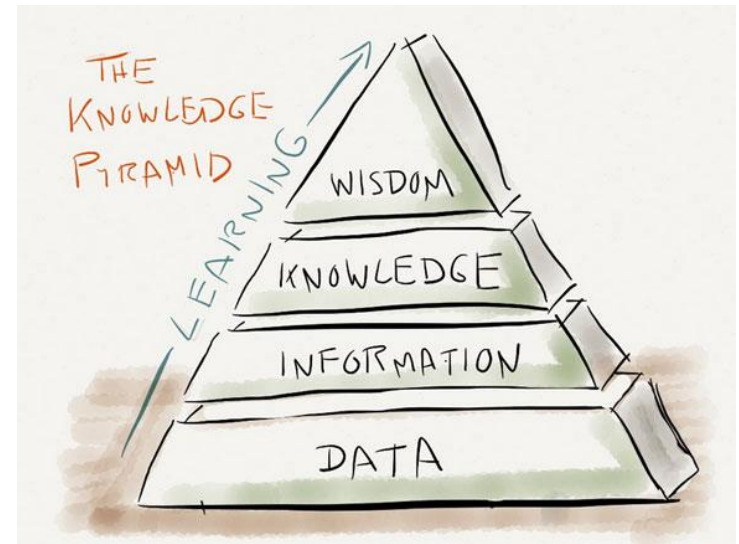
GBIF Initiatives on Data Quality

Dmitry Schigel

*Programme Officer for Content Analysis and Use
GBIF Secretariat*

OUTLINE

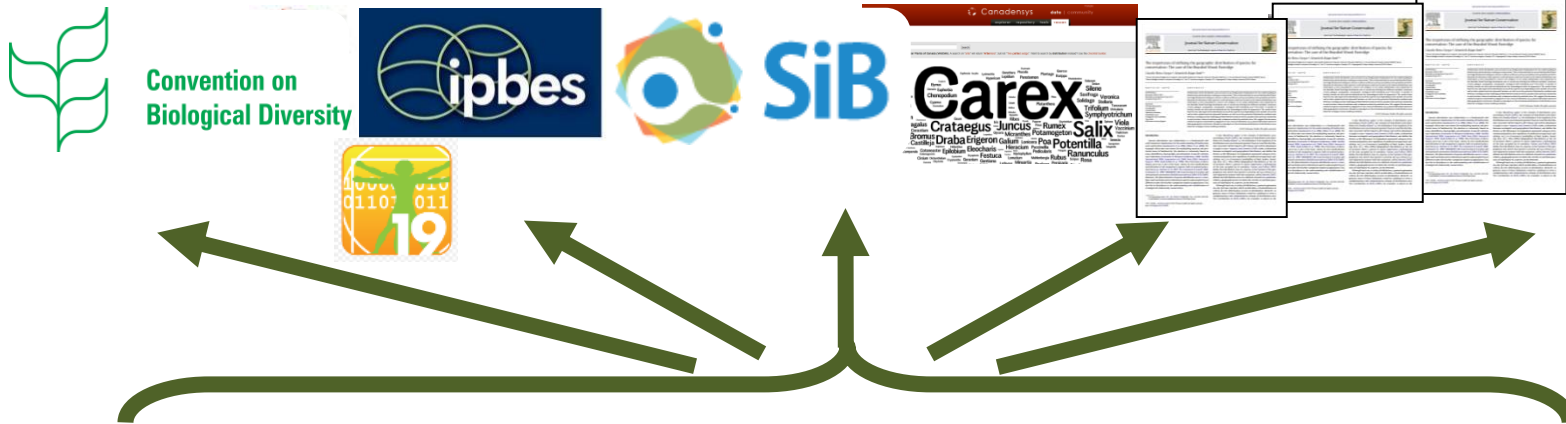
- Global biodiversity data
- GBIF network
- GBIF as a global scientific infrastructure
- Brazil and biodiversity data
- GBIF: fitness for use and data quality
- Future prospects



DATA SOURCES AND DATA USES

Uses of biodiversity data

Taxonomy, conservation, biosecurity, land-use planning, climate change response, crop development, resource management, materials development, forensics ...



Sources of biodiversity data

Collections, field observations, monitoring activities, genomics, citizen science, remote sensing, expert knowledge, historical literature, ...



