Food and Agriculture Organization of the United Nations

The International Treaty ON PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE

Digital Object Identifiers for food crops

Descriptors and guidelines of the Global Information System

Pier Augusto Breccia, 2013 - Detail of the painting "Punto di riunione" - Oil on canvas cm 155x140 - Priv. Coll. USA

Digital Object Identifiers for food crops

Descriptors and guidelines of the Global Information System

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2018

Citation: Alercia, A., López, F.M., Sackville Hamilton, N.R. and Marsella, M., 2018. Digital Object Identifiers for food crops - Descriptors and guidelines of the Global Information System. Rome, FAO

ISBN 978-92-5-130418-1

CONTENTS

© FAO/Juerrija Chaves

FOREWORD

World population continues to grow and climate change is contributing to the loss of crop biodiversity. In this context, plant conservation and breeding are essential to increase agriculture productivity and meet the food security challenges. Crop varieties that achieve significantly higher yields and that are able to withstand new diseases and extreme weather events will have to be developed and used sustainably.

To tackle these challenges, the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) of the United Nations (FAO) has been facilitating the exchange of crop material, funding conservation and research projects around the world. And now, through its Global Information System (GLIS), it also helps agricultural researchers, plant breeders and farmers to access relevant scientific and technical information.

The Global Information System of the International Treaty provides a standardized automated one-stop shop for plant genetic resources for food and agriculture (PGRFA) information around the world. It facilitates easy access to information on seeds and other crop material for research, training and plant breeding. It does so through the development and promotion of the use of Digital Object Identifiers (DOIs), an international standard adapted to identify plant germplasm worldwide.

Currently, millions of accessions are conserved in germplasm collections and breeding pools; many of them are duplicates with valuable information that is lost when the material is transferred from one holder to another. Additionally, different user communities such as plant breeders, data curators, researchers and extensionists, often follow different methods to assign identifiers, according to their needs. The lack of standardization had prevented the community from exchanging PGRFA data worldwide which had been for years one of the main challenges for the effective conservation and sustainable use of PGRFA. By using the DOIs standards adopted by the Governing Body of the International Treaty, users will be able to identify and document their plant material uniquely and permanently and will facilitate data interoperability among different systems.

We were delighted to launch the DOI Registration module of the GLIS portal in October 2017. The Governing Body welcomed such quick progress at its Seventh Session in Kigali, Rwanda, in October. By the end of the year, more than half a million materials had been registered in the

System and their passport data made available worldwide in a standard format.

This achievement was possible thanks to the work of experts and technical staff involved in the discussions over the last two years, and in particular, thanks to the support and endorsement of the Scientific Advisory Committee of the International Treaty.

This booklet is one means by which the Secretariat seeks to strengthen capacities of Contracting Parties and National Programmes on the operation of the International Treaty. The booklet helps readers to understand how DOIs can be adopted in their daily work and integrated in their institutional workflow, including what data needs to be provided and how to do it. It is structured in two parts:

Part I of the Manual includes the *Guidelines for the optimal use of Digital Object Identifiers as permanent unique identifiers for PGRFA - v.2*, which describe the main features and benefits of DOIs associated to PGRFA, and a set of basic principles for users to determine when to assign them.

Part II, *Data required for the assignation of Digital Object Identifiers in the Global Information System v.2.1*, referred to as "the descriptors", lists the data to be provided when registering PGRFA material in GLIS through a service that assigns Digital Object Identifiers.

Additional technical documentation required for the registration is provided by the Secretariat of the International Treaty upon request: please send an email to pgrfa-treaty@fao.org.

We are confident that these materials respond to the needs of users of the plant genetic resources community and will have a positive impact on the present and future implementation of the International Treaty. We hope they will also contribute to the way we document plant genetic resources for food and agriculture and to the way we exchange this information globally.

Kent Nnadozie
Secretary
International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA)
Food and Agriculture Organization of the United Nations (FAO)

ACKNOWLEDGEMENTS

The Secretariat of the International Treaty gratefully acknowledges many institutions and individuals that have contributed to, and facilitated the undertaking of this work. All this documentation has been developed and published thanks to the support of the project "*Implementation of the pilot phase of the Global Information System on Plant Genetic Resources of Article 17 of the ITPGRFA*", generously funded by the Federal Ministry of Food and Agriculture of Germany.

Our very special thanks and acknowledgements go to the many experts who participated in the various consultations and in the training workshops for their valuable comments and suggestions.

We also wish to acknowledge the role of the members the Scientific Advisory Committee of the International Treaty who provided useful advice for the development of the descriptors and the guidelines for the optimal use of DOIs.

Finally, we would like to thank all the partner institutions, networks and individuals who provided important inputs for the improvement of documents included in this booklet. Including the CGIAR Centres and their research programmes, particularly the Genebank Platform for the initial testing of the DOIs, which helped to improve the guidelines.

We are grateful to the colleagues of the Secretariat for their contribution to the development of this booklet. Thanks go in particular to Adriana Alercia, Ruaraidh Sackville Hamilton, Marco Marsella and Francisco López. Thanks also go to Nadia Pellicciotta for the layout work.

Kent Nnadozie, Secretary of the International Treaty, had the overall responsibility for this publication.

ACRONYMS AND ABBREVIATIONS

CGIAR	Consultative Group on International Agricultural Research
DOI	Digital Object Identifiers
Easy-SMTA	Easy-Standard Material Transfer Agreement
FAO	The Food and Agriculture Organization of the United Nations
FAQs	Frequently Asked Questions
GLIS	Global Information System
ITPGRFA	International Treaty on Plant Genetic Resources for Food and Agriculture
MCPD	List of Multi-Crop Passport Descriptors
MLS	Multilateral System
PGRFA	Plant Genetic Resources for Food and Agriculture
PID	Permanent IDentifier
SMTA	Standard Material Transfer Agreement
WIEWS	World Information and Early Warning System

Guidelines for the optimal use of Digital Object Identifiers as permanent unique identifiers for Plant Genetic Resources for Food and Agriculture - v.2

1. Introduction

These guidelines are based on a broad consultative process and describe the main features and benefits of Digital Object Identifiers (DOIs) associated to Plant Genetic Resources for Food and Agriculture (PGRFA) and a set of basic principles for users to determine when to assign them.

This document along with *Data required for the assignation of DOIs in GLIS* serve as reference guides for the effective use of DOIs.

