

Schriften zu Genetischen Ressourcen

Schriftenreihe der Zentralstelle für Agrardokumentation und -information
Informationszentrum für Genetische Ressourcen

Band 12

Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland

Herausgeber diese Bandes

F. Begemann

S. Harrer

J.D. Jiménez Krause

Herausgeber:

Zentralstelle für Agrardokumentation und -information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)
Villichgasse 17, D – 53177 Bonn
Postfach 20 14 15, D – 53144 Bonn
Tel.: (0228) 95 48 - 202
Fax: (0228) 95 48 - 220
Email: igr@zadi.de

Druck:

Druckerei Martin Roesberg
Geltorfstr. 52
53347 Alfter-Witterschlick

Schutzgebühr 17,- DM

ISSN 0948-8332

© ZADI Bonn, 1999

Inhaltsverzeichnis

Inhaltsverzeichnis

i

Vorwort des Projektträgers

iii

G. WILHELM

Einführung in die Workshop-Dokumentation

1

B. BECKER

Hintergrund und Entstehung von BIG

3

F. BEGEMANN

BIG als Verbundprojekt

6

B. BECKER

Verbundpartner ZADI

11

S. HARRER UND J.D. JIMSENEZ KRAUSE

Verbundpartner IPK

14

H. KNÜPFER, K. BACHMANN, N. BIERMANN UND J.

Verbundpartner RUB

17

T. STÜTZEL UND A. JAGEL

Verbundpartner BfN

19

R. MAY

BIG-Suchoberfläche – Ein Prototyp

21

J.D. JIMÉNEZ KRAUSE UND S. ROSCHER

Thematische Einführung in die Arbeitsgruppen

.....
57

J. OCHSMANN, N. BIERMANN, H. KNÜPFER UND K. BACHMANN

SysTax - ein Datenbanksystem für Systematik und Taxonomie

.....
64

J. HOPPE, E. BOOS UND TH. STÜTZEL

Datengrundlagen des Bundesamtes für Naturschutz für das Bundesinformationssystem
Genetische Ressourcen

.....
80

H.G. FINK UND R. MAY

Bundesinformationssystem Genetische Ressourcen – Konzept und erste Ergebnisse

.....
93

B. BECKER

Pflanzengenetische Ressourcen (PGR) im Informationssystem für Genetische Ressourcen
(GENRES) – Online-Demonstration

105

E. MÜNCH

Ausgewählte internationale Informationssysteme zu pflanzengenetischen Ressourcen im
Internet

113

S. HARRER, J.D. JIMÉNEZ KRAUSE UND F. BEGEMANN

Das Internet-Informationsangebot des N.I. Vavilov Institute of Plant Industry (VIR)

121

S. HARRER UND A. OMELCHENKO

Entwicklung des deutschen Clearing-House-Mechanismus

126

H. FREIBERG UND R. SPECHT

Abkürzungsverzeichnis

AGIS	<i>Agricultural Genome Information System</i>
BAZ	Bundesanstalt für Züchtungsforschung an Kulturpflanzen
BCC	<i>Barley Core Collection</i> – Core-Sammlung von Gerste
BfN	Bundesamt für Naturschutz
BIG	Bundesinformationssystem Genetische Ressourcen
BMBF	Bundesministerium für Bildung und Forschung
BML	Bundesministerium für Ernährung, Landwirtschaft und Forsten
BMU	Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit
BSA	Bundessortenamt
CBD	<i>Convention on Biological Diversity</i> – Übereinkommen über die biologische Vielfalt (ÜBV)
CGIAR	<i>Consultative Group on International Agricultural Research</i> – Beratungsgruppe für Internationale Agrarforschung
CGN	<i>Centre for Genetic Resources</i> – Zentrum für Genetische Ressourcen, Niederlande
CHM	Clearing-House Mechanismus des Übereinkommens über die biologische Vielfalt
CITES	<i>Convention on the International Trade of Endangered Species of Wild Fauna and Flora</i> – Übereinkommen über den Internationalen Handel mit gefährdeten Arten freilebender Fauna und Flora, Washingtoner Artenschutzübereinkommen
EADB	<i>European Avena Database</i> - Europäische <i>Avena</i> -Datenbank
EBDB	<i>European Barley Database</i> - Europäische Gersten-Datenbank
EBI	<i>European Bioinformatics Institute</i>
ECCDB	<i>European Central Crop Database</i> – Europäische Fruchtartendatenbank
ECP/GR	<i>European Cooperative Programme on Crop Genetic Resources Networks</i> Europäisches Kooperationsprogramm zu pflanzengenetischen Ressourcen
EMBL	<i>European Molecular Biology Laboratory</i>
EST	<i>Expressed sequenced tagged sites</i>
EVA	Informationssystem für Evaluierungsdaten pflanzengenetischer Ressourcen
FAO	<i>Food and Agriculture Organisation</i> – Landwirtschafts- und Ernährungsorganisation der Vereinten Nationen
FFH	Flora-Fauna-Habitat-Richtlinie der Europäischen Union
GIS	Geographisches Informationssystem
GRIN	<i>Genetic Resources Information Network</i> - Informationsnetzwerk zu Genetischen Ressourcen in den USA
ICBN	<i>International Code of Botanical Nomenclature</i> – Internationaler Code für botanische Nomenklatur
IDBB	Internationale Datenbank für <i>Beta</i>
IGR	Informationszentrum für Genetische Ressourcen
IPGRI	<i>International Plant Genetic Resources Institute</i> – Internationales Institut für Pflanzengenetische Ressourcen
IPK	Institut für Pflanzengenetik und Kulturpflanzenforschung
IT	Informationstechnologie

LSV	Landessortenversuch
MLU	Martin-Luther-Universität Halle
NCBI	<i>National Center for Biotechnology Information</i>
NGB	<i>Nordic Gene Bank</i> – Nordische Genbank, Schweden
PCR	<i>Polymerase Chain Reaction</i> – Polymerase-Ketten-Reaktion
PGR	Pflanzengenetische Ressourcen
QTL	<i>Quantitative Trait Locus</i>
RUB	Ruhr-Universität Bochum
ÜBV	Übereinkommen über die biologische Vielfalt
VIR	<i>N. I. Vavilov Institute of Plant Industry in St. Petersburg</i> – Vavilov Institut für Pflanzenzüchtung, St. Petersburg, beherbergt die Genbank
VSK	Vertragsstaatenkonferenz
ZADI	Zentralstelle für Agrardokumentation und -information

Überblick über Konzepte und Entwicklungsschritte zur Dokumentation und Information pflanzengenetischer Ressourcen in Deutschland

Overview of concepts and developments in the area of documentation and information of plant genetic resources in Germany

FRANK BEGEMANN¹

Abstract

This paper describes the steps undertaken in Germany to develop a concept for a comprehensive documentation and information of plant genetic resources. On behalf of the German Ministry for Food, Agriculture and Forestry (BML), BOMMER and BEESE had prepared a concept for a central documentation in 1990. While the authors left it open how such a central documentation should be achieved, there has been a series of four workshops bringing together many stakeholders in this field to discuss central and decentral models.

Today, an internet-based mixed model has evolved. It includes decentral elements of activities offered through GENRES (www.genres.de), a central navigation catalogue PGRDEU (www.genres.de/pgrdeu) of minimum passport descriptors as agreed by FAO, IPGRI and ECP/GR and accession-specific information sources that are held decentrally and are partly offline. This mixed model would ensure a high data quality as the maintenance of databases remains closely with the germplasm holding institution and, at the same time, would enable even uninformed users to quickly retrieve interesting accessions and activities in Germany and abroad. Besides, it would not require standardizing database software. Based on this concept, the BIG-project (www.genres.de/big) would facilitate to include also information related to other genetic resources held *ex situ* in botanic gardens and *in situ*, using *inter alia* geographic information systems (GIS). BIG would also try to get the central catalogue function and the accession-specific data held decentrally „time-wise“ closer together as it develops a central view for searches on decentral servers „on the fly“. BIG would be the platform to link specific moduls as i.e. the EVA-modul for evaluation data (www.genres.de/eva).

In light of international developments in the field of documentation and information of genetic resources as well as the legal and political framework, *inter alia* the importance of property and copy rights and the implications of the Convention on Biodiversity (CBD), it seems important to closely cooperate with the German Clearing House Mechanism of the CBD (www.biodiv-chm.de). The Information Centre for Genetic Resources (IGR) of the German Centre for Documentation and Information in Agriculture (ZADI) which is coordinating the above-mentioned activities and projects appreciates the efficient cooperation with all its partners in this field which is crucial for any success.

¹ Zentralstelle für Agrardokumentation und –information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

Das Übereinkommen über die biologische Vielfalt (CBD) verfolgt drei Ziele: die Erhaltung der biologischen Vielfalt, die nachhaltige Nutzung ihrer Bestandteile und die ausgewogene und gerechte Aufteilung der sich aus der Nutzung der genetischen Ressourcen ergebenden Vorteile, insbesondere durch angemessenen Zugang zu genetischen Ressourcen und angemessene Weitergabe der einschlägigen Technologien unter Berücksichtigung aller Rechte an diesen Ressourcen und Technologien sowie durch angemessene Finanzierung (BMU 1993). Vorteile aus der Nutzung der genetischen Ressourcen können sich nur ergeben, wenn mindestens zwei zentrale Komponenten zusammentreffen, nämlich das biologische Material sowie die Information über das biologische Material. Jede Komponente für sich genommen ist nur eingeschränkt wirtschaftlich verwertbar.

