

Passport data and intake

How a sample becomes a traceable accession

Identifier	SAMEN-PROV0001-P11
Author	Thomas Schüller, Syntagma Forschung Wien · ORCID 0009-0003-9799-6747
Contact	forschung@tschueller.com
Version and date	1.1 · 2026-09-23
Rights	Text all rights reserved; tools and data layers MIT

Abstract

Paper P04 describes the data architecture of conservation from above — from the catalogues. This paper describes it from below: what must be recorded when a sample arrives so that it is later findable, legally usable and reconstructable? The basis is the Multi-Crop Passport Descriptors format. The paper derives from it an intake form that is manageable even for small organisations without a database, and explains which fields are indispensable. The templates are in the appendix of the package. The paper is the basis for chapter 13 of the book.

1. Why intake is decisive

Three things cannot be produced afterwards:

1. **Origin.** Whoever takes in a sample without an indication of origin can later neither fulfil the due diligence obligation under Regulation (EU) No 511/2014 nor make use of a derogation of market law (paper P03).
2. **Initial state.** Without a germination test and moisture determination at intake there is no initial value, no monitoring interval and no regeneration threshold (paper P09).
3. **Relation to other samples.** Whether a sample is a duplicate, a subsample or an origin of its own is known at intake — and not later.

2. The descriptor format

The Multi-Crop Passport Descriptors, version 2.1, jointly issued by FAO and Bioversity International, define the fields with which genebanks describe their samples. They are the common format of Genesys, EURISCO and the national inventories. The most important fields:

Field	Meaning	Required for the catalogue
INSTCODE	institute code of the holding institution according to the FAO register	yes
ACCENUMB	accession number, unique within the institution	yes
GENUS	genus	yes
SPECIES	specific epithet	recommended
ACCENAME	variety or origin name	recommended

Field	Meaning	Required for the catalogue
ACQDATE	date of acquisition	recommended
ORIGCTY	country of origin (three-letter country code)	recommended
COLLSITE	collecting site as text	recommended
DECLATITUDE , DECLONGITUDE	coordinates in decimal degrees	recommended
COLLDATE	collecting date	recommended
SAMPSTAT	biological status, such as wild, landrace, breeding material, cultivar	recommended
COLLSRC	collecting source, such as wild habitat, farm, market, genebank	recommended
DONORCODE , DONORNUMB	donor institution and its number	where available
OTHERNUMB	other numbers of the same material	where available
DUPLSITE	institute code of the safety duplicate	where available
STORAGE	type of conservation, such as short-, medium- or long-term seed, field, in vitro, cryo	recommended
MLSSTAT	status within the Multilateral System of the International Treaty	recommended
PUID	persistent unique identifier, such as a DOI	recommended
REMARKS	remarks	free

Three fields are sufficient for an entry in Genesys: accession number, genus and institute code. That is the lower limit, not the goal. Every further field increases the value of the entry; the origin fields most of all, because they decide usability.

3. The accession number

A good accession number is unique, unchangeable and low in meaning. Three rules:

1. **Never reassign.** A number once used remains used, even if the sample is lost.
2. **Do not build in meaning.** Numbers encoding species, year or place force renumbering as soon as an assignment turns out to be wrong.
3. **Subsamples with a suffix, not a new number.** The format itself does not regulate subsamples; an in-house suffix to the accession number keeps the relation to the original material visible.

4. The intake form

An intake form that an association can keep without a database can be derived from the format. It has three blocks.

Block A — Identifier and origin (at intake, unchangeable): accession number, date of intake, genus, species, variety name, biological status, country of origin, collecting or source site, coordinates, collecting date, collecting source, donor person or institution, their number, evidence of lawful access, material transfer agreement yes or no.

Block B — State on arrival (at intake, once): quantity in seeds or grams, thousand-seed weight, seed moisture content, germination with sample and date, fill or cut test, infestation, seed storage behaviour.

Block C — Conservation (continuously updated): conservation route, storage location, storage conditions, next test date, duplicate at which site, growers for multiplication, cultivation and multiplication notes.

Block A is not changed after intake; corrections are added as notes. Block C is the storage register. The templates in the appendix implement these blocks: as a table whose columns follow the field names of the format, and as a print version for paper filing.

5. From table to catalogue

An intake form that follows the field names of the format can be reported to the national inventory without rework. The route for a small organisation:

1. Apply for an institute code (INSTCODE). The codes are assigned by FAO through its WIEWS information system; in Austria in coordination with the national inventory (genbank.at).
2. Keep the table with the field names of the format.
3. Report the entries to the national inventory; from there they flow to EURISCO and Genesys.
4. Once a persistent identifier has been assigned, add it in the PUID field.

6. What a data record must achieve

The data layer of a holding has three tasks, which follow from the preceding papers:

Task	Fields	Paper
Findability	identifier, taxonomy, name, institute code	P04
Legal usability	origin, evidence of access, material transfer agreement	P03
Reconstructability after loss	origin, coordinates, duplicate site, variety description	P06

The third task is most often overlooked. After the loss of the genebank in Aleppo and the relocation of the Kharkiv collection (paper P06), the data record was what made reconstruction possible: it was known what had existed and where a duplicate was held. Passport data must therefore be secured separately from the material.

Limits of this paper

The paper reproduces the descriptor format in extracts; the published version with its code lists is authoritative. Reporting routes to the national inventory differ by country; for Austria the route via genbank.at is described, and the specific formalities must be requested from the responsible body. The statement on the role of data in reconstruction is derived from the documented cases, not from an analysis of the data holdings concerned.

References

1. Alercia, A.; Diulgheroff, S.; Mackay, M. (2015): *FAO/Bioversity Multi-Crop Passport Descriptors V.2.1 [MCPD V.2.1]*. Rome: FAO and Bioversity International.
2. Genesys PGR: *Accession passport data basics*. <https://www.genesys-pgr.org/documentation/basics>
3. Genesys PGR: *Frequently asked questions*. <https://www.genesys-pgr.org/content/about/frequently-asked-questions>
4. FAO: *WIEWS — World Information and Early Warning System on PGRFA*. <https://www.fao.org/wiews/resources/en>
5. ARCHE NOAH: *Sortenbestände und Sortendaten* (reference to genbank.at). <https://www.arche-noah.at/sortenerhaltung/samenarchiv/sortendaten/>
6. Regulation (EU) No 511/2014 on compliance measures for users from the Nagoya Protocol.
7. Papers P03, P04, P06 and P09 of this strand.