2. Background

Several communities[1] have highlighted the importance of creating and adopting Permanent Unique Identifiers for improved identification of PGRFA[2]. The reasons cited include the difficulty of collaboration on conservation, research and breeding without a common standard for identification, and the difficulty of finding information associated with the material. Following broad consultation, DOIs were selected as the most appropriate, web-resolvable digital identifiers.

In accordance with the requirements set out in the International Treaty on Plant Genetic Resources for Food and Agriculture (the Treaty) for the Global Information System (GLIS), the system under construction will (1) build on and facilitate linkage between existing systems and

[1] Including the genebank community, genomics community, plant breeders and journal editors.

[2] The Treaty defines PGRFA as "any genetic material of plant origin of actual or potential value for food and agriculture". This broad definition encompasses not only accessions conserved in genebanks and PGRFA conserved *in situ*, but also breeding lines, research materials, and protected modern varieties.

(2) allow for registration of DOIs applicable to all types of PGRFA. In addition, (3) GLIS will not replace existing systems or duplicate their functionality but provide new services needed by the user community and missing from existing systems, (4) DOIs will be easy to implement, (5) GLIS will also accommodate DOIs created by other systems, and (6) use of the system will be voluntary. Registration of DOIs for PGRFA will be voluntary, and, except for a small number of essential metadata descriptors, most descriptors are voluntary.

It is hoped that DOIs will become the global standard for public identification of PGRFA, facilitating linkage between the material and diverse sources of information associated with the material.

GLIS is able to accept registration requests for PGRFAs that already have a DOI assigned through some other service. Such DOI will have to be provided in the registration request so that GLIS does not assign a new one but rather records the one already assigned. However, we strongly recommend stakeholders to register PGRFAs and obtain DOIs through the GLIS service because this will ensure a coherent metadata structure across all PGRFAs that is critical to obtain the highest effectiveness from the adoption of DOIs. Having a coherent metadata structure will facilitate the offering of advanced services, such as the automatic harvesting of publications referencing GLIS DOIs, that other DOI registration services or stakeholder systems may be unable or unwilling to implement.

3. Bringing new opportunities

The functionality provided by the DOI system brings the following new opportunities for users:

- It exposes the material to the public or collaborators in a format that can be resolved by humans as well as computers.
- It enables information on the material to be harvested by robots searching publications and online databases that refer to the PGRFA by its DOI, and thus to be made more readily available.
- It facilitates access to the information about the PGRFA and related PGRFA by pointing to websites and systems where detailed information is created, maintained and made available to the public.

- It provides a simple way for recipients of material under an SMTA to comply with their obligations under the provisions of Article 6.9 [3] of the SMTA, simply by using the DOI in their publications and online datasets to refer to the material received.
- It helps developers of PGRFA to maintain their records and comply with their obligations under SMTA article 6.5b[4].
- It enables families of related PGRFA to be identified and thus jointly searched. For example, it enables a genebank manager to easily find all publications and online datasets created by recipients of accessions from the genebank.
- It provides a simple reliable mechanism to identify accessions that are duplicated across genebanks.
- It facilitates interoperability between databases, by providing a single common standard for sample identification used by all communities.
- It enables collaborating laboratories, should they wish, to track samples between them with any appropriate degree of precision while each laboratory continues to use its own in-house sample tracking system, providing assurance that they are working on the same material.

The precision that a holder of PGRFA needs for the above functions is a primary criterion in acquiring DOIs for PGRFA.

[3] SMTA article 6.9 states "The Recipient shall make available to the Multilateral System, through the information system provided for in Article 17 of the Treaty, all non-confidential information that results from research and development carried out on the Material"

[4] "In the case that the Recipient transfers a Plant Genetic Resource for Food and Agriculture under Development to another person or entity, the Recipient shall […] (b) identify, in Annex 1 to the new material transfer agreement, the Material received from the Multilateral System […]".

4. What is identified?

DOIs can be used to identify PGRFA held by any individual or organization[5], including genebanks, plant breeders, geneticists, other plant scientists, extension officers, seed companies, plant variety protection offices, gardeners, farmers, landowners, and land managers.[6]

At the discretion of the holder, within the guidelines set out here, the material identified can be any entity recognized as such by the holder. It may be a single DNA sample extracted from a plant, or a single seed or plant or plantlet, or a seed lot contained within a single packet or set of plantlets in one tissue culture tube, or the whole seed lot or set of clonal material harvested from a plot or field, or even multiple generations. The material can be an F1 hybrid, a segregating population, a pure line selected from a mixture or from a segregating population, a mixture of pure lines, or any other genetically homogenous or heterogeneous entity. It may be a landrace or other genetically heterogeneous variety, a modern released cultivar, a genebank accession. It may be formally conserved, for example as in a genebank, or have a transitory existence.

It is therefore essential to include with the DOI information on the nature or category of the PGRFA being identified. The key aspect of this categorization is the real world event that resulted in the PGRFA becoming an entity managed by the holder, such as the collection of a sample from *in situ* conditions, the accession of a sample into a genebank collection, the creation (harvest of seed) of a genetically distinct sample by breeding, the registration of a cultivar in a country, or the first documentation of presence of the PGRFA in a natural habitat. This event is known in the DOI metadata as the method of acquiring or creating the PGRFA, one of the few mandatory descriptors.

[5] The identity of the "holder" is not necessarily clear. For example, in the case of material managed *in situ* or on farm, the holder may be an individual land owner or manager, or the local community jointly responsible for the land, or an organization that owns or manages the land. The decision will depend on the local context, and the DOI system accommodates any of these possibilities.

[6] It follows immediately that the DOI identifies PGRFA within the context of the individual or organization that holds it, and thus one DOI maps to the combination of (1) the identity of holder of the PGRFA with (2) how the holder identifies the physical PGRFA material among all other PGRFA held by the holder.

The DOI identifies the material itself, not the associated data. This is an important distinction. *Inter alia*, it means that if the data associated with the material change, the holder should correct the data without changing the DOI.

5. How is the material identified?

GLIS is not intended to replace existing information systems, and therefore does not replace existing systems for identifying PGRFA. Existing identifiers will continue to be used. In a publication or online article, the first reference to the material would include both its DOI and the local identifier normally used by the holder; subsequent references within a single publication may specify only the local identifier.