Schon frühzeitig gab es daher auch in Deutschland Bestrebungen, nicht nur die Erhaltung der pflanzengenetischen Ressourcen systematisch voranzutreiben, sondern immer auch das entstehende Wissen über das erhaltene Material wissenschaftlich abgesichert und solide zu dokumentieren und die damit verbundenen Informationen verfügbar zu machen. 1990 legte eine Projektgruppe des Bundesministeriums für Ernährung, Landwirtschaft und Forsten (BML) unter Leitung von Prof. Bommer ein Konzept zur Erhaltung und Nutzung der pflanzengenetischen Ressourcen in Deutschland vor, das gerade auch die Notwendigkeit der Dokumentation unterstrich sowie die Verfügbarkeit und den Austausch von Daten national und international verbessern sollte (BOMMER & BEESE 1990). Das Konzept sah eine zentrale Dokumentationseinheit als Teil des Informationszentrums für Genetische Ressourcen (IGR) der Zentralstelle für Agrardokumentation und –information (ZADI) vor. Es sollte die Datenerfassung und –speicherung in einzelnen Programmen unterstützen und die Kompatibilität bestehender Systeme sicherstellen. Es sollte somit der Kern eines nationalen Informationsnetzes sein. Bommer und Beese sahen alternativ zwei Modelle der Datenverwaltung in diesem Informationsnetz, die in der Zukunft auf ihre Vor- und Nachteile hin überprüft werden sollten (BOMMER & BEESE 1990):

- (1) „Ein Modell, bei dem dezentral fruchtartenspezifische Datenbanken an den für die Koordination zuständigen Institutionen eingerichtet und unterhalten werden. Die jeweiligen Institute wären somit für sämtliche Verwaltungs-, Wartungs- und Entwicklungsaufgaben sowie für die Dienstleistung bei der Informationssammlung, -aufarbeitung und -weitergabe allein zuständig. Die zentrale Dokumentationseinheit hätte lediglich die Funktion, auf die jeweiligen Datenzugriffsmöglichkeiten hinzuweisen. Der internationale Datenaustausch würde auf der Basis der bestehenden Beziehungen zwischen den jeweiligen Institutionen für bestimmte Bereiche pflanzengenetischer Ressourcen ablaufen und in seinen Standards den speziellen Bedürfnissen angepaßt werden.
- (2) Ein Modell der zentralisierten Datenspeicherung, bei dem alle Daten zur Verfügbarkeit und Beschreibung pflanzengenetischer Ressourcen in Deutschland zusammengefaßt werden. Für die Aktualisierung der Datenbestände wären in erster Linie die jeweiligen Koordinatoren für die verschiedenen Bereiche zuständig. Dies würde durch eine *Online*-Verbindung zwischen der zentralen Datenbank und den peripheren PC-Systemen erreicht.

Zur Entlastung der Koordinatoren in den Forschungsinstitutionen wäre die zentrale Dokumentationseinheit des IGR auch in erster Linie für Dienstleistungen bei der Informationsbereitstellung zuständig.“

Obwohl gerade im Bereich der Informations- und Kommunikationstechnologien (ICT) rasante Entwicklungsfortschritte gemacht werden, die auch für das geschilderte Anliegen gravierende Auswirkungen haben, kann dennoch festgestellt werden, daß die seinerzeit aufgeworfenen Fragen und Modelle nach wie vor aktuell sind. Aufbauend auf der Konzeption von 1990 wurden inzwischen vom IGR vier spezielle Arbeitstagen organisiert, die die Vor- und Nachteile der beiden unterschiedlichen Modelle je nach ICT-Entwicklungsstand wie auch nach institutionellen Rahmenbedingungen neu überprüfen sollten und Vorschläge für die weiteren Entwicklungsschritte machten.

Zunächst fand im November 1992 in Braunschweig eine Tagung statt, auf der eine Bestandsaufnahme der verschiedenen Datensammlungen und Dokumentationsaktivitäten erfolgte. Wichtige Teile der Bestandsaufnahme erstreckten sich neben den großen Genbanken in Gatersleben mit seinen Außenstellen in Dresden-Pillnitz, Groß Lüsewitz und Malchow sowie Braunschweig auf die Arbeiten am Geilweilerhof zur Rebe, bei der Deutschen Sammlung für Mikroorganismen und Zellkulturen (DSMZ), den Genbanken für forstliche Genressourcen, den Botanischen Gärten, dem Bundessortenamt und beim Bundesamt für Naturschutz (BfN) (ANON. 1993). Dieses Vorgehen entsprach einem komplett dezentralen Modell. Im Frühjahr 1993 wurde dennoch auf einer zweiten Tagung in Bonn ein erster Entwurf für ein ORACLE-basiertes zentrales Dokumentations- und Informationssystem vorgestellt, das künftig weiter beraten werden sollte (SCHMIDT 1993). Nicht zuletzt aufgrund der rasanten ICT-Entwicklungen in dieser Zeit wurde auch dieses Modell wieder verworfen und im Herbst 1995 auf einer dritten Tagung, die wiederum in Braunschweig stattfand, ein dezentrales Mischmodell mit zentraler Komponente entwickelt, das weitgehend vom sich entwickelnden Internet profitieren sollte. Dieses Modell fand seither weite Akzeptanz und wurde auf der vierten Tagung im Frühjahr 1998 in Bonn weiter diskutiert und ausgestaltet.

Dieses Mischmodell sieht zum einen ein zentrales Informationssystem von Verweisen auf dezentral im Internet angebotene Informationsquellen vor; dieses System GENRES wird vom IGR im Internet unter <http://www.genres.de> angeboten und kontinuierlich aktualisiert. Aufgrund von sich bietenden Rationalisierungseffekten wird GENRES nicht nur für den Bereich der landwirtschaftlichen und gartenbaulichen Kultur- und Nutzpflanzen, sondern auch für Forst- und sonstige Wildpflanzen, Haus- und Nutztiere, Mikroorganismen, aquatische Genressourcen sowie relevante politische, rechtliche und institutionelle Rahmenbedingungen entwickelt. Eine zweite Komponente des Mischmodells ist ein zentraler Suchkatalog im Internet (PGRDEU), der basierend auf international u.a. von FAO, IPGRI und ECP/GR abgestimmten sogenannten „minimalen Deskriptorenlisten“ nur ausgewählte Paßportdaten enthält. PGRDEU wird vom IGR gemeinsam mit den sich daran beteiligenden Partnern unter <http://www.genres.de/pgrdeu> angeboten und erleichtert dem Nutzer das schnelle Auffinden einzelner Akzessionen in deutschen Sammlungen. Für eine Suche nach ausführlichen Infor-

mationen einzelner Akzessionen verweist PGRDEU wiederum auf dezentrale Datenbestände, die z.T. auch nur *offline* vorhanden sind. Dieses Mischmodell ermöglicht eine hohe Datenqualität, da die Daten bei den Institutionen gehalten und aktualisiert werden, die sie erheben und dokumentieren, und gleichzeitig ermöglicht es auch dem nicht-informierten Nutzer das schnelle Auffinden von interessanten Akzessionen bzw. Aktivitäten in Deutschland und darüber hinaus. Außerdem erfordert dieses Modell keine Standardisierung bei der dezentral verwendeten Datenbanksoftware; dies erhöht seine breite Akzeptanz bei den mitwirkenden Partnern.

Eine weitere Entwicklung der Dokumentations- und Informationssysteme erscheint inzwischen auch vor dem Hintergrund der sich durch das Übereinkommen über die biologische Vielfalt veränderten Situation beim Zugang zu genetischen Ressourcen und den relevanten Informationen sowie den schnellen Wissenszuwächsen im Bereich der biotechnologischen Forschung erforderlich. Einerseits muß man wohl davon ausgehen, daß sich der Zugang zu genetischen Ressourcen in Zukunft komplizierter darstellen wird, andererseits erscheint eine größere Nutzbarmachung heimischer *In-situ*-Ressourcen denkbar. Auf diese Veränderungen muß sich ein nutzernahes und effizientes Dokumentations- und Informationswesen einstellen.

Vollkommen in Übereinstimmung mit dem oben geschilderten Mischmodell einer dezentral/zentralen Dokumentation und Information wurden Arbeiten im Rahmen eines vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Verbundvorhabens begonnen, neben den genetischen Ressourcen, die *ex situ* vorhanden sind auch die *In-situ*-Bestände in die Informationssysteme einzubeziehen, soweit dies sinnvoll und möglich ist. An dem Projekt zum Aufbau des Bundesinformationssystems für Genetische Ressourcen (BIG) beteiligen sich das Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben, das Institut für spezielle Botanik und Botanischer Garten der Ruhr-Universität Bochum im Namen des Verbands Botanischer Gärten in Deutschland, das Bundesamt für Naturschutz (BfN) sowie die ZADI. Das IGR der ZADI koordiniert das Projekt und stellt sicher, daß BIG auf den vorhandenen Konzeptionen und Informationssystemen aufbaut und diese ergänzt (<http://www.genres.de/big>) (siehe auch Beitrag BECKER). Der Mehrwert, der durch BIG erreicht werden soll, liegt neben der bestehenden PGRDEU insbesondere in der zusätzlichen Einbindung von Informationen wichtiger weiterer *Ex-situ*-Bestände der Botanischen Gärten sowie der Erschließung von *In-situ*-Ressourcen. Dies wird über eine nutzerfreundliche Abfrage der Taxonomie und der zu einzelnen Taxa bzw. Akzessionen gehörenden Eigenschaften sowie die Integration von internet-gestützten, interaktiven geografischen Informationssystemen (GIS) ermöglicht. Die neueren ICT-Entwicklungen lassen es auch zu, die zentralen und dezentralen Komponenten des gesamten Informationssystems zeitlich näher aneinanderzurücken. Dies ist vorgesehen über eine zentrale Abfragemaske (oder View) im Internet auf dezentrale *Online*-Datenbestände in den Partnerinstitutionen, die der Nutzer von seinem Bildschirm aus in einem vertretbaren Zeitrahmen durchführen kann. Er realisiert dabei nicht notwendigerweise, daß er zwischen verschiedenen Datenbankservern hin- und herwechselt. Im Zuge der neuesten ICT-Fortschritte wird die Lokalität der Server immer mehr von nachrangiger Bedeutung; an Bedeutung werden jedoch das Leistungsvermögen, die Übertragungsge-

schwindigkeit und die Datenqualität in Informationssystemen gewinnen. Vorausschauend wird BIG bereits jetzt in einer Form entwickelt, daß es sich künftig auch als Informationsplattform für die Verknüpfung von Daten zu Haus- und Nutztieren, Mikroorganismen, forstlichen sowie aquatischen Genressourcen verwenden läßt. BIG ist demzufolge eine Plattform, an die sich einzelne Informationsmodule anbinden lassen.