GLOBAL BIODIVERSITY INFORMATICS OUTLOOK



GBIF NETWORK: PARTNERSHIPS AND AFFILIATIONS



Convention on
Biological Diversity



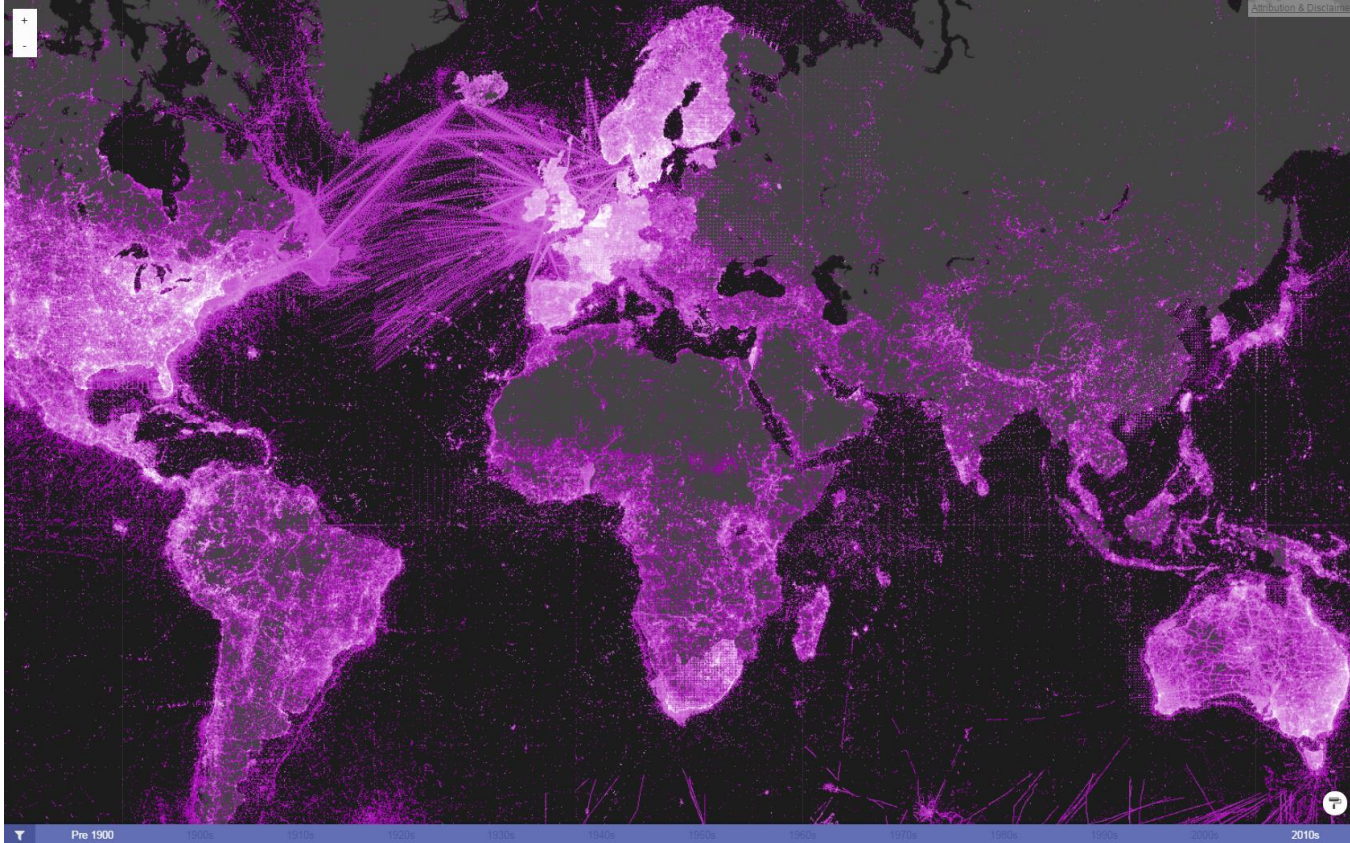
United Nations
Framework Convention on
Climate Change



THE IUCN RED LIST
OF THREATENED SPECIES™



GLOBAL BIODIVERSITY INFORMATION FACILITY



- Data integration - >656 million records for open access download
- Evidence of species occurrence in time and space – any source
- Expanding to accommodate sample event / abundance data
- Web service interfaces



Global Biodiversity Information Facility

Free and Open Access to Biodiversity Data

648,524,213

OCCURRENCES

1,628,227

SPECIES

15,529

DATASETS

799

DATA PUBLISHERS

Data ▾

News ▾

Community ▾

About ▾

Sharing biodiversity data for re-use

[Learn about GBIF](#)

[Publish your data through GBIF](#)

[Technical infrastructure](#)

Providing evidence for research and decisions

[Using data through GBIF](#)

[Enabling biodiversity science](#)

[Supporting global targets](#)

Collaborating as a global community

[Current Participants](#)

[How GBIF is funded](#)

