

FOUNDATIONAL BIODIVERSITY INFORMATION PROGRAMME (FBIP)

DATA RELEASE REQUIREMENTS

April 2017

The SANBI Biodiversity Information Policy Framework provides legal principles and guidelines on managing biodiversity information. Through this Framework SANBI strives to ensure easy access to information whilst simultaneously providing protection to sensitive data and maintaining intellectual property rights.

In terms of the release of data from the FBIP grant recipients/data publishers a number of recommendations are provided below:

Public Release of Data

- Biodiversity data accessible via the SANBI network are openly and universally available to all users. All grant holders are to agree in writing that data produced through the use of FBIP funding can be made publically accessible without restriction, but attribution is required so that the data provider is acknowledged by all users of the data.
- Individuals and institutions must agree to make the data produced from FBIP grants available six months after the completion date of the project as stated in the proposal.
- For projects generating DNA barcode data, these are submitted to the Barcode of Life Database (BOLD) held by the international Barcode of Life (iBOL), and are made accessible according to the iBOL data release and resource sharing policy (<http://ibol.org/resources/data-release-policy>). iBOL considers all barcode data in BOLD to be a community resource to be shared publically according to the terms and conditions outlined in its policy, but encourages users of the data to acknowledge the contributor and source.
- Data submitted in BOLD will be made public and transferred to GenBank for public release prior to user initiated publication. Data release will follow a two phase process: Phase 1: quarterly release of all generated sequence data and high level taxonomic information. Phase 2: release of additional data elements that require manual curatorial efforts and detailed taxonomic enquiry.
- In current research practice, a researcher or institution may be granted temporary exclusive use of the data produced. In the event of the open dissemination of data posing a risk to a student thesis, the data provision to the FBIP can be delayed for one to two years on request to the FBIP.
- In the instance of data produced by a student that is withheld from public release the data must still be provided to the FBIP when the project ends. This data is required for monitoring and reporting purposes. In the event that the grant holder / student has improved the data since its submission to the FBIP, a new version should be provided for dissemination.
- Where data are considered to be sensitive (the GPS co-ordinates of a collection locality for a species threatened by over-collecting) or confidential (eg. name of human subjects from which disease samples were taken), then GPS co-ordinates can be disseminated at a coarse scale or the names of human subjects must be removed from data sets. These or similar restrictions must be stated by the grant-holder when data are submitted to the FBIP.
- For FBIP funded projects, data being served via the SABIF/SANBI-GBIF platform, will be associated with one of two licences i.e. **CC0** or **CC-BY** (see additional notes below). All publishers are to choose a licence equivalent to: CC0 or CC-BY:
 - **CC0** - Fully-open public access
 - **CC-BY** - Attribution required

Additional Notes:

GBIF is updating its data sharing / licensing terms for publishers.

At its meeting in India last year, the GBIF Governing Board agreed to a policy (<http://www.gbif.org/newsroom/news/data-licensing-and-endorsement>) to ensure that all species occurrence datasets within GBIF.org would in future be associated with one of three Creative Commons licence designations: CC0, CC BY or CC BY-NC.

This policy change followed a stakeholder consultation earlier in 2014 (<http://www.gbif.org/newsroom/consultations#licensing>), in which respondents indicated the need to clarify and standardize the terms of using data accessed through GBIF.org, both for publishers and for users.

CCO (Copyright Waiver)

Several countries and significant publishers of data within GBIF have already adopted unrestricted **public domain tools** (rather than licences), in particular **CC0**, as the basis for data publishing within their networks, arguing that other options are not legally valid and/or that this option best supports open science. **GBIF should support data publishers adopting public domain tools wherever these are appropriate.**

CC-BY:

A number of institutions publishing data within GBIF currently have policies that do not allow data to be shared without asserting intellectual property rights or defining restrictions on use. Several significant networks publishing data within GBIF have existing data sharing agreements with their publishers and cannot vary the terms associated with these agreements. **Under current circumstances, GBIF cannot move completely to adoption of public domain tools (e.g. CC0) without risk of serious disruption and loss of data.**

CC-BY-NC

Several responses also indicated that, at least at present, some of their data publishers expect to be able to restrict data use to non-commercial activities. However, there is no clear and accepted definition of which activities are to be considered non-commercial. As with other restrictions on data use, it is unlikely that such restrictions could be enforced in a consistent way throughout the network. Creative Commons indicates that non-commercial restrictions do not contribute to "Free Culture". **Under current circumstances, GBIF may face some disruption and loss of data if it excludes the option for data publishers to request exclusion of commercial uses.**