However, when PGRFAs are transferred across organizations, locally assigned identifiers have been insufficient to uniquely identify the material. A globally unique, persistent identifier such as a DOI is preferable to maintain consistency over time, provide for proper recognition of rights and obligations, and facilitate access to research outcomes contributed by subsequent recipients of the material.

The assigned DOI should be used to identify the material publicly, especially in electronic media that can be searched online.

6. The PGRFA holder's commitment

A holder of PGRFA who obtains a DOI for a sample of PGRFA makes a commitment to associate that DOI permanently with the material, and not to use the same DOI for any other PGRFA.

Obtaining a DOI does not require or imply any commitment by the holder to maintain the PGRFA alive, and does not change any commitment the holder may or may not already have. If the PGRFA dies or is lost, the DOI persists as an historical record and will not be reused for other PGRFA. This way, any information accrued when the material was available can still be accessed.

Obtaining a DOI does not require or imply any commitment by the holder to make the PGRFA or associated data available to others, and does not change any commitment the holder may or may not already have about the material or the associated information.

7. Relationship with existing systems

Many holders of PGRFA have some form of inventory management system and/or workflow system, with provision for quality control, tracking, data collection and data management. For genebanks this requires the documentation of accessions and their origins, maintaining records of viability, health, genetic integrity and quantity of seeds or clones, tracking progress through viability tests, characterization, and growing out to rejuvenate or multiply stocks, e.g. GRIN-Global. For plant breeders it requires identification and tracking progress through crossing, selection, multiplication, evaluation and release, e.g. Breeding for Results (B4R) being promoted by the CGIAR Excellence in Breeding Platform. For the seed industry it involves tracking progress through seed maintenance from breeders' seed to seed sold to farmers, with appropriate verification of genetic identity. These systems are the primary sources of information on PGRFA, and would normally be the primary sources of data for GLIS.

In addition, some communities have developed portals to expose data on the material they hold and to enable users to search those data. These are typically data warehouses, or secondary data sources, populated by the holders of PGRFA from their primary sources. An example is Genesys (www.genesys-pgr.org), enabling the public to search for accessions available in participating genebanks across the world. Another is the FAO's World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture (WIEWS: http://www.fao.org/wiews).

GLIS is not designed to replace any of these systems, and hence does not duplicate their functionality. Genebanks, breeders and others lacking such functionality may acquire it through relevant capacity-building initiatives.

In order to link existing systems, GLIS needs to maintain in a central repository the data required to identify the corresponding records in the corresponding system. These are the mandatory descriptors (holder, local identifier, scientific name or common name, method, date), which should be uploaded by the holder of the material from a primary data source.

The DOI system makes no assumption about the nature of the documentation system of a holder of PGRFA. It assumes only that the holder can identify the PGRFA held with sufficient precision and permanence to meet the holder's commitments as described above.

8. When to obtain a DOI?

For holders who have opted to use DOIs to publicly identify the PGRFA that they hold, the basic rule for when to obtain a DOI for a PGRFA is simple:

> ***assign a DOI to any PGRFA that you manage and whose present or past existence you wish or need to make publicly known, at the level of detail you choose***

and, as a corollary

> *do **not** assign a DOI to any PGRFA whose present or past existence you do **not** wish or need to make publicly known*

The DOI system provides for a pre-publication embargo period, during which a DOI is assigned but not published. This opens the possibility of assigning DOIs to PGRFA without immediately making their existence publicly known.

Establishment of more specific rules requires consideration of the reasons, or use cases, for making known the existence of the PGRFA. Many different use cases have been identified, but many are specific to particular stakeholder groups. For simplicity, the following sections suggest standards for implementation by specific groups. These may be refined and selfenforced as appropriate to each group.

9. Implementation for genebank managers

The primary objectives of genebank managers are to conserve PGRFA as accessions and provide samples of those accessions to users. They need to conserve rationally (avoiding unnecessary duplication and undesirable gaps), and to facilitate rational use (facilitating the searching of associated information to find the best match between the material available for distribution and material needed by users). They need to inform potential users about the existence of their accessions, and they need to provide descriptive information about their accessions, preferably including information collected by their users. Depending on the crop, one accession may be genetically more or less heterogeneous, and precise monitoring of genetic identity correspondingly difficult.

Hence for genebanks one DOI would typically correspond to one available accession. Standard basic practice for genebank managers would therefore be:

- Obtain a DOI for each accession that is available to users. If known and not confidential, specify also the DOI of the progenitor of the accession. If the accession was acquired from another genebank, the progenitor would be the DOI of the accession in the provider's genebank: hence there could be two DOIs for samples that are intended to be the same genetic material, but conserved by different genebanks. Alternatively, if the providing genebank's accession has no DOI or its DOI is not known, the provider's accession identifier may be specified.
- If the genetic composition of an accession is changed, whether deliberately or unintentionally, such that the new material would be registered as a new accession, obtain a new DOI for the new material. This could include separating a mixed accession into its components, selecting a pure line out of an accession, or discovering a mislabelled sample.
- Different samples of an accession held by the same genebank would normally share the same DOI, unless the genebank manager has a special need to identify the specific samples publicly. The progenitor DOI must be specified, to record that the sample is part of the accession.
- Remind recipients that SMTA Art 6.9 obliges them to make available, through the GLIS, all non-confidential information that results from research and development carried out on the material received, and inform recipients that they can[7] fulfil this obligation by using DOIs registered in GLIS in all their publications and public datasets. This will help the genebank manager associate users' results to the provider's material.

10. Implementation for recipients of PGRFA

Recipients of PGRFA have widely differing needs and capacities. For some users, precise monitoring of genetic identity of the variants can be critically important; the intrinsic genetic variability of genebank

[7] Under SMTA Art 5a, the provider cannot oblige recipients to make their results available to the provider. SMTA Art 6.9 does not state how recipients should make their results available through the GLIS. Hence the provider can do no more than offer DOIs as a mechanism to comply with the recipient's obligations.

accessions can pose significant challenges, and they may need more precise identification than just an accession ID. For other users, this genetic variability may be desirable or unimportant. Some users have their own advanced computerised systems for managing and identifying PGRFA and associated information; others may have nothing.

