

Datenbanken und Informationssysteme

Wo pflanzengenetische Ressourcen nachgewiesen sind und wie man sie abfragt

Kennung	SAMEN-PROV0001-P12
Verfasser	Thomas Schüller, Syntagma Forschung Wien · ORCID 0009-0003-9799-6747
Kontakt	forschung@tschueller.com
Fassung und Stand	2.3 · 2026-09-28
Rechte	Text alle Rechte vorbehalten; Werkzeuge und Datenschichten MIT
Vorbehalt	Schnittstellen und Abfragemuster nach der Dokumentation der Betreiber (Abruf 24.09.2026), nicht nach eigenen Abfragen.

Kurzfassung

Fachpapier P04 beschreibt die Datenkette der Erhaltung, Fachpapier P11 den Datensatz an ihrem Anfang. Dieses Papier beschreibt ihr Ende: die Informationssysteme, in denen pflanzengenetische Ressourcen nachgewiesen sind, und den Weg, sie abzufragen. Es ordnet vierzehn Systeme nach Reichweite und Datenart, zeigt die Kennungen, die sie verbinden, und gibt Abfragemuster, Schnittstellen und Nutzungsbedingungen wieder. Grundlage ist die Datenschicht `datenbanken.json`, auf der auch der Datenbank-Wegweiser beruht. Alle vierzehn Einträge tragen dort den Status „teilgeprüft“: Die Angaben stammen aus der Dokumentation der Betreiber, nicht aus eigenen Abfragen.

1. Die Landschaft

1.1 Nach Reichweite

Abbildung 12 „Datenbanklandschaft“ ordnet die Systeme in drei Ebenen. **National** führen Genbanken ihre Bestände: in Österreich genbank.at der AGES, in Deutschland GBIS/I des IPK Gatersleben, in den USA GRIN-Global. **Europäisch** fasst EURISCO die nationalen Inventare zusammen; AEGIS ist die darauf aufbauende virtuelle europäische Genbank. **Weltweit** stehen Genesys als Passdatenportal, WIEWS als Berichts- und Codesystem der FAO, GLIS als Vergabestelle dauerhafter Kennungen und mehrere benachbarte Bestände, die nicht für Genbanken gebaut wurden, aber für die Erhaltung gebraucht werden.

1.2 Nach Datenart

System	Reichweite	Datenart	Schnittstelle
Genesys PGR	weltweit	Passdaten	ja, mit Anmeldung
FAO WIEWS	weltweit	Berichte, Einrichtungen	nicht belegt
GLIS	weltweit	Kennungen (DOI)	ja, Anmeldung nur zum Registrieren
Seed Information Database	weltweit	Samenmerkmale	nicht belegt
GBIF	weltweit	Vorkommen, Taxonomie	ja, Suche ohne Anmeldung
Crop Ontology	weltweit	Merkmale, Vokabular	ja
NCBI GenBank / ENA	weltweit	Sequenzen	ja, ohne Anmeldung
IUCN Rote Liste	weltweit	Gefährdung	ja, mit Token

System	Reichweite	Datenart	Schnittstelle
World Flora Online	weltweit	Taxonomie	ja, ohne Anmeldung
EURISCO	Europa	Passdaten, Merkmale	keine dokumentierte
ECPGR AEGIS	Europa	Verbund	nein
GRIN-Global	USA	Passdaten, Taxonomie, Merkmale	nicht belegt
IPK GBIS/I	Deutschland	Passdaten, Bestellung	BrAPI, Einzelheiten nicht geprüft
genbank.at	Österreich	Passdaten, Merkmale, Bestellung	nicht belegt

„Nicht belegt“ heißt, dass die Datenschicht keine dokumentierte Schnittstelle nachweist, nicht, dass keine besteht.