Ein erstes wichtiges Modul, das insbesondere für die Pflanzenzüchtung von Interesse ist, ist das Informationssystem zu Evaluierungsdaten (EVA) (<http://www.genres.de/eva>). Es wird als Verbundprojekt vom IPK, der Bundesanstalt für Züchtungsforschung an Kulturpflanzen (BAZ), der Martin-Luther-Universität in Halle und der ZADI durchgeführt; das IGR der ZADI koordiniert auch dieses Projekt und kann damit sicherstellen, daß die Entwicklungsschritte kompatibel mit den weiteren Bausteinen im gesamten Informationsverbund genetischer Ressourcen sind. Beispielhaft an den Fruchtarten Gerste, Kartoffel und Apfel werden akzessionsbezogene Daten zu Werteeigenschaften, d.h. Evaluierungen, über das Internet in nutzerfreundlicher Form zur Verfügung gestellt. Eine Ausdehnung auf andere Fruchtarten sowie auf weitere Evaluierungsergebnisse dieser Fruchtarten zur fortlaufenden Aktualisierung wird angestrebt.

Alle diese Entwicklungsschritte und Informationsbausteine müssen nicht zuletzt auch unter Berücksichtigung von Informationssystemen außerhalb Deutschlands sowie in enger Übereinstimmung mit den politisch-rechtlichen Rahmenbedingungen erfolgen. Von besonderer Bedeutung sind hier die Wahrung von Eigentumsrechten, den damit verbundenen Copyrights sowie die sich aus dem Übereinkommen über die Biologische Vielfalt ergebenden Anforderungen. Insofern erscheint auch eine enge Abstimmung mit dem vom Bundesumweltministerium (BMU) beim IGR der ZADI in Auftrag gegebenen Projekt zum Aufbau des sogenannten Clearing-House-Mechanismus, einem Informations- und Kommunikationssystem, ratsam (<http://www.biodiv-chm.de>).

Die weiteren Beiträge in diesem Band geben einen ausführlichen Überblick über ausgewählte dezentrale und zentrale Komponenten des deutschen Informationsverbundes im Bereich pflanzengenetischer Ressourcen. Dieser Informationsverbund wird erfolgreich sein, wenn auch weiterhin alle beteiligten Partner eng zusammenarbeiten. Die in den letzten Jahren in Deutschland erzielten Erfolge sind auch im internationalen Vergleich beachtlich.

Literatur

- ANON. (1993): Dokumentation Genetischer Ressourcen in der Bundesrepublik Deutschland. Beiträge zur ersten Arbeitstagung „Dokumentation Genetischer Ressourcen“ am Institut für Pflanzenbau der FAL, Braunschweig, 24.11.1992. BML-Schriftenreihe, Reihe A: Angewandte Wissenschaft, Heft 422, S. 77-162.
- BMU (1993): Gesetz zu dem Übereinkommen vom 5. Juni 1992 über die biologische Vielfalt vom 30. August 1993. BGBl, Teil II, 9.9.1993, Nr. 32, S. 1741ff.
- BOMMER, D.F.R. UND K. BEESE (1990): Pflanzengenetische Ressourcen – Ein Konzept zur Erhaltung und Nutzung für die Bundesrepublik Deutschland. BML-Schriftenreihe, Reihe A: Angewandte Wissenschaft, Heft 388, S. 5-190.
- SCHMIDT, F. (1993): Dokumentation Genetischer Ressourcen in der Bundesrepublik Deutschland – Ein Konzept für ein zentrales Informationssystem. BML-Schriftenreihe, Reihe A: Angewandte Wissenschaft, Heft 422, S. 165-188.

Dokumentation pflanzengenetischer Ressourcen in der Genbank des IPK Gatersleben

Documentation of plant genetic resources in the genebank of the IPK in Gatersleben

HELMUT KNÜPFER¹

Einführung

Die Genbank des IPK Gatersleben umfaßt am Standort Gatersleben ca. 85.000 Akzessionen von Kulturpflanzen und verwandter Wildpflanzen. Diese gehören zu mehr als 2.300 Arten, 670 Gattungen und 84 Familien (s. Tabellen 1-3). Dazu kommen noch die Sammlungen in den Außenstellen in Dresden-Pillnitz, Groß Lüsewitz und Malchow (Poel). Ihre Dokumentationssysteme, die aufgrund der historischen Entwicklung unabhängig von der IPK-Genbank erarbeitet wurden, wurden 1993 vorgestellt (FISCHER 1993, SCHÜLER 1993, WILLNER 1993). Die erstgenannten beiden Außenstellen konnten ihre Dokumentation im Rahmen des EVA-Projektes (HARRER 1999) vervollkommen. Die folgenden Ausführungen beziehen sich lediglich auf die Genbankdokumentation in Gatersleben.

Über die Aufgaben und Entwicklung der Genbank wurde an anderer Stelle ausführlich berichtet (z.B. HAMMER ET AL. 1994).

Tab. 1: Pflanzenfamilien, die mit mehr als 500 Akzessionen in der Genbank vertreten sind

Tab. 1: Plant families with more than 500 accessions in the genebank

Akzessionen	Familie	Akzessionen	Familie
39.980	Gramineae	1.661	Linaceae
2.0565	Leguminosae	1.561	Umbelliferae
6.238	Solanaceae	1.043	Chenopodiaceae
3.298	Cruciferae	885	Papaveraceae
2.817	Liliaceae	822	Labiatae
2.339	Compositae	792	Scrophulariaceae
2.073	Cucurbitaceae		

¹ Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK)
Abt. Genbank

Tab. 2: Gattungen, die mit mehr als 1.000 Akzessionen in der Genbank vertreten sind

Tab. 2: Genus with more than 1.000 accessions in the genebank

Akzes-sionen	Gattung	Familie	Akzes-sionen	Gattung	Familie
16.753	<i>Triticum</i>	Gramineae	2.789	<i>Allium</i>	Liliaceae
12.859	<i>Hordeum</i>	Gramineae	2.155	<i>Brassica</i>	Cruciferae
7.486	<i>Phaseolus</i>	Leguminosae	2.018	<i>Secale</i>	Gramineae
3.548	<i>Lycopersicon</i>	Solanaceae	1.662	<i>Linum</i>	Linaceae
3.207	<i>Vicia</i>	Leguminosae	1.381	<i>Capsicum</i>	Solanaceae
3.155	<i>Pisum</i>	Leguminosae	1.352	<i>Zea</i>	Gramineae
2.946	<i>Glycine</i>	Leguminosae	1.213	<i>Aegilops</i>	Gramineae
2.904	<i>Avena</i>	Gramineae	1.026	<i>Triticale</i>	Gramineae

Tab. 3: Arten, die mit mehr als 500 Akzessionen in der Genbank vertreten sind

Tab. 3: Species with more than 500 accessions in the genebank

Akzes-sionen	Art	Akzes-sionen	Art
12.109	<i>Triticum aestivum</i> L. em. Fiori et Paol.	1.337	<i>Zea mays</i> L.
11.577	<i>Hordeum vulgare</i> L.	1.047	<i>Capsicum annuum</i> L.
7.027	<i>Phaseolus vulgaris</i> L.	1.025	<i>Triticale</i>
3.010	<i>Pisum sativum</i> L. s.l.	969	<i>Brassica oleracea</i> L.
2.915	<i>Glycine max</i> (L.) Merr.	783	<i>Lactuca sativa</i> L.
2.758	<i>Lycopersicon esculentum</i> Mill.	717	<i>Beta vulgaris</i> L.
2.101	<i>Avena sativa</i> L.	663	<i>Papaver somniferum</i> L.
1.669	<i>Secale cereale</i> L.	617	<i>Vicia sativa</i> L.
1.564	<i>Triticum durum</i> Desf.	543	<i>Triticum aethiopicum</i> Jakubz.
1.556	<i>Linum usitatissimum</i> L.	537	<i>Antirrhinum majus</i> L.
1.530	<i>Vicia faba</i> L.		

In der Arbeitsgruppe Genbankdokumentation des IPK wird ein Informationssystem für pflanzengenetische Ressourcen (PGR) entwickelt, das einerseits der Unterstützung der Arbeitsabläufe in der Genbank (sog. Genbank-Management), und andererseits der Informationsbereitstellung für Nutzer des Materials, z.B. die Pflanzenzüchtung und die Kulturpflanzenforschung, dient. Neben der Dokumentation der eigenen Sammlungsbestände beteiligt sich die Genbank auch am Aufbau institutsübergreifender Datenangebote, sowohl im nationalen als auch im internationalen Rahmen. Darüber hinaus werden kulturpflanzen-taxonomische Datenbanken entwickelt. Über die Aufgaben des Computereinsatzes in Genbanken (KNÜPFER 1983) und die Entwicklung des Informationssystems der Genbank wurde bereits mehrfach berichtet (z.B. KNÜPFER 1993, KNÜPFER UND FREYTAG 1993).

Technische Basis

Als Datenbanksystem wird FoxPro verwendet. Die Teildatenbanken sind in einem lokalen Novell-Netz von allen angeschlossenen Arbeitsstationen aus zugänglich. Seit 1996 besteht Internet-Anschluß.

Paßportdaten

Kern des Informationssystems sind die Paßportdaten, d. h. Angaben zur Identität, Geschichte, geographischen Herkunft und botanischen Zuordnung des Materials. Vereinheitlichung, Korrektur, Ergänzung und Aktualisierung der Daten erfolgen laufend in Zusammenarbeit mit den Mitarbeitern der Genbank.