In the case that a recipient receives material for which the provider has already obtained a DOI, GLIS offers the recipient three options to identify the material received:

1. Use the DOI for the material as registered by the provider;
2. Obtain and use a new DOI to identify the recipient's sample as a distinct entity from the provider's material, and specify that the DOI of the provider's sample is the progenitor of the new DOI;
3. Obtain and use a new DOI to identify the recipient's sample, without specifying the DOI of the progenitor.

Option 1 could be considered the preferred default option for many general users. It would enable associated information to be collated through GLIS without regard for the source of data. However, the principle is similar to asking recipients to provide reference to the provider's accession ID, which has met with limited success in the past. Moreover, it comes with a number of consequences which may be undesirable in some cases:

- Any online data or online publications by the recipient will be associated directly to the provider's material.
- Attribution of the data to the user will be possible only if an independent mechanism is developed to identify the source of data.
- Similarly, separating data created under the control of the provider from other data will be possible only if an independent mechanism is developed to identify the source of data.
- Similarly, if a problem is detected with quality control or standards or genetic integrity, it not will be possible to identify the extent of the problem unless an independent mechanism is developed to do so.
- If the user obtains multiple copies of the same accession from the same provider, it will not be possible to distinguish between the copies unless an independent mechanism is developed.
- If the user distributes the material to a third party using the original provider's DOI, and the third party assigns a new DOI to

the material received, it will be linked to the original provider's DOI, and the role of the original recipient as intermediary will not be public.

Option 2 is preferred when any of the above consequences of option 1 would cause problems. It would be preferred by recipients that want or need separate public identification for material under their management, or recognition through the DOIs of the associated information that they publish. It would be preferred in partnerships where provider and recipient use independent data management systems and need to track exchanges of samples, and wish to do so through DOIs. The optional link to the provider's DOI would enable full flexibility over the scope of searches for information associated with the genetic material: searches could be limited only to the provider's DOI or only to the recipient's DOI, or could include both.

Option 3 would be for recipients who want to make public the existence of their sample and/or provide associated information, but who do not want to publicly disclose the source of their material.

11. Relationships between DOIs

The core novel function of GLIS through the implementation of the DOI System will be to point to various existing databases as a stable and unique reference. The DOI Module of GLIS has the capacity to establish relationships between DOIs, thus connecting records across systems. They are of two kinds:

- A DOI for a PGRFA is related to DOIs for resources containing information about the PGRFA. The latter include online datasets and publications containing data and information about the PGRFA. The user may explicitly declare them within GLIS as "links to associated information" (see the descriptor R01). In addition, GLIS will systematically trawl the web for resources containing references to the DOI for the PGRFA, and will automatically add these to the DOI's links to associated information. This will enable users to easily discover online data and information associated with the PGRFA.
- A DOI for a PGRFA is related to its progenitor(s), which, being themselves PGRFA, may also have a DOI (see the descriptor R02 DOIs of progenitors). The genetic relationship between a PGRFA and its progenitors can be one of several types depending on how the PGRFA came into existence (see

descriptor M04 Method): the DOI may be a genetic copy of, or a variant of, or novel PGRFA incorporating, its progenitor(s). This will enable users to search whole sets of PGRFA: for example, a set of PGRFA that are at least intended to be copies, or the set of PGRFA that are variants of one specified PGRFA, or the set of PGRFA that incorporates one specific progenitor. In the case that a provider transfers a PGRFA to a recipient and the recipient chooses to obtain a separate DOI using the GLIS tools, the provider's DOI will be automatically identified as the progenitor of the recipient's DOI. This will help assuring the accurate documentation of transfers of PGRFA between providers and recipients. In the case that a holder of PGRFA changes the genetic composition of a PGRFA, or wishes to ensure against possible changes in genetic composition (see the next section), it will ultimately be the user's responsibility to ensure that each DOI is correctly associated with its progenitor(s), although GLIS will provide tools to help the user.

12. Managing and using DOIs

- Germplasm holders will prepare to adopt DOIs by adding an additional field in their database that will receive the DOI assigned to each eligible material.
- The Treaty Secretariat will provide an easy-to-use software toolkit that will facilitate the assignment of DOIs and the uploading and correction of associated data.
- If a holder loses a sample for which a DOI has been assigned, the status of the DOI can be changed on the GLIS server to "historic".
- The holder is encouraged to use the DOI in all publications and online articles and databases containing data collected on the germplasm. In a publication or online article, the first reference to the germplasm should include both its DOI and the local identifier normally used by the holder; subsequent references within a single publication may specify only the local identifier.

Further information on the Global Information System and DOIs can be found in FAQs (Frequently Asked Questions)): http://www.fao.org/plant-treaty/areas-of-work/global-information-system/faq/en/

© FAO/Afshaan Shafi

Data required for the assignation of Digital Object Identifiers in the Global Information System v.2.1

This document describes data to be provided when registering Plant Genetic Resources for Food and Agriculture (PGRFA) in the Global Information System (GLIS) through service that assigns Digital Object Identifiers (DOIs). It does not include detailed formatting instructions, which will be provided by the Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) of FAO in forthcoming technical documentation; nor does it provide guidelines on the use of DOIs, which are addressed in a separate document[8]. This document builds upon the results of the Global Survey on Descriptors required for PGRFA and recommendations provided by the Scientific Advisory Committee members[9].

Based on the information contained in this document, additional training material will be developed along with a few more use cases. The first use case for genebanks is provided in Table 1 as an example.

To facilitate the contribution of information to GLIS reducing the burden on participating institutions, collaboration agreements with other existing data management systems will be established to harvest relevant data wherever possible.

Users having already assigned a DOI and willing to join GLIS should register their material providing the already-assigned DOI in the registration request along with other descriptors described in this document. GLIS will record that DOI and not create a new one.

[8] Please see http://www.fao.org/plant-treaty/areas-of-work/global-information-system/guidelines/en/

[9] Report and Analysis of the Global Survey on Descriptors required for PGRFA material, 2015. FAO: http://www.fao.org/3/a-bp470e.pdf

In descriptor R02: DOI(s) of Progenitor(s), references are made to DOIs of other PGRFAs related to the one being registered. Obviously, in order to specify such DOIs, the corresponding PGRFA must have been registered in advance.

Descriptors associated to the DOI in GLIS are grouped into the following classes[10]:

1. CORE DESCRIPTORS

These descriptors are essential for GLIS to offer its core functions. Some of them are fundamental for the registration of material (*Mandatory descriptors*). Others, although not mandatory, should be provided whenever available to support GLIS's more advanced functions.

