

Task Group on Data Fitness for Use in Distribution Modelling

The discovery, access and adequate use of primary biodiversity data is critical for distribution modelling, as this rich set of methodologies is among the most important to understand the patterns and trends in global biodiversity, and to inform decision making. Since aggregated data, by its nature, is never comprehensive, a ‘dots-to-surfaces’ extrapolation is the best available approach to generalize the evidence, providing a foundation for biodiversity knowledge. Understanding and predicting spatial biodiversity patterns and trends through modelling of the past, present and future distributions is essential for sustainable use of resources and for addressing many of the world’s key challenges such as species and habitat loss and environment-driven changes. It is estimated that various institutions collectively house several billion specimens of more than a million species and cultivated varieties, genetic samples and other important evidence of patterns and trends in global biodiversity. **Currently**, only a fraction of this vast databank of species information and genetic material is freely and digitally available, and the continuous efforts to use the digitally available information for distribution modelling have resulted in a collection of analytical tools and literature assisting better understanding of the changing environment and biodiversity. From the distribution modelling literature published so far, it is apparent that some of the data handling and data uses examples are based on misinterpretations and insufficient post processing, other are powerful and accurate data use cases leading to scientific insights and adequate support of decision making.

As part of a broader global strategy on fitness for use of biodiversity data, GBIF is convening a Task Group on Data Fitness for Use in Distribution Modelling (DFFU-DM) **aiming** to help improving the fit of data related to distribution modelling needs to the variety of important uses required and requested by this community of research and policy. DFFU-DM will capture the best available experiences, document limitations in existing GBIF services, and suggest improvements in the functionality of GBIF.org for domain-specific needs. Its activities will be informed by ongoing work in GBIF to improve the functionality of its data services by enabling users to filter for relevant data fit for particular purposes, based on pre-defined **data use profiles**.

The Task Group is established with three **objectives**:

1. Based on domain specific data use experience, to make recommendations on improving data availability and data use, data mobilization, data and metadata publishing, and data processing. The Task Group is expected to deliver a vision of the ideal data serving the needs of the distribution modelling community, as well as past, current and expected data modifications, cleaning steps, analyses and visualization needs. Among the outcomes, delimitation of the specific **data use profiles** are expected in the report, defining how GBIF data could be filtered to deliver optimal subsets for particular research applications.
2. To document best practices from ongoing initiatives using distribution modelling related data, and to collect the information on repeatable tools and data management solutions by consulting with ongoing distribution modelling and related initiatives, and the broader biodiversity data community, e.g. through GBIF Community Site and TDWG, relevant academic societies and conferences, in order to bring together the different stakeholders, and catalyse activities around use of biodiversity data for distribution modelling.
3. Based on information from GBIF Secretariat about current developments relating to quality and fitness for use, to make recommendations on GBIF.org improvements, and to

provide guidance in the development of training and outreach materials to help data users to build upon the currently available resources and to share the new developments.

Mandate

1. Develop a schedule and activities for the Task Group
2. Liaise with other experts and define the data use priorities essential for the distribution modelling community and with potential donor organisations that might be interested in funding work on data fitness for distribution modelling
3. Liaise with on-going initiatives and projects to document best practices initiatives using data for distribution modelling
4. Consult with and encourage distribution modelling experts to document and share the repeatable tools and data management solutions
5. Provide guidance in the development of training and outreach materials to help data users to build upon the currently available resources and to share the new developments
6. Consult widely and determine key questions that need to be addressed for the community specific data use needs on data availability and data use improvements, data mobilization, data and metadata publishing, and data processing at institutional, national, regional, and global levels
7. From the user perspective of the data portal develop recommendations for improvements in the GBIF.org functionality by creating targeted data use profiles.

Outputs

The main deliverable will be a set of practical guidelines and recommendations around the issues defined under the Terms of Reference presented in the form of a report to the Executive Secretary of GBIF, on or before December 1, 2015.

Timeline

The Task Group will operate for a period of six months, and is expected to begin in April 2015 after the members have been confirmed by the Executive Committee. A first report will be presented to GBIF Secretariat by 1 August 2015 and circulated among the attendees of the 22nd GBIF Governing Board meeting. The final report will be delivered to GBIF S by 1 December 2015, after which the Task Group shall be dissolved. The final report will be presented to the 23rd GBIF Governing Board in September/October 2016.

<i>Timeline</i>	<i>Activity</i>
March 2015	GBIF EC approves candidate members of the Task Group. Candidate members invited.
April 2015	Establishment of the Task Group. Wireframes for the reports, identification of the priority directions. Initial teleconference. Communication and writing tasks distributed among the Task Group members.
May 2015	First meeting of FFU members to collate initial inputs from Task Group members collated and to refine the work plan. Key stakeholders, GBIF, TDWG and modelling community invited to contribute: consultation open.
July 2015	Comments / inputs from key stakeholders and broader community collated into the first (intermediate) report. Teleconference meeting of the Task Group to collectively draft first round report

August 1, 2015	First report presented to GBIF S and forwarded to GB 22.
November 2015	Consultation closed. Revise report based on inputs. Face-to-face meeting of the Task Group to write up the final report.
December 1, 2015	Deliver final report

Mode of operation

The Task Group will deliver its recommendations to the GBIF Governing Board through the Secretariat. The Task Group will be coordinated by, and report through, the Programme Officer for Content Analysis and Use at the GBIF Secretariat. The Task Group shall mostly operate remotely through email and associated collaborative tools such as wiki, Skype and conference calls; but will hold two face-to-face meetings, focussed on consolidating and writing its report. Travel, accommodation and other essential costs related to these meetings will be met from the budget for the GBIF Work Programme 2014-16. The task force may also use other events of interest to hold additional meetings.

Task Group membership

The Task Group shall be comprised of **6-7 members** reflecting a global representation and the variety of distribution modelling schools. The members are invited based on the priority list of candidates offered by the GBIFS and the Chair of the Task Group for the approval by the GBIF Executive Committee. The Task Group is expected to consult widely within the distribution modelling community and GBIF community, including GBIF stakeholders and the GBIF Secretariat will provide support for this consultation in the form of mailing lists, discussion groups, etc. It is intended to be the nucleus team that will consult widely with other experts, institutions, initiatives & projects. The Task Group is chaired by **Prof Jorge Soberón** of the University of Kansas, USA.

GBIF S contact

Dmitry Schigel, Programme Officer for Content Analysis and Use, e-mail: dschigel@gbif.org/