Das vorhandene Genbankmaterial ist im Internet recherchierbar (<http://fox-serv.ipk-gatersleben.de>); etwa 21% der Saatgutanforderungen des Jahres 1998 basierten auf WWW-Recherchen der Nutzer. Seit Anfang 1997 erfolgten monatlich ca. 430 Abfragen. Die mittlere Antwortzeit liegt bei 9 bis 10 Sekunden. Der WWW-Server der Genbank wurde 1996 und 1997 unter Windows 3.11 und Windows NT sowie dem Datenbanksystem Visual FoxPro 3.0 von den Leipziger Informatikstudenten U. Neubert und Th. Nowotka im Rahmen eines Praktikums am IPK entwickelt (KNÜPFER 1997, KNÜPFER ET AL. 1998a). Wohl einmalig im Bereich von Internet-Datenbanken pflanzengenetischer Ressourcen ist die Möglichkeit, die aktuellen Abfrageergebnisse als DBF-Tabelle herunterzuladen.

Charakterisierungs- und Evaluierungsdaten

Die beim Erhaltungsanbau gewonnenen Charakterisierungs- und Evaluierungsdaten (Bonituren) werden seit etwa 1992 laufend in Dateien erfaßt. In größeren Sortimentgruppen werden hierfür auch mobile Datenerfassungsgeräte eingesetzt. Die rückwirkende Erfassung historischer Boniturdaten, die beim jährlichen Vermehrungsanbau seit 1946 erhoben wurden, wird durch Drittmittel-

projekte unterstützt (s. unten). Daten der sekundären Evaluierung, die hauptsächlich aus Forschungsprojekten von Kooperationspartnern stammen (z.B. WALTHER et al. 1999), werden seit 1995 erfaßt, bisher insbesondere Resistenz- und Qualitätsmerkmale von 38 Fruchtarten (vorwiegend Getreide, Hülsenfrüchte, Gemüse, sowie Arznei- und Gewürzpflanzen; HAMMER ET AL. 1998).

Managementdaten

Zu den Managementdaten gehören alle Informationen, die für interne Arbeitsabläufe in der Genbank benötigt werden (KNÜPFER 1983). Die Saatgutbestände im Kühllager werden in einem eigenen System dokumentiert, und Keimfähigkeitsergebnisse statistisch ausgewertet (z.B. SPECHT ET AL. 1997, 1998). Die Anbauvorbereitung erfolgt computergestützt, und Aussaatlisten und Feldbücher werden gedruckt (FREYTAG ET AL. 1995). Für das *in vitro* gehaltene Material wurde ein eigenes Managementsystem programmiert, mit dessen Hilfe u.a. die laufenden Arbeiten geplant werden können. Eine Bilddatenbank für Fotos sowie Ergebnisse der molekularen Charakterisierung von Genbank-Akzessionen ist im Aufbau. Daten der Wetterstation, darunter seit 1999 auch bodennahe Ozonwerte, werden vollautomatisch in feinem Zeitraster erfaßt und zentral verfügbar gehalten.

Drittmittelprojekte

EVA. Im Rahmen des BML-Verbundprojektes EVA (s. Beitrag HARRER in diesem Band) werden die am IPK auf Karteikarten und in Feldbüchern vorhandenen langjährigen Evaluierungsdaten bei Gerste erfaßt und für ein zentrales Informationsangebot aufbereitet. Dabei werden bis zu 33 Merkmale je Akzession und Anbaujahr aufgenommen. Durch ABM-Kräfte wurden darüber hinaus in den vergangenen Jahren auch rückwirkend Daten von *Phaseolus*-Bohnen, Weizen, Roggen, Erbsen und Lein aufgenommen.

EU-Projekt Gerste. Ein vom Autor koordiniertes EU-Projekt mit 28 Partnern (darunter acht Genbanken, elf Forschungsinstitute, neun Züchtungsfirmen und eine Nichtregierungsorganisation) mit dreijähriger Laufzeit wurde kürzlich bewilligt. Es beinhaltet die gemeinsame und koordinierte Evaluierung von Genbank-Akzessionen (besonders Landsorten und Gersten-Wildarten), einschließlich der Internationalen Barley Core Collection (BCC; s. unten), auf eine Reihe von Pilz- und Viruskrankheiten sowie Blattlaustoleranz und abiotische Stressfaktoren. Die Evaluierungsmethoden werden zwischen den Partnern abgestimmt. Die beteiligten Genbanken werden Duplikate in ihren Beständen ermitteln und Material für die gemeinsame Evaluierung vermehren und bereitstellen. Die besten Akzessionen werden der Züchtung zur Einbeziehung in Kreuzungsprogramme übergeben. IPK wird gemeinsam mit ZADI eine Projekt-Homepage erstellen. Die beim Projektkoordinator IPK betriebene Europäische Gerstendatenbank des ECP/GR (s. unten) wird im Projekt aktualisiert und um die Evaluierungsergebnisse sowie vorhandene Evaluierungsdaten ergänzt, die über das Internet recherchierbar gemacht werden sollen.

BIG. In Zusammenarbeit mit drei anderen Einrichtungen in der Bundesrepublik Deutschland wird das BMBF-Verbundprojekt „Bundesinformationssystem Genetische Ressourcen“ (BIG) bearbeitet (s. Beitrag BECKER in diesem Band). Datenbestände des IPK (Abteilungen Genbank und Taxonomie), der botanischen Gärten Deutschlands und des Naturschutzes sollen dabei über eine gemeinsame Schnittstelle im Internet verfügbar gemacht werden. In diesem Rahmen wurde mit dem Aufbau einer kulturpflanzentaxonomischen Datenbank zu „Mansfeld’s World Catalogue of Cultivated Plants“ begonnen (OCHSMANN et al. 1999).

Weitere Datenbanken

EBDB. Die „*European Barley Database*“ des ECP/GR enthält gegenwärtig 89.000 Nachweise (Prototyp: <http://www.dainet.de/eccdb/barley>) (KNÜPFER UND LÓPEZ 1999). Damit dokumentiert sie die Bestände der Gerstenkollektionen von 36 Genbanken in 25 Ländern. Gemeinsam mit der Außenstelle Malchow wurde die Europäische *Poa*-Datenbank aufgebaut (<http://www.dainet.de/eccdb/poa/>) (KNÜPFER ET AL. 1998b). Ziel ist bei diesen Datenbanken u. a. eine internationale Arbeitsteilung durch Ermittlung von Duplikaten und Unikaten (VAN HINTUM UND KNÜPFER 1995, KNÜPFER ET AL. 1997). Paßportdaten der Genbank werden regelmäßig in einem abgestimmten Format („Multi-Crop Passport Descriptors“, HAZEKAMP ET AL. 1997) an verschiedene internationale PGR-Datenbanken, sogenannte „*Central Crop Databases*“ (KNÜPFER 1995) abgegeben.

Kulturpflanzentaxonomische DB. Zur Unterstützung von Explorations- und Sammlungsaktivitäten in einzelnen Ländern wird eine kulturpflanzentaxonomische Datenbank betrieben (KNÜPFER UND HAMMER, im Druck). Sie enthält Informationen über die kultivierten Pflanzenarten (u.a. von Kuba, Italien und Korea) und ermöglicht die Zusammenstellung von länderspezifischen Checklisten (z.B. Nord- und Mittelitalien, HAMMER ET AL. 1999), die Publikation von speziellen Listen (z.B. der kultivierten Heilpflanzen Koreas, HAMMER ET AL. 1997) sowie den Vergleich von Kulturpflanzenfloren. Diese Datenbank ergänzt das Vorhaben der Datenbank für das Mansfeld-Verzeichnis (OCHSMANN ET AL. 1999) durch länderspezifische Detailinformationen.

Barley Core Collection

Die Genbank des IPK ist maßgeblich am Aufbau der Internationalen *Barley Core Collection* (BCC) beteiligt, einer separaten Sammlung begrenzten Umfangs, jedoch mit möglichst großer Diversität, die aus repräsentativen Mustern vorhandener Sammlungen gebildet wird und durch ihre umfassende Charakterisierung und Dokumentation den Zugang zu diesen erleichtern soll (KNÜPFER UND HINTUM 1995). Das IPK ist eines von weltweit vier „*active BCC centres*“, die für die Abgabe von Saatgut an Interessenten zuständig sind. Die BCC ist Hauptuntersuchungsgegenstand des oben genannten EU-Projektes. Eine Internet-Datenbank für die BCC soll am IPK entwickelt werden.

Literatur

- BECKER, B. (1999): Bundesinformationssystem Genetische Ressourcen – Konzept und erste Ergebnisse. In: F. BEGEMANN, S. HARRER UND J.D. KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. Schriften zu Genetischen Ressourcen, Bd 12, ZADI, Bonn.
- FISCHER, M. (1993): Stand der Dokumentation genetischer Ressourcen in der Genbank-Außenstelle Dresden-Pillnitz. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssituation. Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Angewandte Wissenschaft H. 422, 129-130.
- FREYTAG, U., K. HAMMER, H. KNÜPFER UND H. LUX (1995): Einsatz der EDV für das interne Genbankmanagement, Interdependenzen der Teilprojekte und deren Nutzung. In: F. BEGEMANN UND R. VÖGEL (Hrsg.): *In-situ*-Erhaltung pflanzengenetischer Ressourcen in der Bundesrepublik Deutschland am natürlichen Standort und *on farm*. Schriften zu Genetischen Ressourcen, Bd. 2, S. 216-217, ZADI, Bonn.
- HAMMER, K., U. FREYTAG, U. WALTHER UND H. KNÜPFER (1998): Sekundäre Evaluierung von Genbankmaterial - Ergebnisse und Informationsbereitstellung. In: F. BEGEMANN (Hrsg.): Züchterische Nutzung pflanzengenetischer Ressourcen - Ergebnisse und Forschungsbedarf. Schriften zu Genetischen Ressourcen, Bd. 8, S. 228-234, ZADI, Bonn.
- HAMMER, K., H. GÄDE UND H. KNÜPFER (1994): 50 Jahre Genbank Gatersleben - eine Übersicht. Vortr. Pflanzenzüchtg. 27, 333-383.
- HAMMER, K., H. KNÜPFER UND H.-D. HOANG (1997): Koreanische Heilpflanzen - eine Liste der kultivierten Arten. Drogenreport 10 H. 16, 57-59, i-xxvi.
- HAMMER, K., H. KNÜPFER, G. LAGHETTI AND P. PERRINO (1999): Seeds from the past. A catalogue of crop germplasm in Central and North Italy. Istituto del Germoplasma, Bari, Italy. iv + 258 pp.