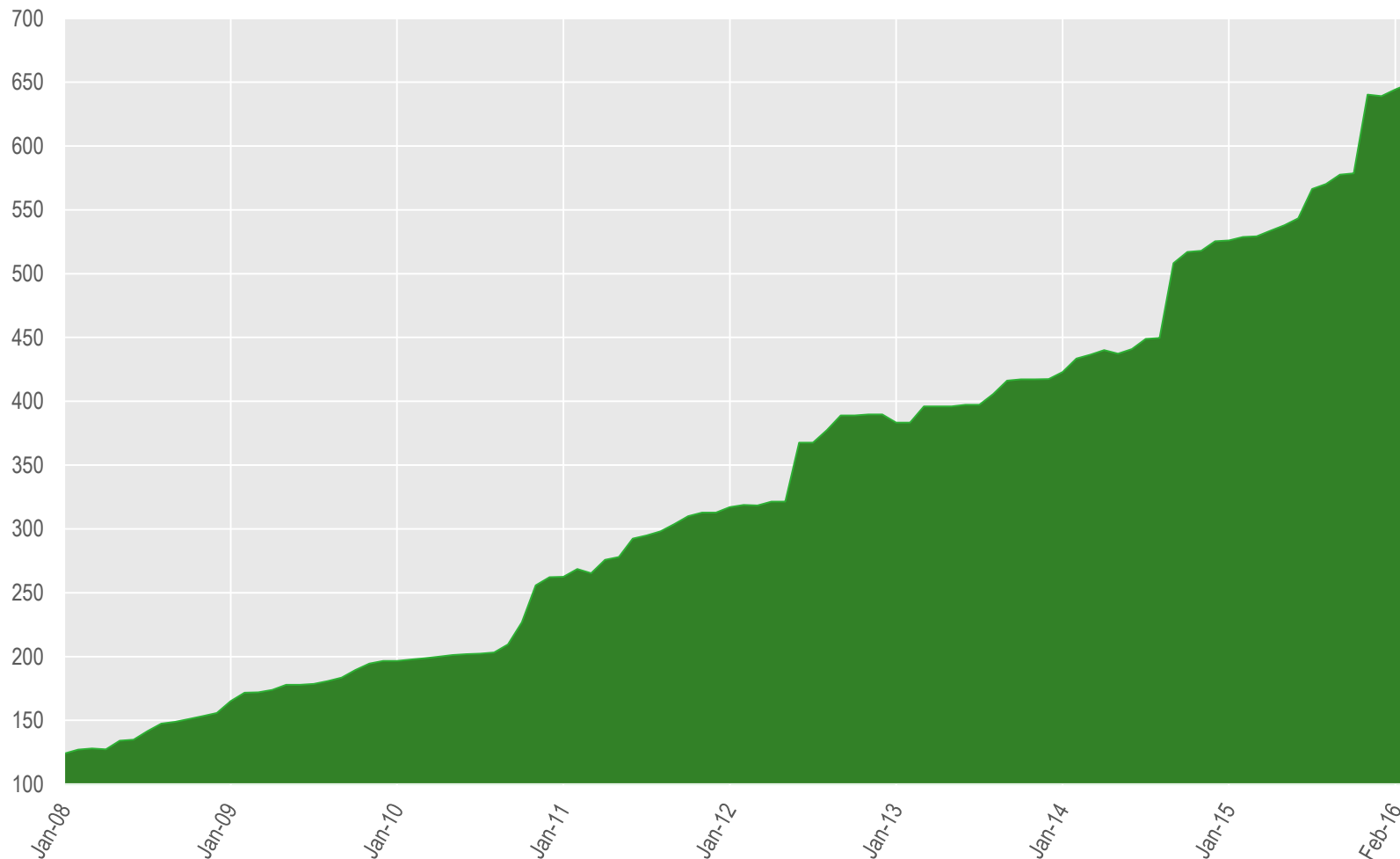
[Enhancing capacity](#)

Search news items and information pages...