2. Was die Systeme leisten und was nicht

2.1 Passdatenkataloge

Genesys PGR des Crop Trust ist der größte Nachweis von Ex-situ-Akzessionen im Format der Multi-Crop Passport Descriptors. Die Startseite nannte am 24. September 2026 rund 4,54 Millionen Akzessionen; der Jahresbericht 2025 des Crop Trust spricht von über 4,46 Millionen Mustern, rund drei Viertel des Weltbestands von etwa 5,9 Millionen (Fachpapier P04). Genesys hält kein Material, sondern vermittelt Anfragen an die haltende Genbank.

EURISCO, vom IPK Gatersleben im Auftrag des ECPGR betrieben, weist über zwei Millionen Akzessionen von über 400 Einrichtungen nach, mit Merkmalsdaten und AEGIS-Kennzeichnung. **AEGIS** umfasst laut ECPGR rund 122.000 Akzessionen, die nach vereinbarten Standards erhalten und nach den Bedingungen des Saatgutvertrags abgegeben werden; die Einzelheiten stehen in EURISCO.

National liefert **genbank.at** die österreichischen Bestände an EURISCO und bietet Pass-, Charakterisierungs- und Evaluierungsdaten sowie die Musteranforderung. **IPK GBIS/I** weist rund 151.000 Akzessionen der deutschen Bundes-Ex-situ-Genbank nach und gibt Material mit Materialübertragungsvereinbarung ab. **GRIN-Global** führt über 600.000 aktive Akzessionen des US-Genbanksystems und die Taxonomie GRIN Taxonomy.

Keiner dieser Kataloge sagt etwas darüber, ob ein Muster keimfähig oder abgabebereit ist.

2.2 Kennungen und Berichte

FAO WIEWS vergibt die Einrichtungscodes (INSTCODE), weist über 17.000 Einrichtungen nach und trägt die Weltzustandsberichte und den Nachhaltigkeitsindikator 2.5.1. **GLIS**, das Informationssystem des Internationalen Saatgutvertrags, vergibt kostenlos DOIs für pflanzengenetisches Material; laut Startseite beim Abruf rund 2,1 Millionen.

2.3 Benachbarte Bestände

- **Seed Information Database** (Society for Ecological Restoration mit Kew): Lagerverhalten, Samengewicht, Keimbedingungen; laut Betreiber rund 182.000 Datensätze zu rund 54.800 Taxa. Sie beantwortet, ob eine Art in die Samenbank kann (Fachpapiere P02 und P07).
- **GBIF**: Vorkommens- und Belegdaten, darunter Wildverwandte von Kulturpflanzen.
- **IUCN Rote Liste**: Version 2026-1 mit rund 176.000 bewerteten Arten, davon über 49.500 bedroht (Fachpapier P18).
- **World Flora Online**: globale Konsensklassifikation mit Namensabgleich (Fachpapier P13).

- **Crop Ontology:** einheitliche Begriffe für Merkmale, Methoden und Skalen je Kulturart, damit Merkmalsdaten verschiedener Genbanken vergleichbar werden.

3. Kennungen als Brücke

Die Systeme sind nicht technisch verknüpft, sondern über Kennungen.

Kennung	Vergeben von	Eindeutig	Verwendet in
INSTCODE, z. B. AUT001	FAO WIEWS	weltweit	allen Passdatenkatalogen
Akzessionsnummer (ACCENUMB)	haltende Einrichtung	nur mit INSTCODE und Gattung	Genesys, EURISCO, nationale Systeme
DOI, Präfix 10.18730	GLIS	weltweit, dauerhaft	Genesys, EURISCO (falls gemeldet), IPK
Taxon-Kennungen (WFO-ID, taxonKey, NCBI-Taxonomie-ID, IUCN-Taxon-ID)	je System	nur im eigenen System	WFO, GBIF, INSDC, IUCN

Der Einrichtungscodex besteht aus dem dreistelligen Ländercode nach ISO 3166 und einer Nummer und ist der Schlüssel des Datenmodells (Fachpapier P04). Die DOI über GLIS ersetzt den zusammengesetzten Schlüssel durch eine einzige dauerhafte Kennung, die im Passdatenfeld **PUID** eingetragen wird (Fachpapier P11). Für Namen gibt es keine solche Brücke: Der Übergang zwischen den Taxon-Kennungen läuft über den wissenschaftlichen Namen und ist anfällig für Synonyme.