1a. Mandatory descriptors. Values for these descriptors must be provided for every PGRFA material in GLIS. The holder is not given the option to enter "unknown", "not applicable" or "other", or to leave the field empty; the PGRFA cannot be registered until valid values have been entered for all mandatory fields.

1b. Highly recommended descriptors. These descriptors may not be always available. Valid values must be entered if known, but missing values are allowed.

2. ADDITIONAL DESCRIPTORS

These descriptors are relevant in the appropriate context, depending on the values assigned to some core descriptors; values should be provided if available and applicable. Missing values are allowed.

[10] These classes form an objective classification intended to enhance data quality through helping the registrant to enter correct data. They do not correspond to any subjective classification by "importance" of the descriptors. A mandatory descriptor is not necessarily any more "important" than a highly recommended or additional descriptor. For example, the name of a variety or the country of provenance of a sample may be considered critically important; but they are not always known or applicable and therefore cannot be treated as mandatory.

1a. MANDATORY DESCRIPTORS

M01. Organization/individual conserving the PGRFA

Specify the organization, individual or legal entity conserving the PGRFA by selecting the first applicable form among the following:

1 FAO-WIEWS Institute code[11] (use this if you are acting for an organization and your organization is registered in WIEWS) *[INSTCODE]*[12]

2 Your PID from Easy-SMTA (use this if you have a PID but not a FAO-WIEWS Institute code)

3 Organization name and address (use this if you are acting for an organization and your organization does not have a FAO-WIEWS Institute code or a PID). For *in situ* material this may be the protected area authority, or the owner of the area where material is conserved

4 Individual name and address (use this if you are acting as an individual in your own right and you do not have a PID).

M02. PGRFA unique identifier

The identifier that you use to identify your PGRFA material to distinguish it from other PGRFA conserved by you. Specify precisely one identifier for the material. Make sure that it is unique among the PGRFA conserved by you, and sufficient to enable you to identify the PGRFA when you see this identifier, for example if a future GLIS user contacts you or your successor about it.[13]

For example, it could be an accession ID *[ACCENUMB]*, Selection ID, Derivative name, Population ID, Seed lot ID, Catalogue entry, or any other designation specific to the material.

---

[11] http://www.fao.org/wiews

[12] Italicized terms in square brackets refer to the equivalent descriptor in the FAO/Bioversity Multi-Crop Passport Descriptors (MCPD, v2.1): http://www.bioversityinternational.org/e-library/publications/detail/faobioversity-multi-crop-passport-descriptors-v21-mcpd-v21/)

[13] GLIS will register an error if you attempt to register a second PGRFA with the same PGRFA unique identifier; and a warning if this unique identifier duplicates any other identifier you have registered for another PGRFA.

M03. Date *[ACQDATE]*

Date on which you became the holder of the PGRFA. Date fragments are also accepted, e.g. when only year or year and month are defined.

M04. Method[14]

Considering the date given for the previous field (Date), what event occurred on that date that resulted in you becoming the PGRFA holder?

Choose one of the following:

1 **Acquisition**: you acquired the PGRFA from someone else. In this case descriptor M03 above should contain the date you acquired it. A few examples for illustration:

- You are a genebank manager and you acquired a new accession (from any source, *ex situ* or *in situ*).
- You are a farmer and you bought a variety from the local market or obtained it from some other source.
- You are a breeder or other researcher and obtained the PGRFA from a collaborator, genebank or any other source outside your own breeding or research programme.

2 **In-house copy**: you created the PGRFA by subsampling or taking a harvest from another material that is under your management, with the intent that the new PGRFA should be, as far as possible, a genetic copy of the parent, and you wish to register a new DOI for the new PGRFA rather than use the DOI of the parent. In this case, descriptor M03 above should contain the date you took the material or harvest from the parent, and the descriptor R02 below should contain the DOI of the parent. Some examples:

- You are a genebank manager and you want to publish information about specific components or samples of an accession using a DOI to identify the specific component or sample, and therefore choose to assign different DOIs to each PGRFA of each accession.

[14] Note: this is a key field whose value determines what context-specific descriptors are applicable. It has no equivalent in the MCPD because all genebank accessions have the same value. Some specific examples are given by way of illustration, but these are not exhaustive: you are not expected to try to fit your case into these examples, and GLIS will not record the specific details.

- You are a breeder or researcher using DOIs as the basis of interoperability between your and your collaborators' databases, and you need the specific material (DNA sample, leaf tissue, seed lot, etc.) transferred between you and your collaborators to be identifiable and linked in both databases.

3 **In-house variant**: you created the PGRFA by subsampling or taking a harvest from another material that is under your management, where the parental sample is genetically variable (such as a segregating population and/or physical mixture), and the PGRFA you are registering is a genetic subset of the variability present in the parent. In this case, descriptor M03 above should contain the date you subsampled or harvested the new PGRFA from the parent, and the descriptor R02 below should contain the DOI of the parent. A few examples for illustration:

- You found that the parental material was a mixture of distinctive types, and selected one of those distinct types to be the PGRFA that you are registering.
- You found that the parental material was continuously variable in one or more traits, and selected one type to be the PGRFA that you are registering.
- To enhance the genetic purity of the parental material, you subjected it to one or more generations of purification methods such as single-seed descent, single-plant selection, forced self-pollination or the creation of dihaploid plants.
- You attempted to create a PGRFA that is a genetic copy of its parent, but, based on comparing the DNA or traits of the parent and offspring sample, you discover that genetic drift and/or selection has significantly changed the composition of the offspring sample, and you therefore decide it needs a different DOI.

4 **Novel distinct PGRFA**: The PGRFA is a novel variety or breeding or research material that you have created from one or more parental samples under your management, distinct from its parents, through a process that includes at least one innovative step such as crossing, mutation, or genetic modification. In this case, descriptor M03 above should

contain the date you first harvested the novel distinct material that you are registering, and the descriptor R02 below should contain the DOI(s) of the parent(s). Some examples:

- You deliberately cross-pollinated two or more parents or induced a mutation, and selected a new variety from them.
- You attempted to create a sample that is a genetic copy of its parent, but, based on comparing the DNA or traits of the parent and offspring sample, you discover that there has been unintended cross-pollination from a different variety or a mutant has arisen naturally, and you therefore decide it needs a different DOI.