Search

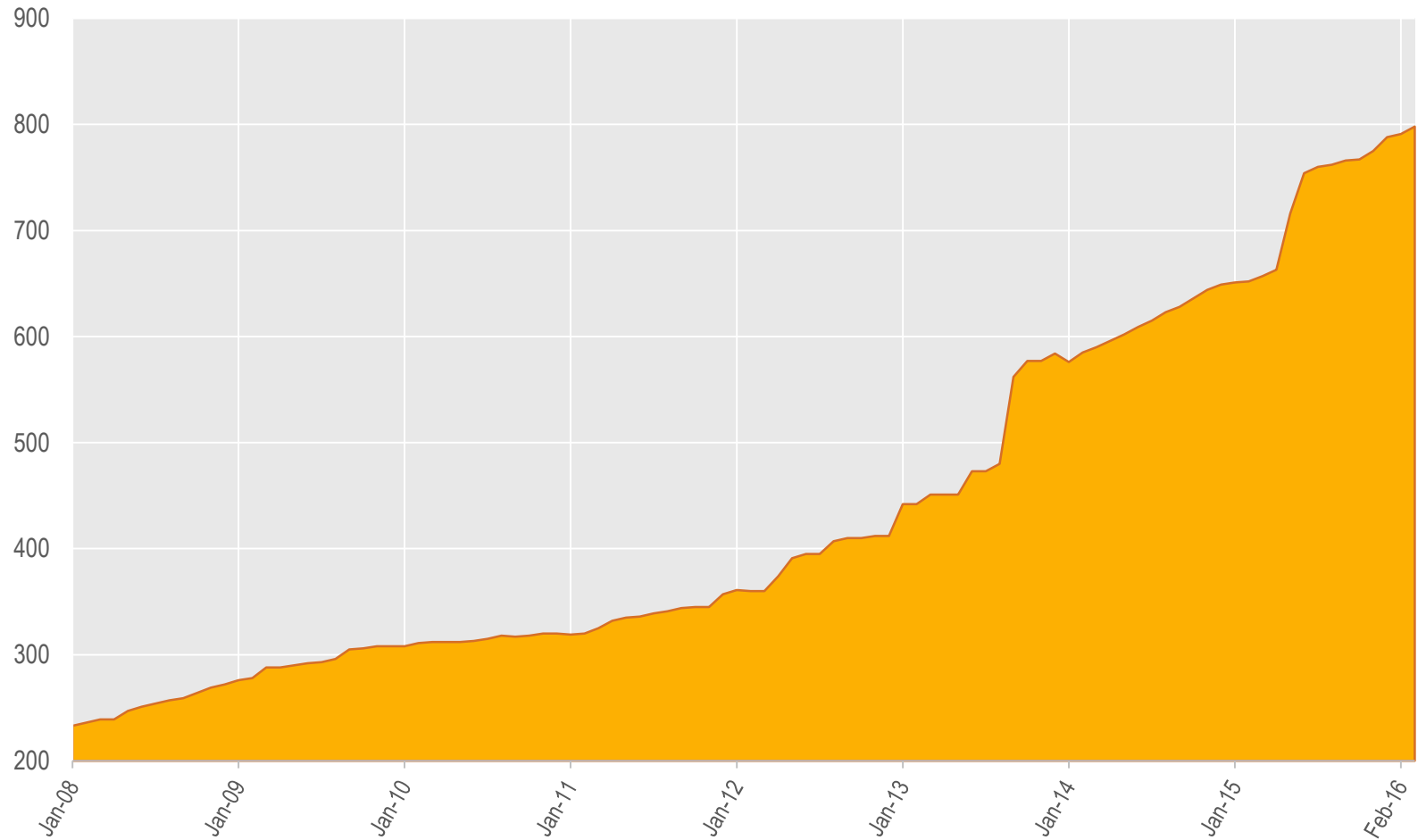
DATA PUBLISHED THROUGH GBIF.ORG

Occurrence records (millions)



DATA PUBLISHERS

Number of institutions registered as GBIF data publishers



GBIF BY THE NUMBERS

648,781,852

species occurrence records

15,516

datasets

798

data-publishing institutions



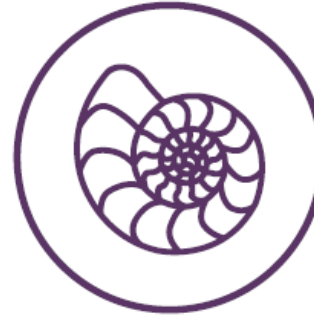
TYPES OF DATA SHARED THROUGH GBIF



Specimen



Material sample



Fossil



Literature occurrence



Observation
Human observation
Living specimen

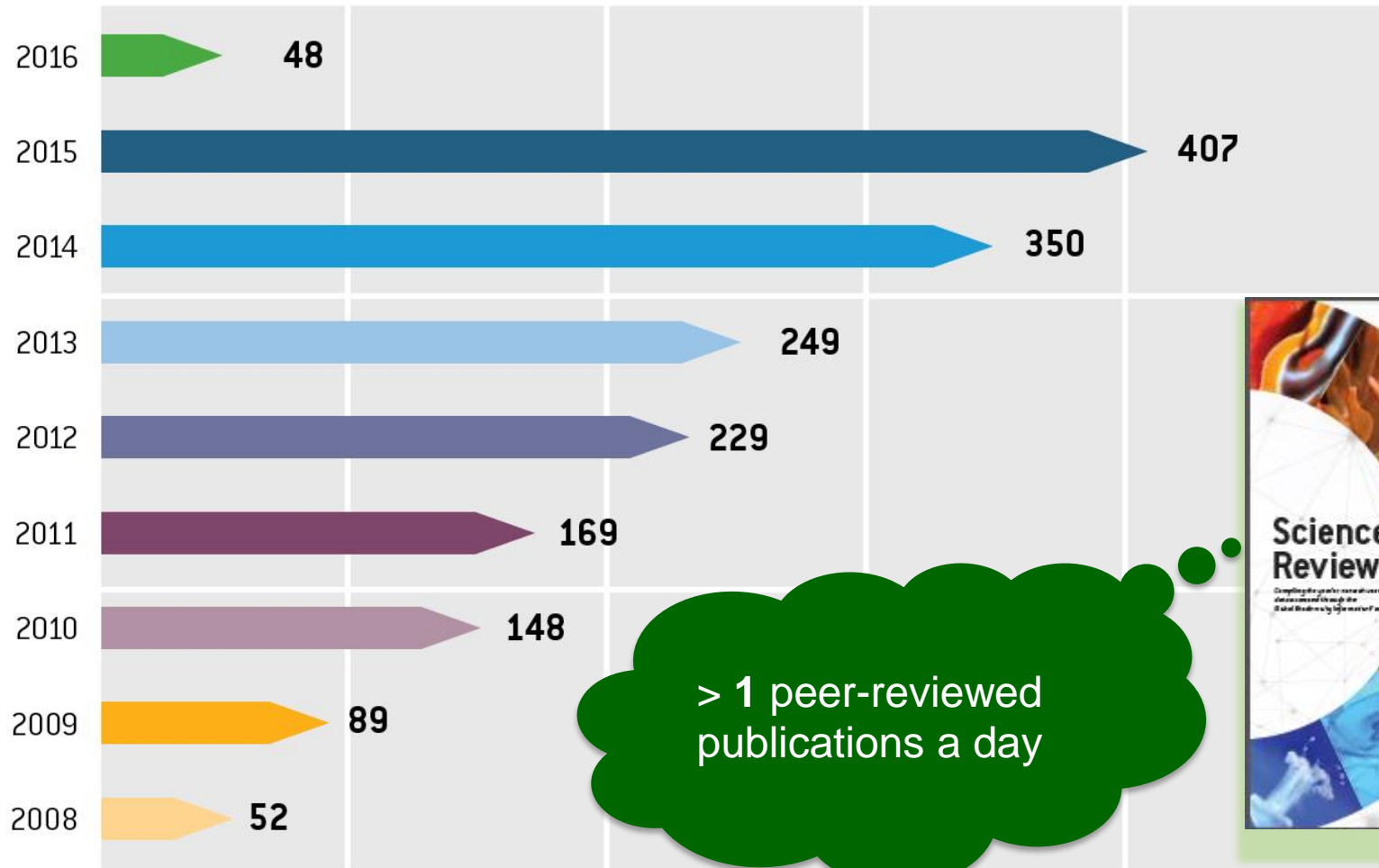


Machine observation



sample-based data

GBIF FOR SCIENCE



> 1 peer-reviewed publications a day



annual number of peer-reviewed publications using GBIF-mediated data

BRAZIL AND GLOBAL BIODIVERSITY DATA

Brazil

A GBIF Associate Country Participant from Latin America
Names of countries, territories and islands are based on the ISO 3166-1 standard.