4. Abfrage in Schritten

Die Datenschicht ordnet den Systemen neun Fragen zu. Daraus ergibt sich eine Reihenfolge:

1. **Namen klären** in World Flora Online oder GRIN Taxonomy, Synonyme notieren.
2. **Akzessionen suchen**, zuerst national (genbank.at, IPK GBIS/I), dann in EURISCO und Genesys.
3. **Kennungen auflösen**: Einrichtungscodex über WIEWS, DOI über GLIS.
4. **Umfeld prüfen**: Lagerverhalten in der Seed Information Database, Vorkommen in GBIF, Gefährdung in der Roten Liste.
5. **Material anfordern** bei der haltenden Genbank, nicht beim Katalog.

Für neun Systeme nennt die Datenschicht Abfragemuster. Als verifiziert gekennzeichnet sind nur Muster, die an der offiziellen Dokumentation nachgesehen wurden; eine eigene Abfrage liegt auch ihnen nicht zugrunde.

System	Zweck	Muster	Verifiziert
Genesys	Übersicht einer Genbank	https://www.genesys-pgr.org/wiews/{INSTCODE}	ja
GLIS	DOI auflösen	https://doi.org/10.18730/{DOI-SUFFIX}	ja
WIEWS	Einrichtungen und Codes	https://www.fao.org/wiews/data/organizations/en/	ja
GRIN-Global	Taxonomiesuche	https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysearch	ja
World Flora Online	Namen abgleichen	https://list.worldfloraonline.org/matching_rest.php?input_string={NAME}	ja

System	Zweck	Muster	Verifiziert
GBIF	Vorkommen, maschinenlesbar	<code>https://api.gbif.org/v1/occurrence/search?scientificName={NAME}&country={LAND2}</code>	ja
GBIF	Vorkommen im Webportal	<code>https://www.gbif.org/occurrence/search?q={NAME}</code>	nein
IUCN	Art nachschlagen, mit Token	<code>https://api.iucnredlist.org/api/v4/taxa/scientific_name?genus_name={GATTUNG}&species_name={ART}</code>	ja
Crop Ontology	Variablen einer Ontologie	<code>https://cropontology.org/brapi/v1/variables/{ONTOLOGIE}</code>	ja
ENA	Sequenzdaten eines Taxons	<code>https://www.ebi.ac.uk/ena/portal/api/search?result=read_run&query=tax_tree({TAXID})</code>	ja
NCBI	Nukleotidsuche	<code>https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=nuccore&term={NAME}</code>	nein

Für EURISCO, AEGIS, genbank.at, IPK GBIS/I und die Seed Information Database führt der Weg über die Suchmaske.

5. Schnittstellen und Lizenzen

Genesys bietet eine Schnittstelle unter `api.genesys-pgr.org` mit Anmeldung über OAuth2, als Login oder mit Client-Zugangsdaten über den Helpdesk; auch BrAPI verlangt Anmeldung. Für R gibt es das Paket `genesysr`. Die Nutzungsbedingungen verlangen Quellenangabe und untersagen die Weitergabe der Daten an Dritte; eine Veröffentlichung verweist auf das Portal statt auf mitgelieferte Rohdaten.

EURISCO hat keine dokumentierte Schnittstelle (so auch Frontiers in Plant Science 2026); Suche und Export laufen über die Weboberfläche, die Nutzungsbedingungen sind weltweit, lizenzfrei und nicht exklusiv. **IPK GBIS/I** dokumentiert BrAPI-Endpunkte; der Passdaten-Schnappschuss 2025 steht unter CC BY-SA 4.0. **GRIN-Global** ist als Werk der US-Regierung gemeinfrei. **GLIS** verlangt Anmeldung nur zum Registrieren und Aktualisieren.