- HARRER, S. (1999): EVA – Aufbau eines Informationssystems für Evaluierungsdaten pflanzengenetischer Ressourcen in der Bundesrepublik Deutschland. In: F. BEGEMANN, S. HARRER UND J.D. KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. Schriften zu Genetischen Ressourcen, Bd 12, ZADI, Bonn.
- HAZEKAMP, T., J. SERWINSKI AND A. ALERCIA (1997): Multicrop Passport Descriptors (final version). In: E. LIPMAN, M. W. M. JONGEN, TH. J. L. VAN HINTUM, T. GASS AND L. MAGGIONI (compilers): Central Crop Databases: Tools for Plant Genetic Resources Management, pp. 75-78. International Plant Genetic Resources Institute, Rome, Italy; CGN, Wageningen, Netherlands.
- HINTUM VAN, TH.J.L. AND H. KNÜPFER (1995): Duplication within and between germplasm collections. I. Tracing duplication on the basis of passport data. Genet. Resour. Crop Evol. 42, 127-133.
- KNÜPFER, H. (1983): Computer in Genbanken - eine Übersicht. Kulturpflanze 31, 77 143.
- KNÜPFER, H. (1993): Dokumentation pflanzengenetischer Ressourcen. In: S.C. JUTZI UND B. BECKER (Hrsg.): Pflanzengenetische Ressourcen. Erhaltung und multiple, nachhaltige Nutzung. Beiträge zur 21. Witzenhäuser Hochschulwoche, 2.-4. Juni 1993. Witzenhausen. Tropenlandwirt, Beih. 49, S. 79-88.
- KNÜPFER, H. (1995): Central crop databases. In: TH.J.L. VAN HINTUM, M.W.M. JONGEN AND T. HAZEKAMP (eds.): Standardization in plant genetic resources documentation, pp. 51-62. Centre for Genetic Resources The Netherlands, Wageningen.
- KNÜPFER, H. (1997): Options and approaches to providing on-line access to databases. Part III. FoxPro on-line databases on an institution server. Report, European Central Crop Databases (ECCDB) On-line Databases Training Workshop, 8-10 June 1997, Bonn, Germany. ECP/GR & ZADI. (<http://www.cgiar.org/ecpgr/platform/ECCDB1/Fox-serv.htm>).
- KNÜPFER, H., L. FRESE AND M.W.M. JONGEN (1997): Using central crop databases: searching for duplicates and gaps. In: E. LIPMAN, M. W. M. JONGEN, TH. J. L. VAN HINTUM, T. GASS AND L. MAGGIONI (compilers): Central Crop Databases, Tools for Plant Genetic Resources Management. IPGRI, Rome, Italy; CGN, Wageningen, The Netherlands, pp. 59-68.
- KNÜPFER, H. UND U. FREYTAG (1993): Dokumentation pflanzengenetischer Ressourcen in der Genbank des Instituts für Pflanzengenetik und Kulturpflanzenforschung Gatersleben. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationsituation. Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Angewandte Wissenschaft H. 422, 113-123.
- KNÜPFER, H. AND K. HAMMER (in press): Agricultural biodiversity: a database for checklists of cultivated plant species. In: Proc. Third International Symposium on the Taxonomy of Cultivated Plants, Edinburgh.
- KNÜPFER, H., S. HARRER, J.D. JIMÉNEZ KRAUSE, L. LÓPEZ, U. NEUBERT, T. NOWOTKA UND E. WILLNER (1998): Informationsangebote der Genbank im Internet. In: F. BEGEMANN (Hrsg.): Züchterische Nutzung pflanzengenetischer Ressourcen - Ergebnisse und Forschungsbedarf. Schriften zu Genetischen Ressourcen, Bd. 8, S. 239-243. ZADI, Bonn.
- KNÜPFER, H., S. HARRER AND E. WILLNER (1998): The European *Poa* Database. In: L. MAGGIONI, P. MARUM, R. SACKVILLE HAMILTON, I. THOMAS, T. GASS AND E. LIPMAN

- (compilers): Report of a Working Group on Forages. Sixth Meeting, 6-8 March 1997, Beitostølen, Norway, pp. 33-36. International Plant Genetic Resources Institute, Rome, Italy.
- KNÜPFFER, H. AND TH.J.L. VAN HINTUM (1995): The Barley Core Collection - an international effort. In: T. HODGKIN, A.H.D. BROWN, TH.J.L. VAN HINTUM AND E.A.V. MORALES (eds.): Core Collections of Plant Genetic Resources. Wiley, Chichester, 171-178.
- KNÜPFFER, H. AND L. LÓPEZ (1999): Status Report on the European Barley Database (EBDB). In: L. MAGGIONI, H. KNÜPFFER, R. VON BOTHMER, M. AMBROSE, K. HAMMER AND E. LIPMAN (compilers): Report of a Working Group on Barley. Fifth meeting, 10-12 July 1997, Alterode/Gatersleben, Germany. International Plant Genetic Resources Institute, Rome (in press).
- OCHSMANN, J., N. BIERMANN, H. KNÜPFFER UND K. BACHMANN (1999): Aufbau einer WWW-Datenbank zu "Mansfeld's World Manual of Agricultural and Horticultural Crops" (Mansfeld-Verzeichnis, 3. Aufl.). In: F. BEGEMANN, S. HARRER UND J.D. KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen in Deutschland. Schriften zu Genetischen Ressourcen, Bd 12, ZADI, Bonn..
- SCHÜLER, K. (1993): Stand der Dokumentation genetischer Ressourcen in der Genbank - Außenstelle Groß Lüsewitz. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssituation. Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Angewandte Wissenschaft H. 422, 126-128.
- SPECHT, C.-E., U. FREYTAG, K. HAMMER AND A. BÖRNER (1998): Survey of seed germinability after long-term storage in the Gatersleben genebank (part 2). Plant Genet. Resour. Newsl. 115, 39-43.
- SPECHT, C.-E., E.R.J. KELLER, U. FREYTAG, K. HAMMER AND A. BÖRNER (1997): Survey of seed germinability after long-term storage in the Gatersleben genebank. Plant Genet. Resour. Newsl. 111, 64-68.
- WALTHER, U., A. HABEKUß, D. KOPAHNKE, G. PROESELER AND E. SCHLIEPHAKE (1999): Evaluation of IPK's barley germplasm for disease resistance. In: L. MAGGIONI, H. KNÜPFFER, R. VON BOTHMER, M. AMBROSE, K. HAMMER AND E. LIPMAN (compilers): Report of a Working Group on Barley. Fifth meeting, 10-12 July 1997, Alterode/Gatersleben, Germany. International Plant Genetic Resources Institute, Rome (in press).
- WILLNER, E. (1993): Stand der Dokumentation genetischer Ressourcen in der Genbank-Außenstelle Malchow/Poel. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssituation. Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Angewandte Wissenschaft H. 422, 124-125.

Das Informationsangebot der Genbank der BAZ – *Online-Demonstration*

The information service of the BAZ genebank - online presentation

STEFAN BÜCKEN¹ UND LOTHAR FRESE¹

Die heutige Genbank der Bundesanstalt für Züchtungsforschung an Kulturpflanzen (BAZ) war zur Zeit ihrer Gründung im Jahre 1970 als eine Institution zur Unterstützung der Pflanzenzüchtung und Pflanzenzüchtungsforschung bei der systematischen Suche nach genetischen Ressourcen für ausgewählte Züchtungs- oder Forschungsprogramme geplant (Abb. 1). Heute würde man eine solche Einrichtung als „*Crop Introduction Centre*“ bezeichnen. Neben der aktiven Sammlung von Wild- und Kulturformen und der Vermittlung von Material anderer Einrichtungen an die praktische Züchtung und die Züchtungsforschung, erhielt die Braunschweiger Genbank den Auftrag, in der Bundesrepublik Deutschland dezentral entstandene Sammlungen pflanzengenetischer Ressourcen von Forschungseinrichtungen zentral zu lagern und zu erhalten (BOMMER 1972). Hierdurch entwickelte sich eine große Getreidesammlung, die insbesondere in den achtziger Jahren durch Sammlungen potentieller Industriepflanzenarten sowie Körnerleguminosen und Gräsern ausgebaut wurde. Die Schwerpunkte der Sammlung liegen zur Zeit bei den Getreidearten, Industriepflanzen, Leguminosen sowie verschiedenen Arten der Gattungen *Brassica* und *Sinapis*. Weiterhin werden die deutsch-niederländische *Beta*- und Zichoriensammlung sowie Teile der niederländisch-deutschen Kartoffelsammlung betreut. Besondere Aktivitäten werden bei den Leitfruchtarten *Beta* und *Avena* entwickelt.

Im internationalen Kontext erhielt die Braunschweiger Genbank das Mandat zum Aufbau und zur Pflege der Internationalen Datenbank für *Beta* (IDBB) (IPGRI 1987) und der Europäischen *Avena* Datenbank (EADB) (UNDP/IBPGR 1984). Beide Datenbanken dienen dazu, die weltweit bzw. europaweit in Genbanksammlungen verfügbaren Muster der Gattung *Beta* und *Avena* zu dokumentieren und auf diese Weise einen vereinfachten zentralen Informationszugriff zu gestatten.

