##### **PRT-CGN-PG-503 PROTOCOL FOR PUBLISHING DATA VIA THE INTEGRATED PUBLISHING TOOLKIT (IPT)**

**Introduction**

The Global Biodiversity Information Facility (GBIF) is an international non-profit organisation to provide free and universal access to data regarding the world's biodiversity.

A wide range of countries and organisations participate in GBIF and have made their data on biodiversity available via the portal of GBIF. The tool that is recommended by GBIF to publish data via the GBIF data portal is the Integrated Publishing Toolkit (IPT). Information on passport data of the collection of plant genetic resources of CGN can be searched for via the GBIF data portal, see: <https://www.gbif.org/>.

**IPT**

For this functionality CGN is using an IPT that is hosted and maintained by Naturalis Biodiversity Center. IPT can generate data archives which can be harvested, indexed and published by GBIF on their data portal. In the future, also other portals (e.g. EURISCO) might use the IPT to harvest, index and publish data from CGN.

The Dutch node of GBIF (NLBIF) can be approached for all questions regarding the IPT. There are also extensive manuals available from the GBIF website.

The IPT can be approached via URL: https://ipt.nlbif.nl. CGN can log in at its own data sources to refresh the data.

Making data (from GENIS) available via IPT has two main aspects:

* Availability of (Oracle generated) text files that contains the data for publishing. It is possible to publish several datasets.
* Configuration and creation of archives by mapping of schemas (e.g. DawinCore, DarwinCoreGermplasm extension) and entering meta data on these data files within the IPT.

**Datasets**

It is possible to publish several data sets through the IPT. Different data sets might have their own tables, views and/or scripts. They are described below

1. *Centre for Genetic Resources, the Netherlands, PGR passport data, published via GBIF data portal*

* File location: N:\Documentatie\CGN\_Databases\IPT
* Script: CreateIPT.sql
* Oracle table: GNS\_IPT
* Text file: ipt\_cgn\_passport\_data.txt
* Schemas: DawinCore and DarwinCoreGermplasm
* Manual: N:\Documentatie\CGN\_Databases\IPT\info.docx

This dataset should be updated every two months to ensure that GBIF can harvest recent data. Therefor the script must be executed to generate a new text file. A new archive must be generated within the IPT from that text file.

1. *Centre for Genetic Resources, the Netherlands, Indigenous Trees and Shrubs, published via GBIF data portal*

* File location N:\Documentatie\CGN\_Databases\IPT\BenS\WUR-CGN-BenS-dwc.txt
* Manual to create this file: N:\Documentatie\CGN\_Databases\IPT\info.docx