



Food and Agriculture
Organization of the
United Nations



The International Treaty
ON PLANT GENETIC RESOURCES
FOR FOOD AND AGRICULTURE

Implementing DOIs

1st Scientific Advisory Committee meeting
24-25 November 2016, Rome

Marco Marsella
Secretariat of the International Treaty, FAO

Overview

- How DOIs were selected
- Descriptors associated to DOIs
- The GLIS DOI registration service
- Integration Toolkit
- Updates on initial adoption of DOIs
- DOI System initiatives potentially relevant to GLIS
- Final remarks and feedback received
- Advice sought

The DOI selection process

- Request from the community for PUID - **long standing issue**
- Late 2014, technical analysis of most promising PUID types
- Presentation of findings at COGIS in January 2015
 - DOI, ARK and LSID shortlisted
- Further investigation and expert consultation
 - 17 participants
 - scoring each PUID type against 20 indicators
- DOIs selected as the most promising PUID for GLIS
- Service agreement with DataCite to offer DOI minting **free-of-charge** to the GLIS user community

GLIS descriptors for PGRFA (1/2)

- Why do we need descriptors?
 - resolution: from DOI to PGRFA record
 - discovery: from one or more PGRFA attributes to DOI
- Based on **MCPD 2.1** and **conceptual work behind IRIS**, with **generalisations** for breeders and other user communities
- **Global survey** with over 200 participants from 60 countries held in July-August 2015
- **Focus group** of 15 experts followed to validate survey results

GLIS descriptors for PGRFA (2/2)

- **Minimal** set of mandatory descriptors on **readily available** information
 - mandatory (essential to identify PGRFA)
 - highly recommended (should be provided if available)
 - context dependent (should be provided if available and applicable e.g. info on collection or pedigree)
- **First version** of the descriptors was published in October 2016 for comments
- **Guidelines** on the adoption of DOIs also published in October 2016 for comments
- Descriptors and guidelines **to be validated** with breeders and other user communities
 - Indonesia BSF project is a first opportunity

DOIs for PGRFA in the GLIS context

- The DOI is **associated to the physical object** not to its description
 - the description can change if required due to
 - incorrect identification
 - taxonomy changes
- Effects of material transfer on DOIs
 - the recipient's material get a new DOI
- Recipient's DOI and provider's DOI are related
- A GLIS DOI will look like **10.18730/12345678**
 - the prefix **10.18730/** will be the same for all GLIS DOIs
 - the rest is not much longer than most Accession Numbers

Targets

- The way to “attach” information to PGRFAs registered in GLIS
- Targets are **links to websites** where additional information on the PGRFA can be found
- Targets can point to virtually anything
 - passport data
 - characterization and evaluation data
 - details on collecting missions
 - genomics, phenomics and other “omics”
 - pictures and diagrams
 - traditional knowledge, recipes
 - trade statistics, quarantine regulations
- Targets are **classified** according to a controlled vocabulary on the kind of information that can be found at their destinations

Relations

- Relations among PGRFAs are maintained by GLIS
 - creation from ancestry
 - derivation/selection from mixtures or populations
 - transfer from providers
- Targets associated to PGRFAs related to the current one are made available as “related targets”
 - access to information generated by recipients
 - access to information available on ancestry

The GLIS DOI registration service (1/2)

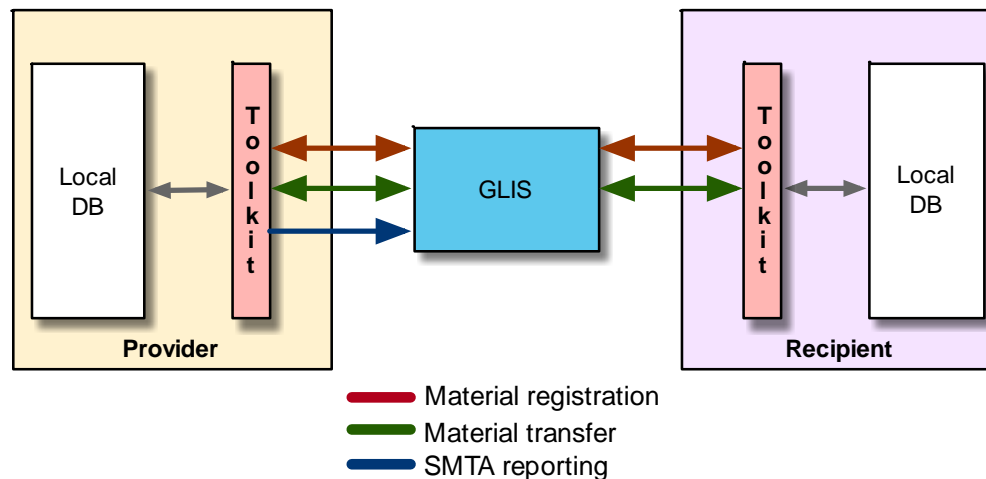
- Initial GLIS module designed to respond to community's needs
- Mapping of GLIS descriptors to DataCite metadata 4.0
 - cleared by the DataCite Metadata Working Group
- Several expressions of interest on the service and descriptors
- What if the PGRFA being registered already has a DOI assigned?
 - some institutions are planning to mint their own DOIs
 - GLIS descriptors are considered for adoption
 - no problem, GLIS will accept already assigned DOIs
- Registration and update options:
 - Small collections: multilingual web user interface
 - Medium collections: batch initial registration, new registrations and updates through the web user interface
 - Large collections: published XML-based integration protocol