5 **Observation - Natural**: You hold the PGRFA material *in situ* or on farm and it appeared on your land without your intervention. In this case, descriptor M03 above should contain the date you first observed it on your land.

6 **Inherited**: The PGRFA is one that you have inherited from your predecessor in title. In this case, descriptor M03 above should contain the date you inherited it. A few examples for illustration:

- You have taken on the role of managing a research collection and you don't know how your organization acquired or created the PGRFA.
- You are a farmer and the PGRFA is a traditional variety which you have been maintaining on-farm since you became responsible for the farm.

M05. Scientific name or common name

Specify Genus (e.g. *Manihot*) and specific epithet (e.g. *esculenta*) or common name (e.g. cassava)' as follows:

1 **Genus** *[GENUS, SPECIES]*

Specific epithet

Specific epithet of the scientific name. If unknown, 'sp.' is allowed.

2 **Common name** *[CROPNAME]*

Common name of the taxon, e.g. 'malting barley', 'macadamia', 'maïs'. More than one common name may be provided.

1b. HIGHLY RECOMMENDED DESCRIPTORS

R01. Links to associated information

One or more URLs where further information about the PGRFA can be obtained. This should be the specific page containing information on the PGRFA rather than a web site's landing page[15]. Examples of such URLs are:

1 A URL to your own web site page were you have published information about the PGRFA

2 A URL to any other public data repository where you have placed your data

3 A URL to e-journals, online datasets or other online resources where you have published experimental results using the PGRFA. If a DOI of such resource is available, please provide it in the canonical URL form http://doi.org/{DOI} (e.g. http://doi.org/10.123445/67890).

R02. DOI(s) of progenitor(s)

If you know the DOI of the PGRFA or PGRFA materials from which your PGRFA was derived, specify it here. What the progenitor is depends on how you obtained or created your PGRFA (see M04 above):

- If the method is *Acquisition*, this is the DOI of the PGRFA conserved by the provider. This will be known only if the provider had previously obtained a DOI for his/her PGRFA.
- If the method is *In-house copy*, this is the DOI of the parental PGRFA from which you created your copy.
- If the method is *In-house variant*, this is the DOI of the variable material from which you selected your new PGRFA.
- If the method is *Novel distinct PGRFA,* list the DOIs of any materials that you used to create this novel distinct PGRFA. DOIs can be specified for all direct ancestral materials, regardless of the precise relationship between the ancestral DOIs and the DOI of your novel distinct PGRFA. The relationship can include any type or combination of innovation and selection.

[15] For example, https://www.genesys-pgr.org/acn/id/596419 should be entered instead of https://www.genesys-pgr.org

R03. Biological status *[SAMPSTAT]*

This describes the conditions of provenance of the PGRFA. It follows the MCPD classification. This and descriptor M04 Method above determine what context-specific descriptors (see later on) are applicable.

R04. Additional taxonomic category

Provided the scientific name is specified in M05 above, finer levels of classification should be given if known:

1 **Species authority** *[SPAUTHOR]*
Authority for the scientific name

2 **Subtaxa** *[SUBTAXA]*
Any additional infra-specific taxon: subspecies, variety, form, Group

3 **Subtaxon authority** *[SUBTAUTHOR]*
Authority for the subtaxon at the most detailed level

R05. Names *[ACCENAME]*

One or more registered names or other designations, such as the name of a landrace, traditional variety or modern cultivar, or some other name or designation used to identify a breeder's selection or elite line or variety. This should not duplicate information provided in M02.

R06. Other identifiers *[OTHERNUMB]*

Any other identifiers that have been assigned to identify the PGRFA. It does not include identifiers for other PGRFA materials that you believe maybe similar to this PGRFA.

R07. MLS status *[MLSSTAT]*

The status of the PGRFA with regard to the Multilateral System of Access and Benefit-Sharing (MLS) of the International Treaty on Plant Genetic Resources for Food and Agriculture.

0. Not available under the MLS

1. Available under the MLS; may be further specified as follows:

1.1 The PGRFA is of a crop listed in Annex I and is under the management and control of a Contracting Party to the Treaty and in the public domain

1.2 The PGRFA is in an international collection under Article 15 of the Treaty

1.3 The holder received the PGRFA with an SMTA

1.4 The holder has voluntarily placed the PGRFA in the MLS

1.5 The PGRFA is derived from, and distinct from, material previously received from the MLS, is still under development and not yet ready for commercialization, and may be made available at the discretion of the developer, with an SMTA.

R08. Historical PGRFA

This descriptor indicates whether or not the material identified by the DOI currently exists. For example, users may wish to obtain DOIs to identify historical material referred to in online publications or datasets. If a PGRFA is permanently lost after being registered, the PGRFA holder should change the descriptor value to "1=Yes".

0. No (material currently exists)

1. Yes (material no longer physically exists)

© FAO/Aishaan Shafi

2. ADDITIONAL DESCRIPTORS

2.1 Context: material you acquired from someone else

These descriptors apply when method is *Acquisition* (see M04). Their use is particularly encouraged when the provider's DOI (see R02) is unknown[16], to provide an alternative link back to the provider's PGRFA material. In future, GLIS will attempt to reconcile missing DOIs by using the information provided here.

A01. Provider's location *[DONORCODE], [DONORNAME]*

Location or name of the person or organization that provided the PGRFA to you, following equivalent rules to M01).

A02. Provider's PGRFA unique identifier *[DONORNUMB]*

Unique identifier used by the provider to identify the PGRFA under the provider's management.

A03. Country of provenance[17] *[ORIGCTY]*

The country in which the PGRFA material was either collected or bred or selected, or the first country in the known history of the PGRFA.

2.2 Context: material previously collected from in situ conditions

These descriptors apply to all PGRFA that you hold *ex situ* and that were previously collected from *in situ* conditions, either collected by you or your organization, or collected by someone else and subsequently transferred to you.

[16] And especially encouraged when information on provenance is not available through a target (R01), for example, a page in Genesys.

[17] Equivalent in some cases to the "country of origin". Country of provenance is used here with a practical definition applicable to PGRFA that avoids potential confusion with country of origin as defined in the Convention on Biological Diversity.