Country Report

1/1/15 - 31/12/15

Download

Summary

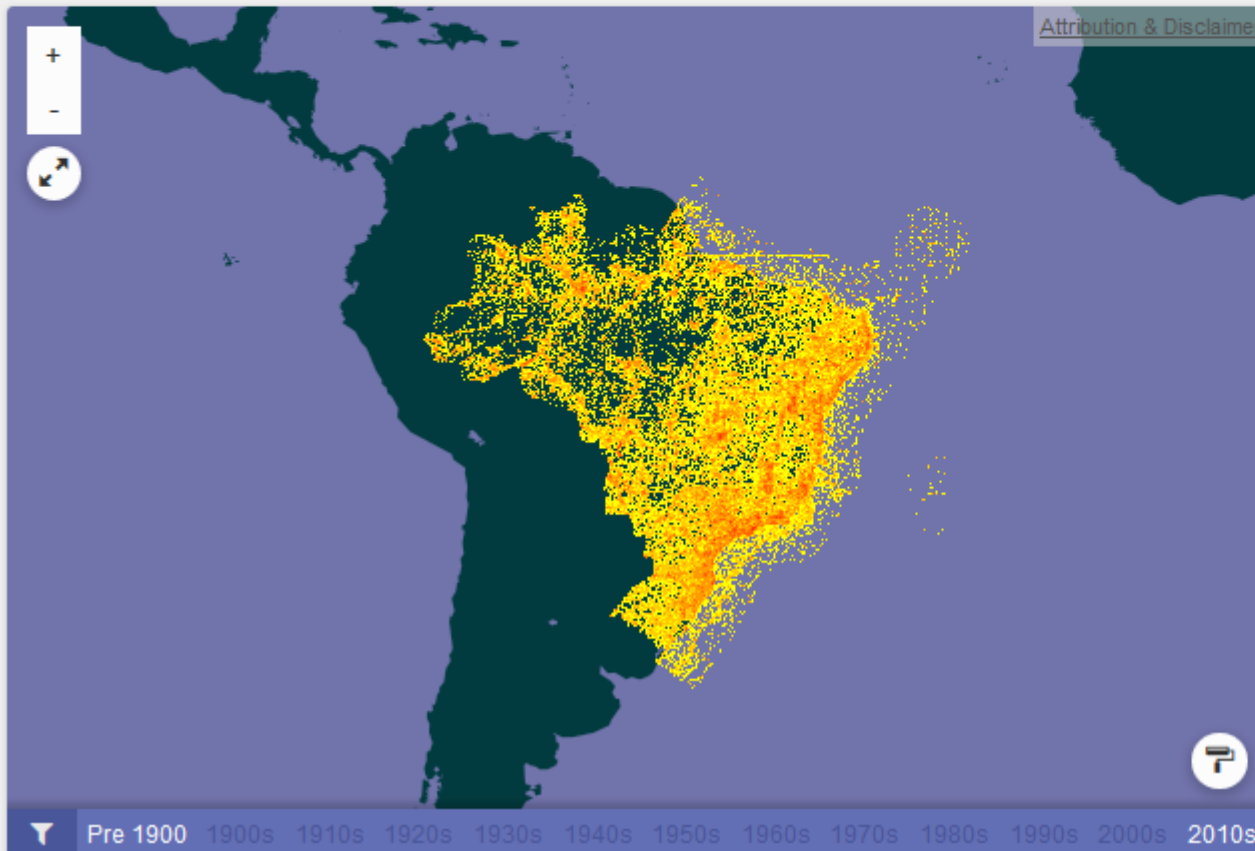
Data About

Data Publishing

Participation

News and Events

Publications



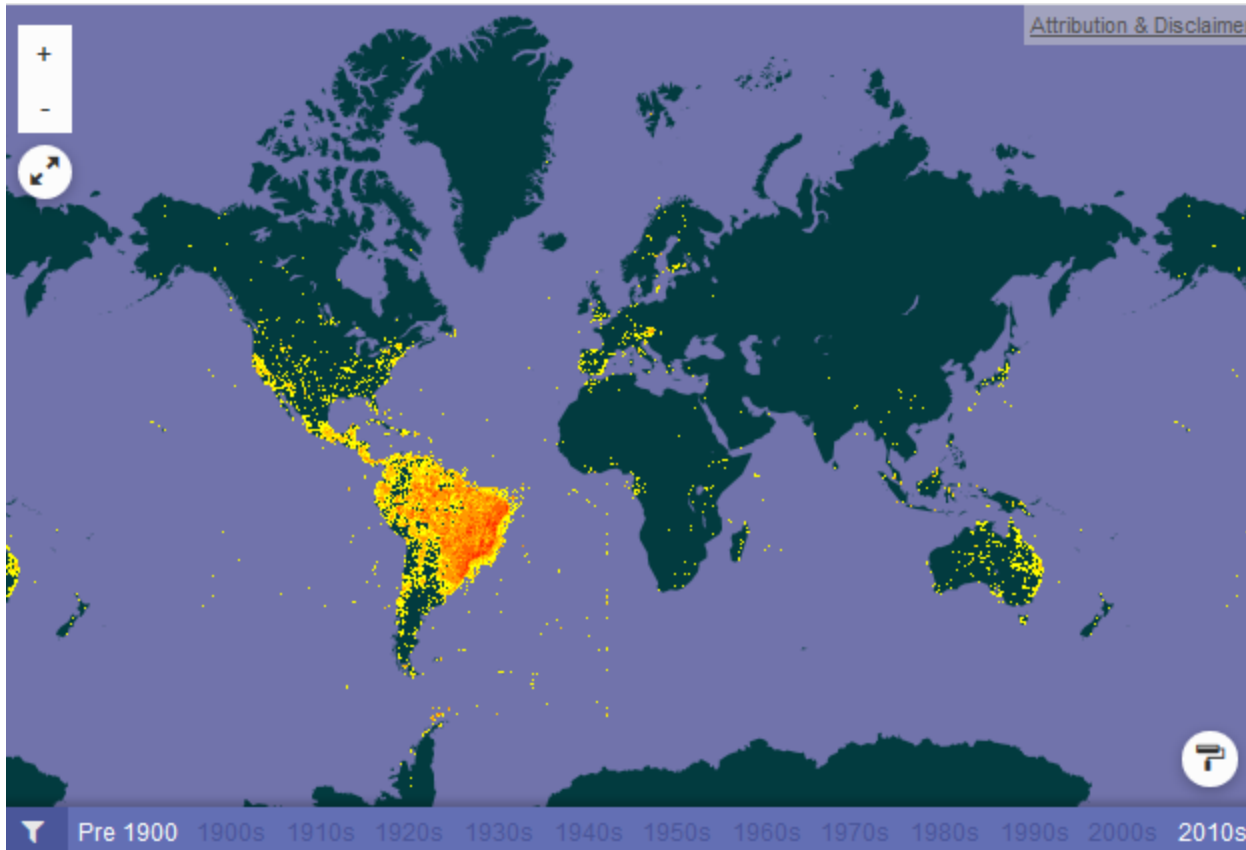
Data about Brazil

- 1,218 occurrence datasets with 7,597,166 records.
- No checklist datasets.
- No metadata-only datasets relevant to Brazil.
- 37 countries contribute data about Brazil.



[View records shown on the map](#)

BRAZIL AND GLOBAL BIODIVERSITY DATA



BRAZIL: A SUPERPUBLISHER COUNTRY



Mammals
52,069
occurrences
+ 45%



Birds
727,079
occurrences
+ 61%



Bony fish
176,139
occurrences
+ 74%



Amphibians
77,851
occurrences
+ 107%



Insects
476,895
occurrences
+ 27%



Reptiles
81,786
occurrences
+ 82%