GBIF erlaubt die Suche ohne Anmeldung, Downloads mit Konto. Die Lizenz gilt je Datensatz: CC0, CC BY 4.0 oder CC BY-NC 4.0. Mischt eine Abfrage diese, bestimmt die engste, was mit dem Ergebnis geschehen darf. **World Flora Online** steht unter CC0. Die Sequenzarchive sind frei zugänglich; **NCBI** erlaubt drei Anfragen je Sekunde, mit Schlüssel zehn. Die **IUCN** verlangt ein Token und erlaubt nur nicht kommerzielle Nutzung mit Zitierpflicht.

6. Datenqualität

Dubletten. Dasselbe Material liegt oft in mehreren Genbanken. Die Felder `DONORCODE`, `OTHERNUMB` und `DUPLSITE` halten das fest, sofern sie ausgefüllt sind (Fachpapier P11). Hinzu kommt die Dublette der Systeme: Ein Datensatz aus genbank.at erscheint auch in EURISCO und Genesys. Wer Zählungen verschiedener Portale addiert, zählt ihn mehrfach.

Veraltete Taxonomie. Genesys stützt sich auf die Taxonomie des US-Genbanksystems (Fachpapier P04), andere Systeme auf eigene Listen. Ein vor Jahren vergebener Name kann heute ein Synonym sein; eine Suche nur nach dem aktuellen Namen übersieht ihn.

Fehlende Koordinaten. Sammelort und Koordinaten sind im Passdatenformat empfohlen, nicht verpflichtend. Einträge ohne sie zählen mit, lassen sich aber keinem Herkunftsgebiet zuordnen. Eine Messung ihres Anteils liegt diesem Strang nicht vor.

Genbankdaten und Vorkommensdaten. Ein Passdatensatz beschreibt ein Muster, das eine Einrichtung führt. Ein Vorkommensdatensatz in GBIF beschreibt, dass ein Organismus an einem Ort zu einer Zeit beobachtet oder belegt wurde. Das eine belegt verfügbares Material, das andere eine Population, die heute erloschen sein kann. Für Wildverwandte zeigt ein Vorkommen ohne Genbankmuster eine Sammellücke an.

Akzessionszahl und keimfähige Akzessionen sind verschiedene Größen (Fachpapier P04): Die bayerische Wildpflanzen-Genbank fand bei 21 Prozent der Akzessionen mehr als die Hälfte der Samen leer, befallen oder tot. Kein Katalog zeigt das.

7. Sequenzdaten

Die Archive der International Nucleotide Sequence Database Collaboration (NCBI, EMBL-EBI, DDBJ) führen Sequenzen, Reads und Assemblies und sind nach ihrer Richtlinie frei zugänglich. Das macht sie zum Kern der Debatte um digitale Sequenzinformationen: Der Vorteilsausgleich wurde für physisches Material entworfen, die Sequenz ist frei verfügbar (Fachpapier P03). Das Lenkungsorgan des Saatgutvertrags fand im November 2025 dazu keinen Konsens. Nach Beschluss 16/2 der Biodiversitätskonvention sollen große Nutzer als Richtwert 1 Prozent des Gewinns oder 0,1 Prozent des Umsatzes in den Cali-Fonds zahlen; die Mittelaufbringung blieb bis 2026 weit hinter den Erwartungen zurück (Datenschicht recht.json).

Die Archive arbeiten mit eigenen Kennungen, der INSDC-Zugangsnummer und der NCBI-Taxonomie-ID. Eine gemeinsame Kennung zwischen Genbankmuster und Sequenz weist die Datenschicht nicht nach; der Einstieg ist das Taxon, wie im Muster für ENA.