In den siebziger und achtziger Jahren wurde an der Genbank ein EDV gestütztes Informationssystem für die eingelagerten pflanzengenetischen Ressourcen aufgebaut (BOMMER 1972, SEIDEWITZ ET AL. 1993). Insbesondere die hierzu entwickelte Dokumentationsstrategie kann als richtungsweisend für die Dokumentationsarbeit an einer zentralen Informations- und Lageeinrichtung für pflanzengenetische Ressourcen mit beschränkten Charakterisierungs- und Evaluierungskapazitäten gelten. Neben sogenannten Passportdaten, welche die eingelagerten Muster zusammen mit den erfaßten Charakterisierungsdaten eindeutig bestimmen, wurde ein besonderer Schwerpunkt auf die Dokumentation dezentral erhobener Evaluierungsdaten gelegt.

¹ Bundesanstalt für Züchtungsforschung an Kulturpflanzen (BAZ)
Genbank

Bei den Daten der BAZ-Genbank handelt es sich somit um einen Datenpool. Dies bedeutet, daß zu den eingelagerten Mustern sowohl Daten aus eigenen Erhebungen als auch Daten, die von anderen in- und ausländischen Institutionen erhoben wurden, dokumentiert wurden. Da die Richtigkeit und die Qualität von Fremddaten nur bedingt überprüfbar sind, wurde in das Dokumentationssystem jedes Datum ohne Wertung aufgenommen. Die Daten werden vor diesem Hintergrund als Hinweis auf eine Eigenschaft verstanden und nicht als wissenschaftlicher Nachweis für diese Eigenschaft.

Im Jahre 1991 wurde das Informationszentrum für Genetische Ressourcen der Zentralstelle für Agrardokumentation und -information (ZADI) eingerichtet, mit der Aufgabe eine bundesweite Informationsinfrastruktur zu genetischen Ressourcen aufzubauen (SCHMIDT 1993). Im Rahmen der sich hierdurch ergebenden Arbeitsteilung im Ressortbereich des BML konzentriert sich die Genbank der Bundesanstalt für Züchtungsforschung an Kulturpflanzen auf die Aufgaben eines Management- und Servicezentrums für pflanzen genetische Ressourcen, sowie auf die Etablierung eines an den Leitfruchtarten *Beta* und *Avena* orientierten fruchtartspezifischen Informationsangebotes.

Das Informationsangebot der BAZ-Genbank nutzt mehrere Angebotskanäle. An pflanzen genetischen Ressourcen Interessierte können zum einen jederzeit den direkten Kontakt mit der Genbank suchen und Daten zu den eingelagerten Mustern in ausgedruckter Form oder auf Diskette anfordern. Zum anderen entwickelt sich das Angebot von Informationen im WWW (World Wide Web) zu einem wichtigen Informationskanal. Dieses Informationsmedium ermöglicht es, allgemeine Informationen über die Genbank auf der *Homepage* zu vermitteln, während musterbezogene Informationen in verschiedenen *online* und *offline* recherchierbaren Datenbanken angeboten werden.

Eine der tragenden Säulen des WWW-Angebotes der BAZ-Genbank ist die durch ZADI aufgebaute Übersichtsdatenbank zu pflanzen genetischen Ressourcen (PGRDEU), welche im Internet *online* recherchierbar ist. Mittels dieser Datenbank ist es nun möglich, zentral die Passportdaten zu den meisten in Deutschland als Genbankmuster definierten pflanzen genetischen Ressourcen abzurufen (Abb. 1).

Eine weitere Säule des WWW-Angebotes stellen die auf der *Homepage* der BAZ-Genbank angebotenen „*downloadable*“ Fruchtartdatenbanken dar. Diese Datenbanken können auf den eigenen PC kopiert oder auf Diskette angefordert werden. Sie erlauben es, die Passportdaten zu den Fruchtartbeständen der BAZ Genbank auf dem eigenen PC zu recherchieren. Auf diese Weise kann auch Interessierten ohne eigenen Internet-Anschluß ein relativ einfacher Informationszugang gewährt werden. Erforderlich ist in diesem Fall jedoch weiterhin ein PC und eine Software, die in der Lage ist das dBASE - Dateiformat zu lesen (Abb. 2).

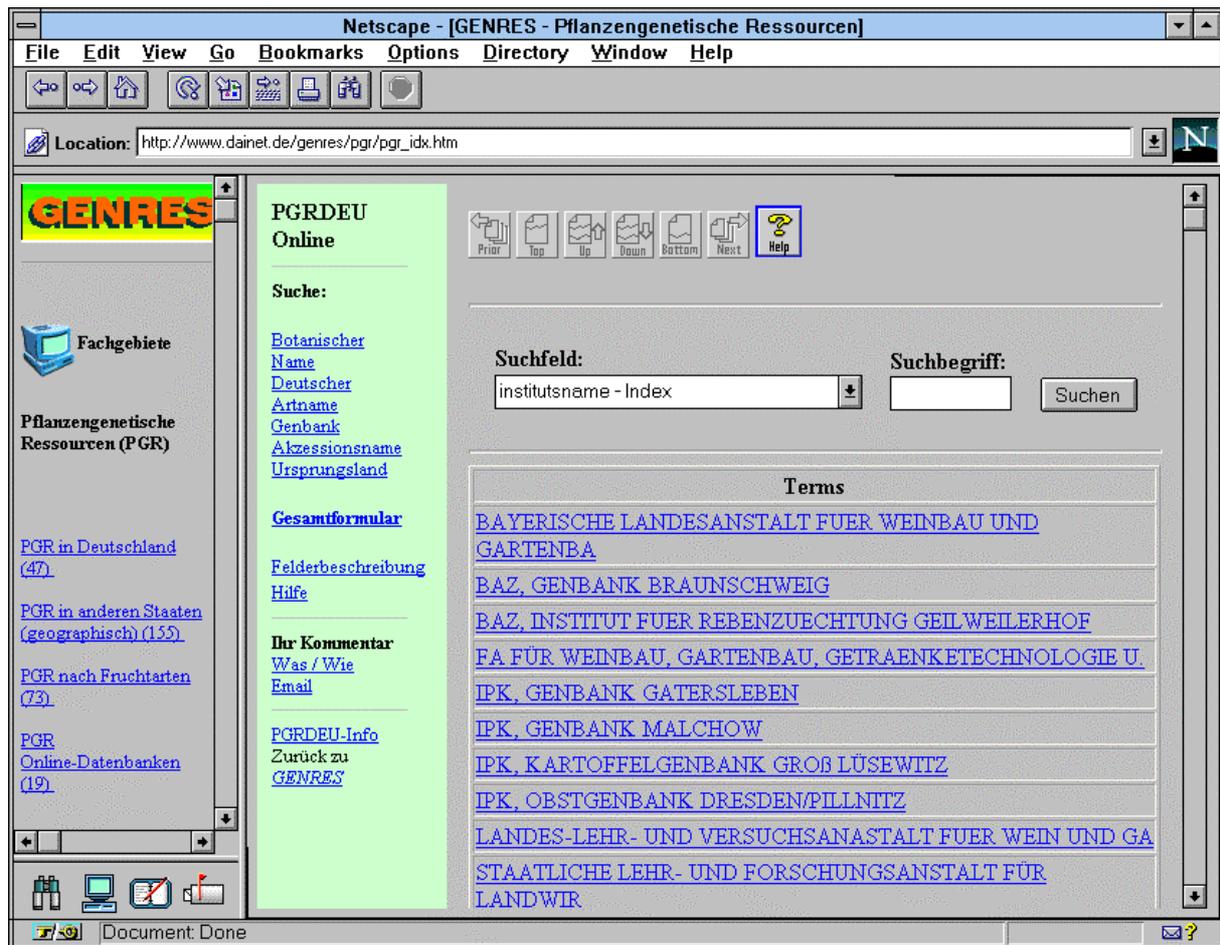


Abb. 1: WWW-Seite der Datenbank zu pflanzengenetischen Ressourcen in Deutschland (PGRDEU) mit Quellenangaben zu den angebotenen Datenbeständen

Fig. 1: Internet page of PGRDEU, the database for plant genetic resources in Germany

Die dritte Säule des Informationsangebotes der BAZ Genbank ist die direkte Beratung und Vermittlung von Evaluierungsdaten aus dem Datenpool. Ergänzend hierzu wird im Rahmen eines Pilotprojektes (EVA) zusammen mit ZADI und anderen Kooperationspartnern an der Repräsentation von Evaluierungsdaten im Internet gearbeitet. Besonders problematisch erscheint in diesem Zusammenhang eine Qualitätsbewertung der Datenbestände im Internet. Diese läßt eine direkte Beratung der Nutzer dieses Informationsangebotes durch die BAZ-Genbank auch weiterhin wünschenswert erscheinen.