Molluscs
35,605
occurrences
+ 31%



Arachnids
65,156
occurrences
+ 29%



Flowering plants
4,825,020
occurrences
+ 466%



Gymnosperms
5,734
occurrences
+ 429%



Ferns
224,671
occurrences
+ 916%



Mosses
84,090
occurrences
+ 374%



Sac fungi
124,499
occurrences
+ 1012%

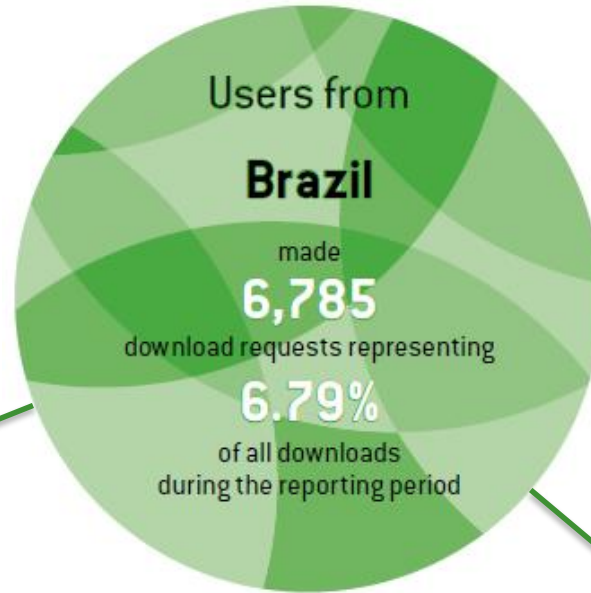


Basidiomycota
72,778
occurrences
+ 495%

BRAZIL: A SUPERUSER COUNTRY

NUMBER OF ARTICLES WITH AUTHORS BY COUNTRY

114	United States
45	United Kingdom
41	Spain
37	Germany
32	Australia
21	Mexico
20	France • Italy
19	Brazil • Netherlands • South Africa
17	Canada
16	Argentina
14	Portugal • Switzerland
13	China
10	Poland
9	Colombia
8	India • Norway
7	Belgium • Sweden
6	Czech Republic • Kenya • New Zealand • Russia