A04. Collector's location *[COLLCODE],[COLLNAME],[COLLINSTADDRESS]*

Location of the home base of the person(s) or organization(s) that originally collected the PGRFA from *in situ* conditions, following equivalent rules to location to M01.

A05. Collector's PGRFA unique identifier *[COLLNUMB]*

Identifier assigned by the collector(s) to the PGRFA collected.

A06. Collecting mission identifier *[COLLMISSID]*

If the PGRFA was collected as part of an organized collecting mission during which other PGRFA materials were collected, and the collectors assigned a code to identify the mission, specify that mission identifier here.

A07. Location where the PGRFA was collected *[COLLSITE]*

Location information below the country level that describes where the PGRFA was collected. This might include the distance in kilometres and direction from the nearest town, village or map grid reference point, (e.g. "7 km South of Curitiba in the state of Parana").

A08. Latitude *[LATITUDE / DECLATITUDE]*

Latitude of the location where the PGRFA was collected.

A09. Longitude *[LONGITUDE / DECLONGITUDE]*

Longitude of the location where the PGRFA was collected.

A10. Uncertainty *[COORDUNCERT]*

Uncertainty of the latitude/longitude coordinates of the location where the PGRFA was collected. This value is typically provided by georeferencing software.

A11. Geodetic *datum* *[COORDDATUM]*

The geodetic *datum* or spatial reference system upon which the latitude/longitude coordinates of the collecting location are based.

A12. Georeferencing method *[GEOREFMETH]*

The method used to estimate latitude/longitude coordinates of the location where the PGRFA was collected.

A13. Elevation *[ELEVATION]*

Elevation of collecting site expressed in metres above sea level.

A14. Collecting date *[COLLDATE]*

Date on which the PGRFA was collected. Date fragments are also accepted when only year or year and month are known.

A15. Collecting source *[COLLSRC]*

A description of the nature of the location where the PGRFA was collected. It follows the MCPD classification.

2.3 Context: material bred

These descriptors apply to all PGRFA that were bred with human intervention and have information on how it was bred, whether by you or by someone else, *in situ* or *ex situ*.

A16. Breeder's location *[BREDCODE], [BREDNAME]*

Location where the material was bred, following equivalent rules to M01.

A17. Ancestry *[ANCEST]*

The pedigree (e.g. 'Hanna/7*Atlas//Turk/8*Atlas') or other description of the ancestry of the PGRFA and how it was bred (e.g. 'mutation found in Hanna', or 'cross involving amongst others Hanna and Irene').


10.18730/59KYG
10.18730/59KXF
GLIS
10.18730/59KM6
10.18730/59KH3

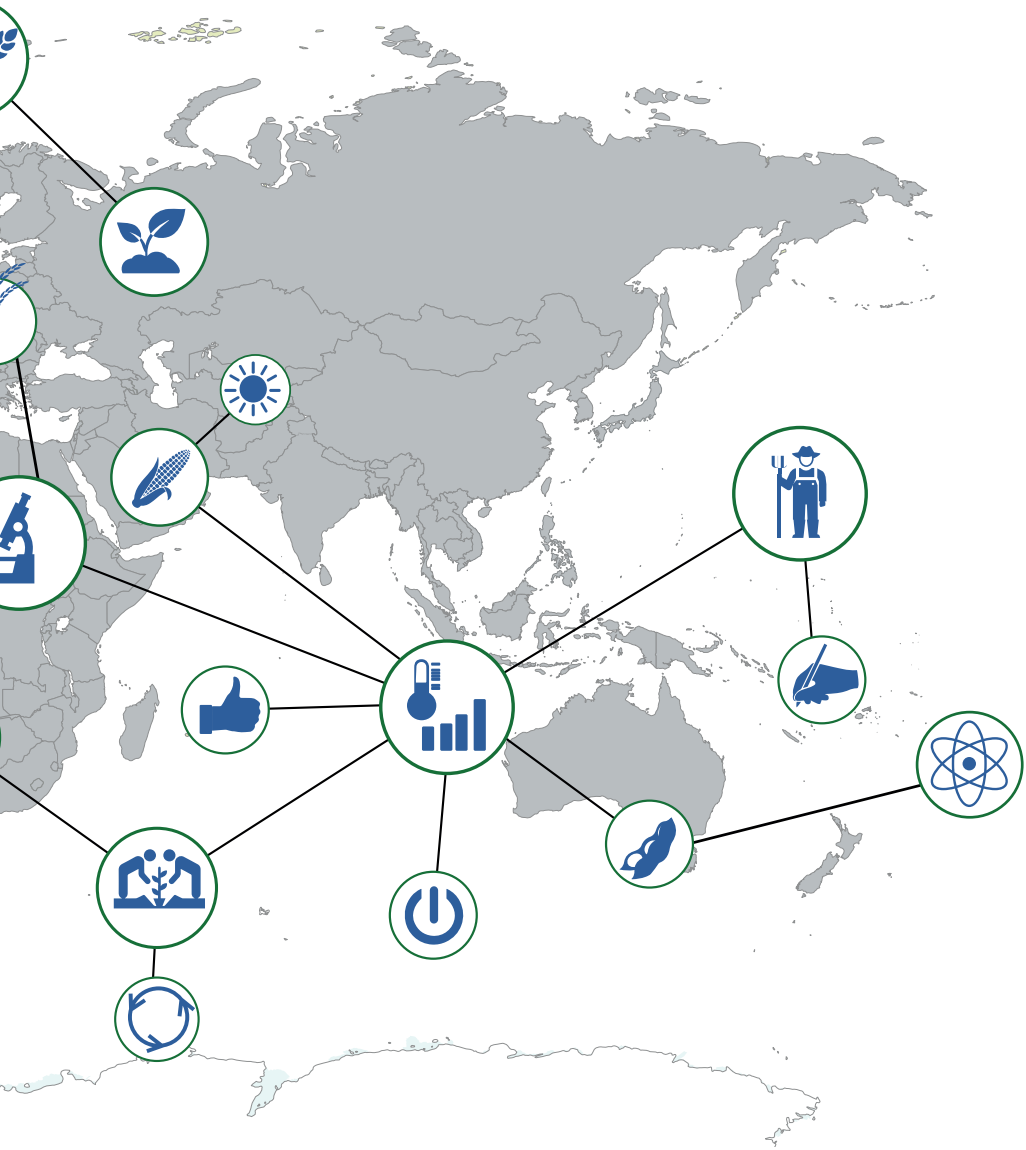


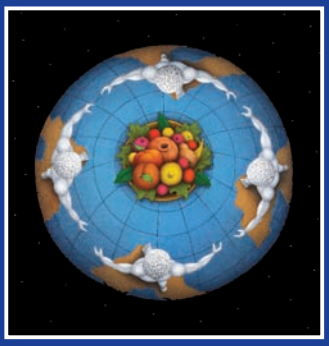
Table 1. Use case for genebanks: Mapping to MCPD v.2.1 descriptors