8. Das Werkzeug

Der Datenbank-Wegweiser geht von der Frage aus, nicht vom System: Man wählt eine der neun Fragen, erhält die zuständigen Systeme mit Schnittstellen, Kennungen und Nutzungsbedingungen und lässt aus Name, Gattung, Art oder Einrichtungscodes Abfrageadressen bauen, etwa für Genesys, GBIF, ENA und die IUCN. Daneben erklärt er die Ordnungssysteme (Fachpapier P13). Eigene Bestände im Passdatenformat führt die Sammlungsverwaltung; der Sammlungsatlas zeigt Sammlungen und Nachweissysteme mit Datenstand und Beleglage.

Abbildungen

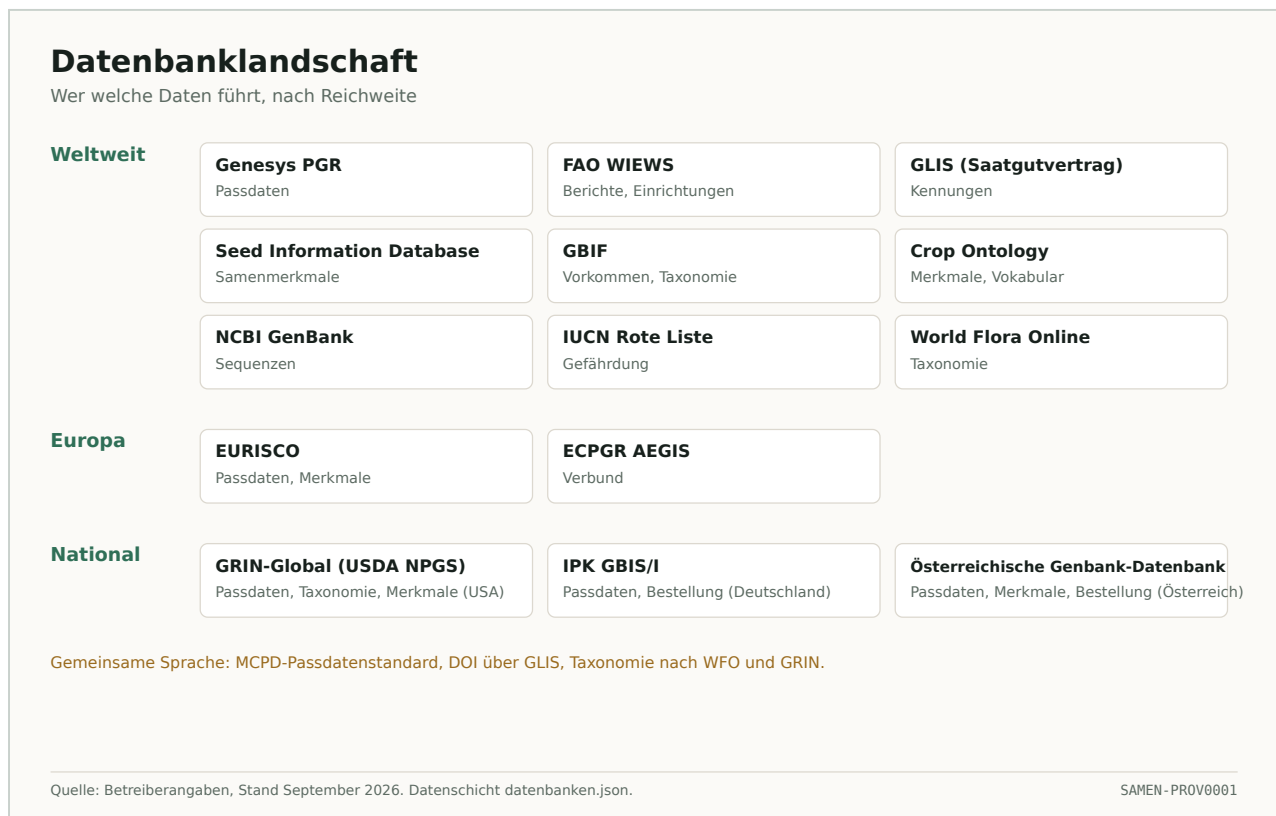


Abbildung 1. Datenbanklandschaft der pflanzengenetischen Ressourcen. Quelle: Datenschicht des Strangs, siehe Fußzeile der Abbildung.