DATA QUALITY, ANALYSIS AND USE

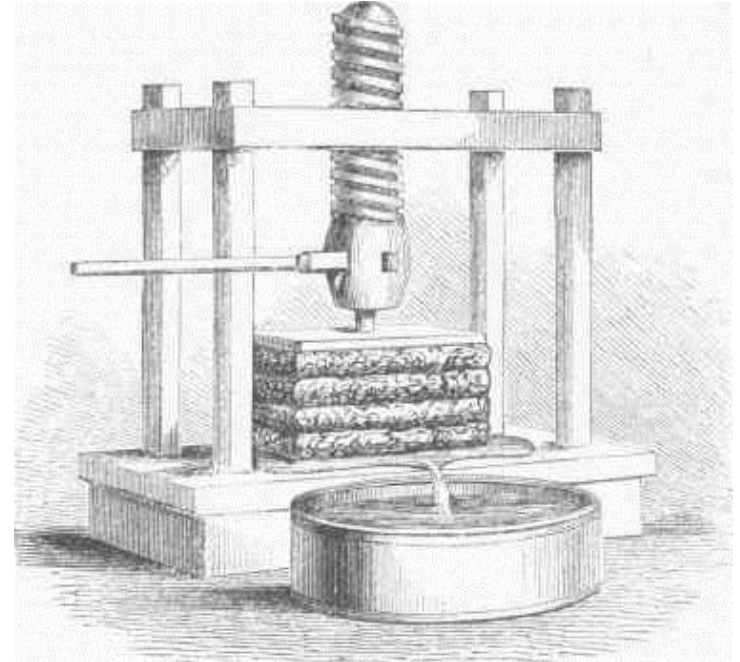


Task groups on **fitness for use**
consulting with pro communities:
recommendations, best practices,
data demands

Promoting data quality culture:
talks, teaching, [publications]

TDWG / GBIF
data quality interest group, chairs:
Antonio Saraiva & Arthur Chapman

Join the work at the **GBIF Community Site!**



Biodiversity
Information
Standards
T D W G

FITNESS FOR USE 2015: AGRO BIODIVERSITY

Agricultural biodiversity and GBIF

Elizabeth Arnaud (Chair), Jean Ganglo Cossi,
Nora Patricia Castañeda-Álvarez, Dag Endresen,
Ebrahim Jahanshiri, Yves Vigouroux



- **Survey** of community needs
- **Final report** on data needs and best practices **ready**
- **50 recommendations**, incl. 7 top priorities
- **Use cases** and template to collect more



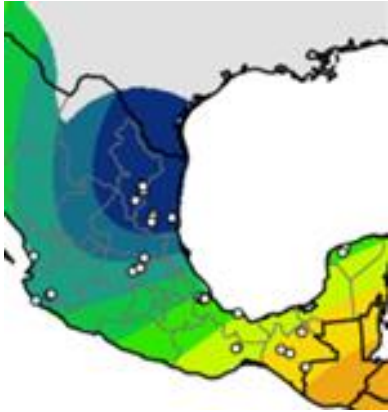
UNIVERSITY OF
BIRMINGHAM



GBIF Norway



FITNESS FOR USE 2015: DISTRIBUTION MODELLING



Distribution modelling and GBIF

Jorge Soberón (Chair), Robert Anderson,
Miguel Araújo, Antoine Guisan, Jorge M. Lobo,
Enrique Martinez, Townsend Peterson

sets of recommendations on improving *access, flagging, completeness, filtering, bias, feedback*, and *archival functions*.

- **Report** ready online
- **survey** of expert experiences and opinions
- **public symposium** videos: <https://vimeo.com/album/3669289>
American Museum of Natural History, 4 Nov 2015

2016: PROFILES, INVASIVES & DNA



2015

- Formalization framework
- São Paulo, Brazil & Co
- Use-specific needs -> **data use profiles**

Why data profiles?

to save time, to reduce data misuse and to lower data use threshold

Fitness for use 2016 topics

- **Invasive alien species**
- DNA evidences for:
 - names = barcoding
 - occurrences = e-DNA
- Marine biodiversity