1. CORE DESCRIPTORS		
1a. MANDATORY DESCRIPTORS		
GLIS descriptor	**GLIS Description**	**MCPD equivalent for genebanks**
M01. Organization/ individual conserving the PGRFA material	The location where the PGRFA is maintained	Institute code *[INSTCODE]*
M02. PGRFA unique identifier	The identifier that you use to identify your PGRFA material to distinguish it from other PGRFA conserved by you	Accession number *[ACCENUMB]*
M03. Date	Date on which you became the holder of the PGRFA	Acquisition date *[ACQDATE]*
M04. Method	Considering the date given for field C03 Date, what event occurred on that date that resulted in you becoming the PGRFA holder	N/A
M05. Scientific name or common name	Genus and specific epithet or common name. Specify one or both	Genus *[GENUS]* Species *[SPECIES]* Common crop name *[CROPNAME]*

1b. HIGHLY RECOMMENDED DESCRIPTORS		
R01. Links to associated information	One or more URLs where further information about the PGRFA can be obtained	N/A
R02. DOI(s) of progenitor(s)	The DOI of the PGRFA or PGRFA materials from which your PGRFA was derived	N/A
R03. Biological status	This describes the conditions of provenance of the PGRFA	Biological status of accession *[SAMPSTAT]*
R04. Additional taxonomic category	Species authority: Authority for the specific epithet Subtaxa: Any additional infra-specific taxon: subspecies, variety, form, Group Subtaxon authority: Authority for the subtaxon	Species authority *[SPAUTHOR]* Subtaxon *[SUBTAXA]* Subtaxon authority *[SUBTAUTHOR]*
R05. Names	One or more registered names or other designations, such as the name of a landrace, traditional variety or modern cultivar, or some other name or designation used to identify a breeder's selection or elite line or variety	Accession Name *[ACCENAME]*
R06. Other identifiers	Any other identifiers that have been assigned to identify the PGRFA material. It does not include identifiers for other materials that you believe maybe similar to this PGRFA.	Other identifiers associated with the accession *[OTHERNUMB]*
R07. MLS status	The status of the PGRFA with regard to the Multilateral System of Access and Benefit-Sharing (MLS) of the International Treaty on Plant Genetic Resources for Food and Agriculture	MLS status of the accession *[MLSSTAT]*
R08. Physical existence	Describes whether the PGRFA is still available or permanently lost after being registered	N/A

2. ADDITIONAL DESCRIPTORS		
2.1 Context: Material you acquired from someone else		
A01. Provider's location	Location or name of the person or organization that provided the PGRFA to you	Donor institute code *[DONORCODE]* Donor institute name *[DONORNAME]*
A02. Provider's PGRFA unique identifier	Unique identifier used by the provider to identify the PGRFA under the provider's management	Donor accession number *[DONORNUMB]*
A03. Country of provenance	The country in which the PGRFA material was either collected or bred or selected, or the first country in the known history of the PGRFA.	Country of origin *[ORIGCTY]*
2.2 Context: Material previously collected from *in situ* conditions		
A04. Collector's location	Location of the home base of the person(s) or organization(s) that originally collected the PGRFA from *in situ* conditions	Collecting institute code *[COLLCODE]* Collecting institute name *[COLLNAME]* Collecting institute address *[COLLINSTADDRESS]*
A05. Collector's PGRFA unique identifier	Identifier assigned by the collector(s) to the PGRFA collected	Collecting number *[COLLNUMB]*
A06. Collecting mission identifier	The identifier, if any, of the mission during which the PGRFA was collected	Collecting mission identifier *[COLLMISSID]*
A07. Location where PGRFA was collected	Location information below the country level that describes where the PGRFA was collected	Location of collecting site *[COLLSITE]*

A08. Latitude	Latitude of the location where the PGRFA was collected	Latitude of collecting site *[DECLATITUDE] [LATITUDE]*
A09. Longitude	Longitude of the location where the PGRFA was collected	Longitude of collecting site *[DECLONGITUDE] [LONGITUDE]*
A10. Uncertainty	Uncertainty of the latitude/longitude coordinates of the location where the PGRFA was collected	Coordinate uncertainty *[COORDUNCERT]*
A11. Geodetic *datum*	The geodetic *datum* or spatial reference system upon which the latitude/longitude coordinates of the collecting location are based	Coordinate datum *[COORDDATUM]*
A12. Georeferencing method	The method used to estimate latitude/longitude coordinates of the location where the PGRFA was collected	Georeferencing method *[GEOREFMETH]*
A13. Elevation	Elevation of collecting site	Elevation of collecting site *[ELEVATION]*
A14. Collecting date	Date on which the PGRFA was collected	Collecting date of sample *[COLLDATE]*
A15. Collecting source	A description of the nature of the location where the PGRFA was collected	Collecting/acquisition source *[COLLSRC]*
2.3 Context: Material bred		
A16. Breeder's location	Location where the material was bred	Breeding institute code *[BREDCODE]* Breeding institute name *[BREDNAME]*
A17. Ancestry	The pedigree or other description of the ancestry of the PGRFA and how it was bred	Ancestral data *[ANCEST]*



Access to information is a non-monetary benefit of the International Treaty on Plant Genetic Resources for Food and Agriculture. The Global Information System was launched in 2017 and can be accessed at

https://ssl.fao.org/glis

FOR MORE INFORMATION CONTACT:

International Treaty on Plant Genetic Resources for Food and Agriculture
Food and Agriculture Organization of the United Nations
Viale delle Terme di Caracalla • 00153 Rome • Italy
Tel: +39 0657055430 • Fax: +39 0657056347
E-mail: pgrfa-treaty@fao.org
URL: www.planttreaty.org and www.fao.org/plant-treaty/en

ISBN 978-92-5-130418-1

I8840EN/1/05.18