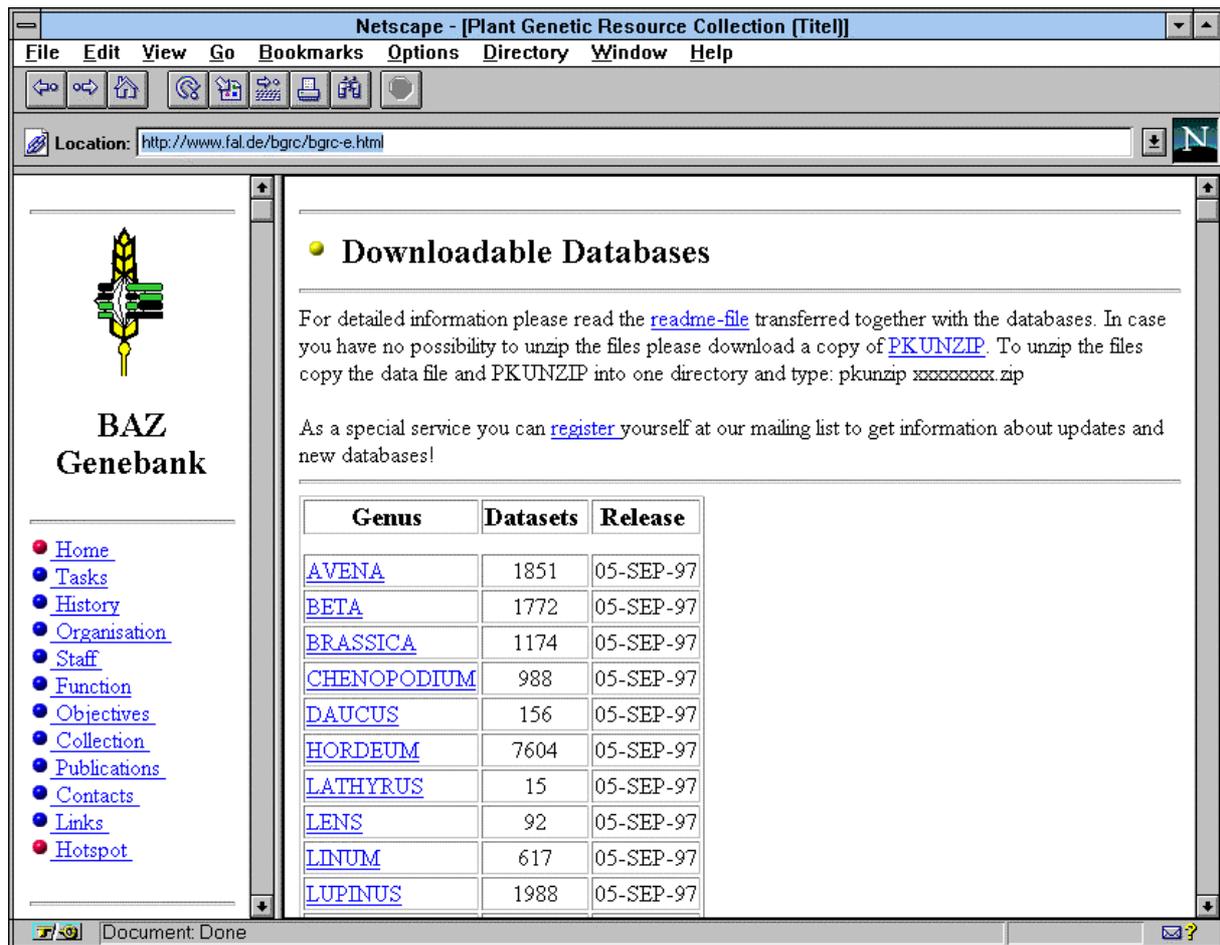


Abb. 2: WWW-Seite der „downloadable“ Fruchtartdatenbanken der BAZ-Genbank

Fig. 2: Internet page of downloadable crop specific databases of the BAZ genebank

Neben den nationalen Informationsangeboten ist die BAZ-Genbank durch das Mandat zur Pflege der Internationalen Datenbank für *Beta* (IDBB) und der Europäischen *Avena*-Datenbank (EADB) in das europäische Informationsangebot durch das Europäische Kooperationsprogramm zu pflanzengenetischen Ressourcen (ECP/GR *European Cooperative Programme for Crop Genetic Resources Networks*) eingebunden. Hervorzuheben ist hierbei der technisch dezentrale Ansatz, der mittels des Internets realisiert werden kann. Während die *on-line* recherchierbaren Datenbestände von IDBB und EADB technisch auf einem leistungsstarken Datenbankserver der ZADI angeboten werden, können die *offline*-Versionen der Datenbanken direkt von der *Homepage* der BAZ-Genbank auf den eigenen PC geladen werden. Obwohl das Informationsangebot demnach geografisch weit verstreut gepflegt wird, erlaubt das WWW einen zentralen Einstieg in das europäische Informationssystem zu pflanzengenetischen Ressourcen über die ECP/GR *Homepage*, ohne daß der Internet-Benutzer etwas von der verteilten Informationsstruktur bemerkt (Abb. 3).

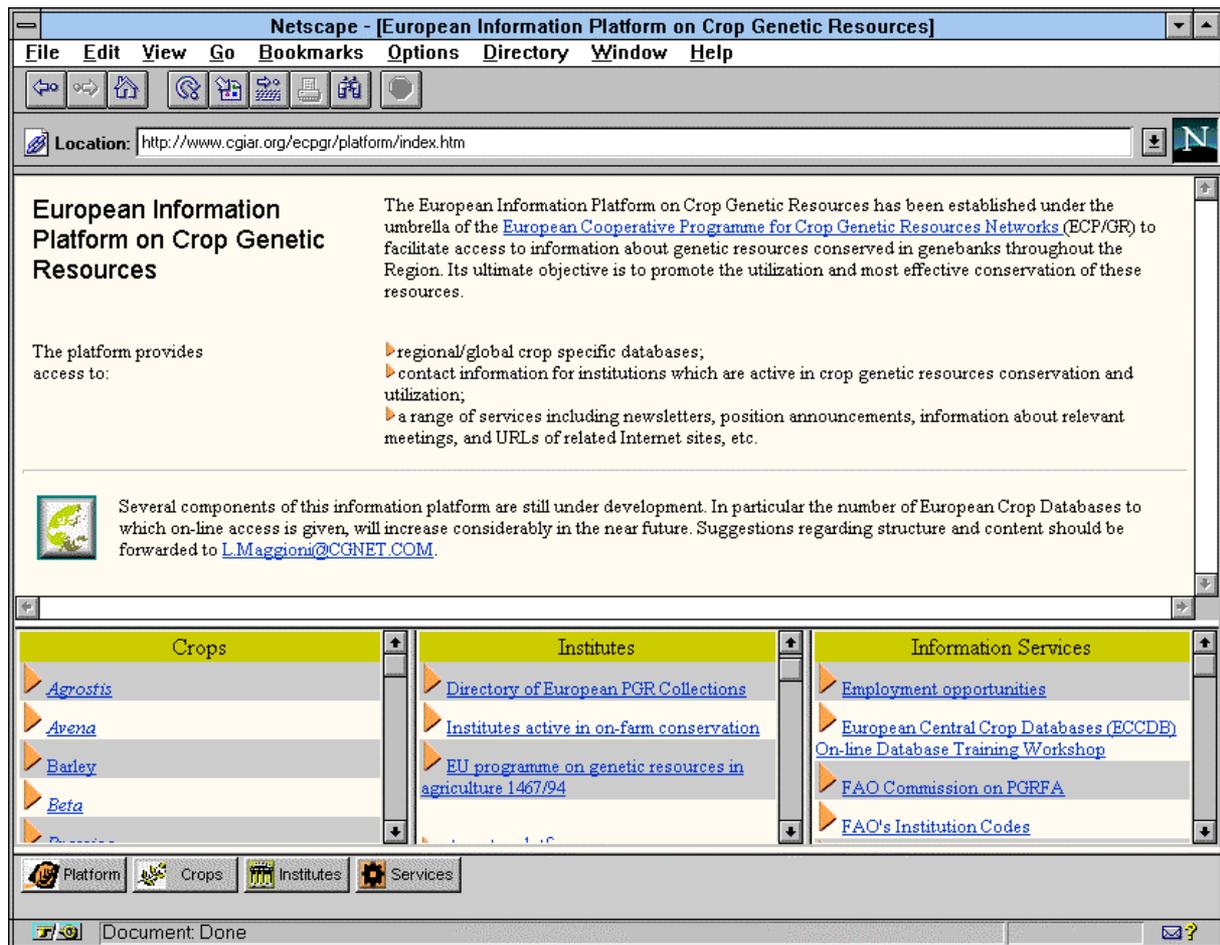


Abb. 3: Homepage der Europäischen Informationsplattform, über die zentral z.B. auf die räumlich dezentral gehaltenen und gepflegten Datenbestände der Europäischen Avena Datenbank zugegriffen werden kann

Fig. 3: Homepage of the ECP/GR Information platform with central view on decentralized data

WWW-Links

<http://www.dainet.de/genres/genres.htm>

<http://www.dainet.de/genres/pgrdeu/>

<http://www.dainet.de/genres/infos/igrinfo.htm>

<http://www.bafz.de>

<http://www.fal.de/bgrc/bazgb-d/dbank-g.html>

<http://www.fal.de/bgrc/bazgb-d/idbb-g.html>

<http://www.fal.de/bgrc/bgrc-g.html>

<http://www.fal.de/bgrc/bazgb-d/eadb-g.html>

<http://www.cgiar.org/ecpgr/index.html>

Literatur

- BOMMER, D. (1972): Entwicklung und Bedeutung von Genbanken in der Welt und in der Bundesrepublik Deutschland. SAFA, 24, Nr. 15 und 16.
- IPGRI (1987): Report on a *Beta* Workshop. European Cooperative Programme for the Conservation and Exchange of Crop Genetic Resources. International Board for Plant Genetic Resources, Rome.
- SCHMIDT, F. (1993): Dokumentation Genetischer Ressourcen in der Bundesrepublik Deutschland - Ein Konzept für ein zentrales Informationssystem. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssysteme. Schriftenreihe des BML, Angewandte Wissenschaft, Heft 422, S. 165- 188.
- SEIDEWITZ, L., L. FRESE UND M. DAMBROTH (1993): Datendokumentation und computergestützte Informationsbereitstellung der Sammlung Pflanzengenetischer Ressourcen am Institut für Pflanzenbau. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssysteme. Schriftenreihe des BML, Angewandte Wissenschaft, Heft 422, S. 131- 143.
- UNDP/IBPGR (1984): Report of a Working Group on Oat held at the Aegean Regional Agricultural Research Institute, Menemen, Izmir, Turkey, 25-27 September 1984. UNDB/IBPGR European Cooperative Programme for the Conservation and Exchange of Crop Genetic Resources, International Board for Plant Genetic Resources, Rome.