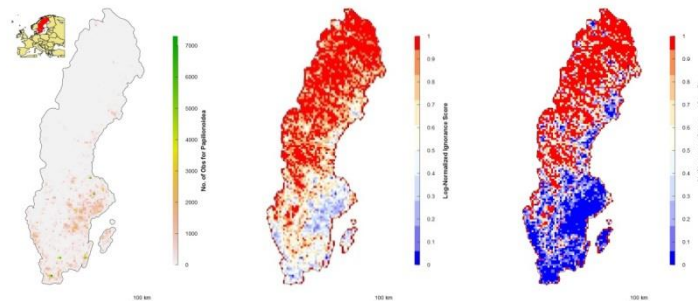
2016

GAP ANALYSIS AND IGNORANCE MAPS

Arturo Ariño & co-authors

GBIF gap analysis guide reviewed
2015: online release for feedback

16 data gap analyses summarized
6 GBIF related cases discussed in detail

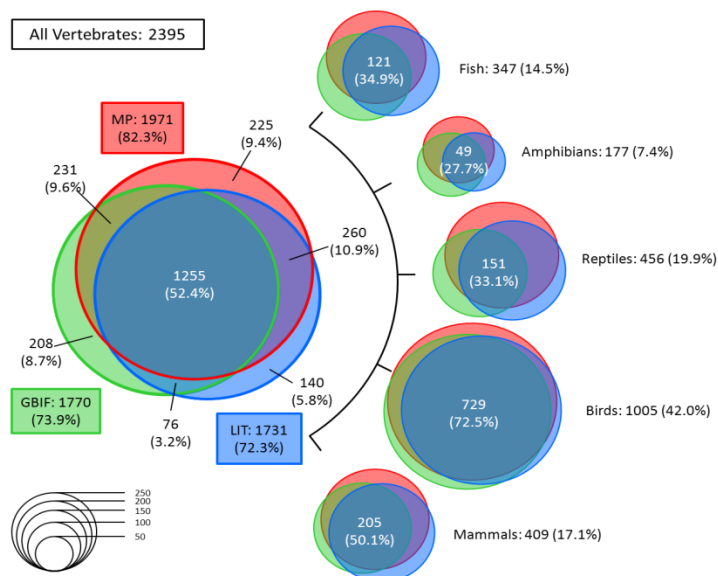


Alejandro Ruede & co-authors

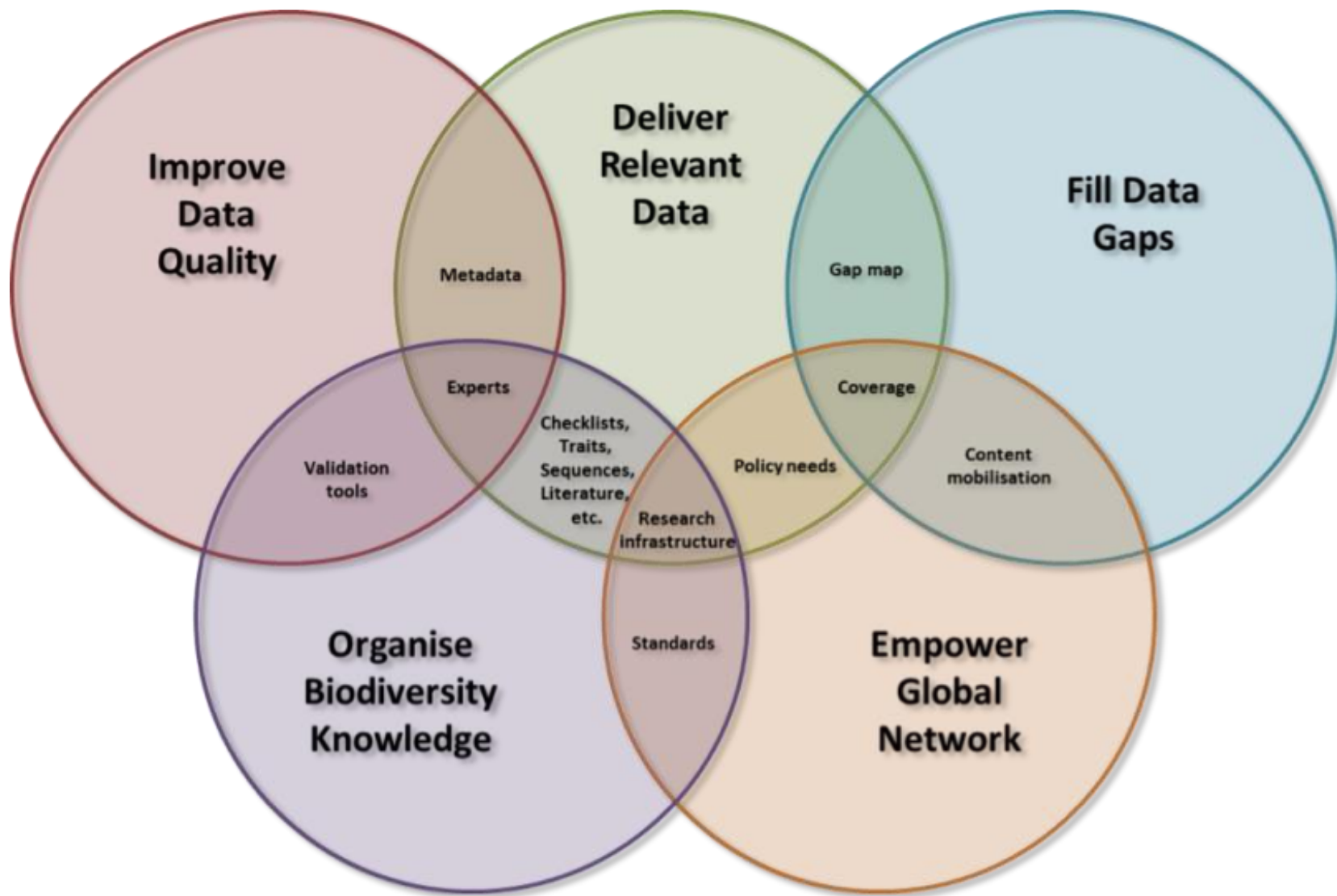
Displaying **bias in sampling effort** of data accessed from biodiversity databases using **ignorance maps**

- to visually explore data quality
- To filter out unreliable results

= **where do citizen scientists go?**



GBIF STRATEGIC PLAN 2017-2021



Muito obrigado