The GLIS DOI registration service (2/2)

- Advanced query API
 - content negotiation (XML, JSON)
 - response filtering and expansion
 - rate limitation
 - HATEOAS
- Core services
 - registration of PGRFA to get the newly assigned GLIS DOI
 - registration of PGRFA with DOI already assigned by registrant
 - update of descriptors associated to a GLIS DOI
 - recording of material transfer (through Easy-SMTA)
 - resolution of DOIs and discovery of material from descriptors
 - PGRFA detail page with relations to other PGRFAs and related targets

The GLIS Integration Toolkit

- Published XML-based integration protocol
- To further promote GLIS integration, a Toolkit is proposed
 - easily deployed and configured to integrate with a local DB
 - extensive security, configuration and logging functions
 - GLIS transactions
 - Easy-SMTA reporting transactions
 - local DB can be published using standard formats and protocols



Collaboration with FAO's IT Division (CIO)

- CIO is collaborating with the Secretariat to:
 - provide a hosting solution for the GLIS web-based platform
 - develop GLIS using FAO standard technologies and methodologies
 - ensure corporate support and long term sustainability for the GLIS IT infrastructure
 - facilitate the integration of GLIS with relevant FAO initiatives
- **First output** will be the Toolkit prototype to be tested in the context of the Indonesia BSF project early next year

Adoption of DOIs

- **Voluntary**: it will succeed only if DOIs meet users' needs
- Initiatives already ongoing to promote DOIs and GLIS
 - **Indonesia BSF project**: Collaboration to ensure GLIS DOI registration and material transfer transactions meet user needs
 - **CGIAR** is planning to:
 - incorporate DOIs into genebank data management systems, GRIN-Global and Genesys
 - register DOIs for all CGIAR accessions by GB7
 - **National genebanks**: Genesys/Eurisco model + DOIs extended to 2 other regions
 - **Breeders**: CGIAR's Excellence in Breeding platform to establish data exchange standards for breeding, in principle based on DOIs
 - **Genomics community**: DivSeek approved adoption of DOIs in principle

Potentially relevant initiatives related to DOIs

- DataCite/CrossRef **EventData**
 - will allow GLIS to find publications where its own DOIs are cited and make them available as targets in the corresponding PGRFA page
 - GLIS may be unable to do so for DOIs minted by other parties (to be verified)
- DataCite/CrossRef/ORCID **Organization Identifier Project**
 - fills the gap between content (DOIs) and contributors/researchers (ORCID)
 - could be considered in addition to or as replacement for Easy-SMTA PIDs and WIEWS institute codes

Final remarks (1/3)

- The DOI is perpetually associated to the physical material
 - If material is lost, DOI stays
- DOI is immutable, but associated description of the material can change as required
- Genebanks don't need to replace Accession Numbers with DOIs
- Communities not using Accession Numbers or other suitable identifiers are encouraged to use DOIs to uniquely and permanently identify their materials
- New DOI is assigned to recipient's material and linked to provider's DOI
- DOIs are not a requirement for material transfer under the MLS

Final remarks (2/3)

- Recipients are encouraged to store **provider's DOI (if available)** along with provider's Accession Number and register a new DOI for the material received
- Use of DOIs is encouraged in publications and registration of related DOIs
 - this is being promoted by publishers to facilitate accurate identification of the material object of the research
 - basis of advanced services offered by the DOI System (e.g. EventData)
- Providers and recipients should use their own DOIs in publications
 - the DOI of the material should be cited at least once in the publication

Final remarks (3/3)

- Publications are moving from paper-based media meant for humans to far richer digital content for humans and machines alike
- Where there is no system in place, the Treaty will promote selected systems to facilitate adoption of DOIs and improve data quality
 - GRIN-Global for genebanks
 - ??? for breeders
 - EURISCO model for regions
 - built-in Toolkit functions

What's in it for users?

- Facilitated access to information
 - providers are not aware of outcomes of research carried out by recipients
 - recipients struggle to find information on the material received
 - information is dispersed over many systems
- Improved and more homogeneous data quality
- Support of formats and protocols for machine access to information
- Better integration among genebanks, breeders and other user communities
- Easier compliance with SMTA stipulations on information sharing
- Access to tools and systems to address local shortcomings
- Access to advanced services developed by the DOI community

Advice sought

- Discussion and validation of descriptors and guidelines with other user communities
- Standard formats and protocols that GLIS should support
- Further work on controlled vocabularies for the classification of targets
- Challenges and opportunities to promote the adoption of DOIs, particularly with breeders
- Other advice for an effective and successful adoption of DOIs