

Four-month fixed term assignment to carry out a systematic review of GBIF using research on human health and diseases

Summary

The GBIF Secretariat seeks applications to carry out a systematic review of GBIF using scientific literature on human health and human diseases. This will involve: 1) developing, refining and, in consultation with the GBIF Secretariat and GBIF task group on mobilization and use of biodiversity data for research and policy on human diseases, agreeing on the optimal methodology and research questions for the systematic review, 2) on this basis, carrying out a review of 200+ health-related research papers citing use of GBIF-mediated data, published since GBIF introduced DOIs for data citation, and any additional literature as relevant, and 3) as a first author, to prepare a report in the form of a submission-ready manuscript of the review paper, and a corresponding presentation.

Background

The discovery, access and appropriate use of primary biodiversity data are critical for research on human health, zoonotic and vector-borne diseases. This field represents a particular type of use of species occurrence data published by institutions around the world through GBIF, and the demand for broader data in this sector is repeatedly indicated in the scientific literature. As part of a broader global strategy on improving the completeness, relevance and fitness for use of biodiversity data, GBIF has convened a task group on mobilization and use of biodiversity data for research and policy on human diseases aiming to improve the coverage and representation of such data, supporting the variety of uses required and requested by this research community. Among the task group [objectives](#) are identification of sources and contacts to improve data coverage to help research on human health, screening existing GBIF uses to define a clear set of questions that can be addressed with data mediated by GBIF, as well as evaluation of fitness for use of data content, standards, vocabularies, and GBIF services to support research on human diseases. The systematic review will support the work of this task group.

Use of GBIF-mediated data in research is taking place at the rate of ca. 3 peer-reviewed papers a day. GBIF Secretariat, through literature tracking, has built a bibliographic database of GBIF-using literature, and prepares annual GBIF Science Reviews, highlighting individual publications in the main themes where GBIF plays a role in the data foundation. The Science Reviews include basic analytics of yearly literature collections, and there are opportunities for deeper analyses of these studies at various levels including advanced bibliometrics for better understanding geography, demography, open access, journal and other spectra, as well as deeper systematic review of methods applied to the GBIF mediated data, data quality practices, use of analytical tools, and data combinations.

The study

To direct user-oriented development of the GBIF network and its tools, and in particular the work of the GBIF task group on mobilization and use of biodiversity data for research and policy on human diseases, the GBIF Secretariat would like to understand the patterns and trends of the scientific data use in the recent research literature on human health, and also to identify opportunities for growth and development of data use, to analyse existing shortcomings identified in the literature, and to summarize opportunities for diversification of the data use portfolio of GBIF in research, and to highlight promising directions for overall growth of usefulness for the GBIF users, such as biomedical researchers and health professionals.

GBIF requests that the study develops and applies rigid review methodology to carry out **analyses** to understand the data use space that GBIF supports, and how this space relates to the priorities of potential users. In addition, the review should look into research areas and relevant individual high-profile studies which have not cited use of GBIF-mediated data, including cases where GBIF was used indirectly, cases where alternative sources were used, and cases where use of GBIF was rejected or not considered. This study should also include an analytical **assessment** of the current strengths and weaknesses to determine how best to position GBIF strategically to current and potential research users, and how GBIF data products could be improved or extended to increase our overall relevance. A key deliverable will be a systematic review analysis developed with GBIF to engage users and to support strategic development of GBIF and, through data and service improvements, its relevance for the human health sectors.

The range of questions to be addressed by the review:

- Range of **pathogens** and **diseases** as subjects of research
- Analyze the range and the roles among the hosts, vectors, reservoirs and pathogen **taxa** involved, as relevant to data access through GBIF
- Domestic animals, role of **livestock** in transmission.
- Geographic and temporal **spread** of studies
- Country spread of researcher **affiliations**
- Range of topics (e.g. public health), approaches (e.g. niche modelling), disciplines (e.g. epidemiology) explored (**positives**)
- Research ISN'T being done using the GBIF mediated data, and the reasons (**negatives**)
- Analyze data **complementarity**. What types of data are used in the studies? GBIF mixed with? Are GBIF mediated data primarily used **in isolation, or mixed** with other data sources, if yes, with what (possible categories include legal statuses of species at various scales, images, genetic sequence data, species traits, etc.)?
- What alternative data are being used where GBIF data hasn't been used? Capture complementary, potentially publishable data **targets** from these papers. Are there databases of interest that could become discoverable through GBIF and improve access?
- Any commonly reported data **issues** (e.g. temporal/spatial information too limited for analyses, questions about identification reliability, spatial bias etc)
- Range of analytical, statistical, visualization **methods** applied
- **Applications**, where and how data and scientific outcomes are used in practice?
- What are the **limitations** to GBIF data usage by researchers and health professionals?
- How to improve **access** to GBIF data by researchers and health professionals?

Target studies

GBIF uses a DOI based system of literature tracking, which enables, on the basis of data citations and mentions of GBIF, to identify studies including those where GBIF mediated data have been used. The resulting collection is available [online](https://www.gbif.org/resource/search?contentType=literature&topics=HUMAN_HEALTH&relevance=GBIF_USED) at GBIF.org. A subset of this literature collection has been tagged for human health relevance and this subset, linked below, is the primary target for the contracted review:

https://www.gbif.org/resource/search?contentType=literature&topics=HUMAN_HEALTH&relevance=GBIF_USED

A relevant description of the compilation process of the “positives” will be available from GBIF for use as part of Materials and methods. The contractor is expected to come up with a systematic approach to identify and explore other relevant studies where GBIF mediate could have been used, the “negatives”. One potential source of insights on the “negatives” are GBIF-mentioning (but not GBIF using) [studies](#) on human health; other approaches to explore potential growth of GBIF relevance are welcome. The snapshot of a complete (as used in the review) bibliographic database of the positives and negatives should be archived and serve as supplementary materials to the review paper.

The ideal contractor would have:

- Understanding of the role and mission of GBIF, and its research users
- Personal experience of using GBIF mediated data
- Proven record of bibliographic, text mining, review work
- Familiarity with biodiversity informatics landscape
- Experience of writing and publishing research and review papers
- Ability to understand, interpret, and analyze research carried out by others
- Understand the key analytical methods and approaches in the field

Biomedical background and relevant doctoral degree are advantageous, but not required - this is a postdoc level, strictly time-limited task analytical and writing task.

Application and deadline

Deadline for proposal submission	30 April 2021
Expected start of contract	1 - 31 May 2021
Duration of the contract	4 months
Delivery of first draft	15 July 2021
Delivery of final draft	by the end of contract, August 2021

Location

The selected candidate shall carry out the work at their home base.

Fee

The contract is paid by a fee of €10,000 from the GBIF Secretariat in two or more installments to cover all costs of the contractor.

Your application

Please include the following in your application:

1. A summary of the approach you will take to the work, including a potential outline for the report (no more than two pages inclusive of possible schemas)
2. A full, max 2 pages CV or similar, plus a list of publications
3. Examples and other means of demonstration of your fit to the task where you led or actively participated in the past five years that are relevant to this proposal.

Please send your proposal to info@gbif.org.

If you have any questions relating to the specifications for this contract, please contact Dmitry Schigel (dschigel@gbif.org).