Grenzen dieses Papiers

Das Papier gibt den Stand der Schnittstellen vom September 2026 wieder, erhoben aus der Dokumentation der Betreiber am 24. September 2026; alle Systeme tragen den Status „teilgeprüft“. Sieben Abfragemuster wurden zusätzlich durch einen Abruf erprobt: Die Muster für Genesys (Einrichtungsseite), die GBIF-Programmierschnittstelle und die ENA-Suche lieferten gültige Antworten; das Muster der World Flora Online lieferte nur die Dokumentation, jenes der Crop Ontology einen Serverfehler, NCBI und die GBIF-Weboberfläche lassen keinen automatischen Abruf zu. Das Ergebnis steht im Datenbank-Wegweiser beim jeweiligen Muster. Adressen, Abfragemuster und Nutzungsbedingungen ändern sich. Mengenangaben weichen je nach Quelle und Stand ab, wie bei Genesys. Die Auswahl der vierzehn Systeme ist nicht vollständig; weitere nationale Inventare, Herbaratenbanken und kulturartspezifische Systeme fehlen. Die Datenqualität der Kataloge wurde nicht geprüft.

Literatur

1. Genesys PGR: *APIs und Terms and conditions of use*. <https://www.genesys-pgr.org/documentation/apis>; <https://www.genesys-pgr.org/content/legal/terms>
2. Crop Trust, Jahresbericht 2025: *Digitizing genebanks — Genesys and GGCE*. <https://report.croptrust.org/2025/our-work/digitizing-genebanks-genesys-and-ggce/>
3. IPK Gatersleben: *Genebank documentation*. <https://www.ipk-gatersleben.de/en/research/genebank/genebank-documentation>

4. Aguilar, C. H. M.; Oppermann, M.; Weise, S. (2026): FAIRness at all costs? Practical limitations in applying the FAIR principles to plant genetic resources data. *Frontiers in Plant Science* 17, 1888670. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2026.1888670/pdf>
5. ECPGR: *AEGIS — European accessions*. <https://www.ecpgr.org/aegis/european-collection/european-accessions>
6. IPK Gatersleben: *Information and ordering; Breeding API*. <https://www.ipk-gatersleben.de/en/infrastructure/gen-e-bank/information-and-ordering>; <https://fair-ipk.ipk-gatersleben.de/public/breedingapi.html>
7. AGES: *Österreichische Genbank*. <https://www.genbank.at/en/>
8. USDA Agricultural Research Service: *GRIN*. <https://www.ars-grin.gov/>
9. FAO: *WIEWS*. <https://www.fao.org/wiews/en/>
10. Internationaler Saatgutvertrag: *GLIS*, Startseite und *FAQs*. <https://glis.fao.org/glis/>; <https://glis.fao.org/glis/static/en/faqs.html>
11. Society for Ecological Restoration: *Seed Information Database*. <https://ser-insr.org/seed-information-database>
12. GBIF: *API documentation; Terms of use*. <https://techdocs.gbif.org/en/openapi/>; <https://www.gbif.org/terms>
13. Crop Ontology: *API help*. https://cropontology.org/api_help
14. EMBL-EBI: *ENA policies*. <https://ena-browser-docs.readthedocs.io/en/latest/about/policies.html>
15. NCBI: *Entrez Programming Utilities Help*. <https://www.ncbi.nlm.nih.gov/books/NBK25497/>
16. IUCN: *Red List API; The IUCN Red List of Threatened Species*. <https://api.iucnredlist.org/>; <https://www.iucnredlist.org/>
17. World Flora Online: *Background*. <https://wfoplantlist.org/background>
18. CBD: *Decision 16/2*. <https://www.cbd.int/doc/decisions/cop-16/cop-16-dec-02-en.pdf>
19. Fachpapiere P02, P03, P04, P07, P11, P13 und P18 dieses Strangs.