PGRDEU - Der zentrale online-Katalog deutscher *Ex-situ*-Sammlungen pflanzengenetischer Ressourcen

PGRDEU – a central online catalogue of German germplasm collections

JULIÁN DANIEL JIMÉNEZ KRAUSE¹

Zusammenfassung

Die zentrale Dokumentation pflanzengenetischer Ressourcen (PGR) in Deutschland ist eine Aufgabe des Informationszentrums für Genetische Ressourcen (IGR) der Zentralstelle für Agrardokumentation und -information (ZADI) in Bonn. Seit der Gründung des IGR 1991 sind dazu mehrere Konzepte erarbeitet worden – von der Einrichtung eines zentralen PGR-Rechenzentrums bis zum heutigen Konzept der dezentralen Dokumentation mit einem zentralen Informationsdienst, bestehend u.a. aus dem online-Katalog PGRDEU deutscher *Ex-situ*- Sammlungen pflanzengenetischer Ressourcen, der hier näher beschrieben wird.

Abstract

The central documentation of German germplasm has been a concern of the Information Centre for Genetic Resources (IGR) of the Centre for Documentation and Information in Agriculture (ZADI) since its foundation in 1991. Different concepts for a central germplasm documentation in Germany have been developed – from a totally centralized approach to today's system of decentralized documentation with a central platform consisting, among other components, of an online germplasm catalogue, PGRDEU, which is described here.

Einführung

In BEGEMANN (1999, in diesem Band) wird die Entwicklung von Konzepten für eine bundesdeutsche zentrale Dokumentation pflanzengenetischer Ressourcen (PGR) beschrieben. Das jetzige Schema der dezentralen Dokumentation mit einem zentralen „Informationsdienst“ resultierte dialektisch aus den vorherigen Konzepten einer vollkommen dezentralisierten und einer sehr stark zentralisierten PGR-Dokumentation in Deutschland (SCHMIDT 1993). Es ist aber auch das Resultat der rasanten Entwicklung der Informationstechnologie im Internet. Besonders das World Wide Web (WWW) mit seiner graphischen und einfach zu bedienenden Benutzeroberfläche bot sich als Kommunikations- und Informationsmedium geradezu an. Viele deutsche PGR-Einrichtungen haben sich in den letzten Jahren Zugang zu und Präsenz im WWW geschaffen.

¹ Zentralstelle für Agrardokumentation und -information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

Der dezentrale Aspekt der PGR-Dokumentation spiegelt sich in der Tatsache wider, daß der Nutzer über das WWW direkt, einfach und zunehmend billiger an die Datenquellen gelangen kann, um dort relevante, aktuelle Informationen „aus erster Hand“ zu erhalten.

Aber im informationsüberladenen WWW wird es immer schwieriger, einzelne isolierte Datenquellen aufzufinden. Sie gehen unter in der Datenflut des Cyberspace. Daher nimmt die Bedeutung von thematischen, zentralen WWW-Informationssystemen zu, die als Orientierungs- und Vermittlungsstellen funktionieren und dem Nutzer beim Auffinden relevanter Informationen und Datenquellen behilflich sind. Genau an dieser Stelle entwickelt das IGR der ZADI einen Informationsdienst, der als bundesdeutsche Informationsdrehscheibe im Bereich der PGR fungiert. PGR-Einrichtungen, die noch keine eigene Darstellung im WWW haben, sind im Verweissystem des zentralen Informationsdienstes im WWW mit ihrer postalischen Adresse, Tel.- und Faxnummern u.s.w. präsent.

Der zentrale online-PGR-Informationsdienst besteht u.a. aus einem Metadaten-Informationssystem (GENRES) (siehe Beitrag MÜNCH) und aus der Faktendatenbank PGRDEU als online PGR Katalog deutscher *Ex-situ*-Sammlungen.

PGRDEU heute und morgen

In PGRDEU werden die wichtigsten Informationen zu den Mustern der *Ex-situ*-Sammlungen in Deutschland zusammengefaßt und als online-Datenbank im WWW angeboten.

Das IGR der ZADI entwickelt und pflegt den online-Katalog PGRDEU. Das bedeutet u.a. Entwurf und Realisierung der Datenbankstruktur, das Konvertieren und Kompilieren von Informationen aus verschiedenen Quellen (zuzüglich Teilüberprüfung auf Dateninkonsistenzen) und das Erarbeiten eines Aktualisierungsmechanismus.

Inhalt

PGRDEU enthält Informationen auf zwei Ebenen – Auf Taxonebene werden die Felder Familie, Gattung, Art, infraspezifische Angaben, deutscher Name und Fruchtartgruppe nach IPGRI-Kategorien erfaßt. Auf Akzessionsebene (z. B. Genbankmaterial) werden Daten zu den Merkmalen Akzessionsname, Akzessionstyp, erhaltende Genbank, Akzessionsnummer in der erhaltenden Genbank, sowie Ursprungsland gespeichert.

Momentan registriert PGRDEU 156.200 Akzessionen zu 6.137 Taxa, die in 38 deutschen Einrichtungen erhalten werden (s. Tab. 1).

Tab. 1: Einrichtungen, deren Passportdaten z.T. in PGRDEU enthalten sind

Tab. 1: Institutions present in PGRDEU

Ort	Einrichtung
Ahrensburg	Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Institut für Zierpflanzenzüchtung
Alzey	Landesanstalt für Rebenzüchtung
Bamberg	Amt für Landwirtschaft und Ernährung Bamberg
Berlin	Humboldt-Univ. Berlin, Inst.für Gärtnerischen Pflanzenbau, Fachgebiet Obstbau
Braunschweig	Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Genbank
Burghaun	Rhöner Apfelbüro
Dodow	Fruchtquell GmbH & Co. KG
Dresden	Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Institut für Obstzüchtung Dresden – Pillnitz
Dresden	Institut für Pflanzengenetik und Kulturpflanzenforschung, Genbank, Außenstelle Süd
Dresden	Sächsische Landesanstalt für Landwirtschaft. Referat Obstbau
Dummerstorf	Mecklenburger Landschafts- und Territorientwicklung. Warnow-Ost e.V.
Erfurt	Lehr- und Versuchsanstalt Gartenbau Erfurt
Freiburg	Staatliches Weinbauinstitut
Freising	Inst. Obstbau und Baumschule, Staatl. Versuchsanstalt Gartenbau, FH Weihenstephan
Gatersleben	Institut für Pflanzengenetik und Kulturpflanzenforschung, Genbank
Geisenheim	Forschungsanstalt Geisenheim. Fachgebiet Rebenzüchtung und Rebenveredlung
Groß-Lüsewitz	Institut für Pflanzengenetik und Kulturpflanzenforschung, Genbank, Aussenstelle Nord
Hundisburg	Verein KULTUR-Landschaft Haldensleben-Hundisburg e.V.
Jork	OVA-Jork Betrieb Esteburg
Köln	Lehr- und Versuchsanstalt für Gartenbau, Fachschule für Floristik
Malchow/Poel	Institut für Pflanzengenetik und Kulturpflanzenforschung, Genbank, Aussenstelle Nord
Marquardt	Bundessortenamt, Prüfstelle Marquardt
Nebra	Naturpark Saale-Unstrut-Triasland e.V. Nebra
Neustadt (W)	Landes- Lehr- und Versuchsanstalt für Landwirtschaft, Weinbau und Gartenbau
Oppenheim	Staatliche Lehr- und Versuchsanstalt für Landwirtschaft, Weinbau und Gartenbau
Osnabrück	Fachhochschule Osnabrück, Fachbereich Gartenbau
Quedlinburg	Lehr- und Versuchsanstalt für Gartenbau und Technik des Landes Sachsen-Anhalt
Ravensburg	Univ. Hohenheim, Institut für Obst, Gemüse- und Weinbau, Fachgebiet Obstbau - Bavendorf
Rostock	LFA für Landwirtschaft und Fischerei. Forschungsstelle für nachwachsende Rohstoffe und Gartenbau
Sieboldingen	Bundesanstalt für Züchtungsforschung an Kulturpflanzen, Institut für Rebenzüchtung,
Trier	Landes-Lehr- und Versuchsanstalt fuer Wein und Gartenbau
Vechta	Versuchs- und Beratungsstation für Obst- und Gemüsebau Langförden der Landwirtschaftskammer Weser-Ems
Veitshöchheim	Bayerische LA für Weinbau und Gartenbau. Sachgebiet Obstbau Baumschule
Weidenbach	Landwirtschaftliche Lehranstalten Triesdorf
Weinsberg	Staatliche Lehr- und Versuchsanstalt für Wein- und Obstbau
Würzburg	Bayerische Landesanstalt für Weinbau und Gartenbau
Wurzen	Bundessortenamt, Prüfstelle Wurzen

Die Daten zur Aktualisierung von PGRDEU werden in regelmäßigen Abständen (etwa jährlich) von mehreren Einrichtungen geliefert. Dabei liefert nicht jede Einrichtung direkt an ZADI. Zum Beispiel werden bei Obstarten die Daten von 30 obstsortenerhaltenden Einrichtungen bereits von der IPK Genbank Außenstelle Süd in Dresden-Pillnitz zum *Bundesobstsortenregister* kompiliert. Diese Daten werden zwecks online Verfügbarmachung an die ZADI geliefert und sowohl als eigenständiges Datenangebot im WWW präsentiert (www.genres.de/bosr) als auch in PGRDEU übernommen.

Die Daten werden, nach geringfügiger Standardisierung, additiv in PGRDEU eingetragen, d.h. Daten, die einer bestimmten Akzession zugeordnet sind, können mehrfach in PGRDEU vorkommen, wenn sie z.B. aus unterschiedlichen Quellen entstammen. Ebenso treten potentielle Duplikate (gleichnamige Akzessionen der selben Fruchtart, die in verschiedenen Sammlungen erhalten werden; siehe Abb. 1) auch mehrfach auf. Die Duplikatenfrage, d.h. das Bestimmen, ob zwei getrennt gehaltene Akzessionen die gleiche genetische Ressource darstellen, ist eine Aufgabe u.a. der erhaltenden Einrichtungen – PGRDEU liefert hierzu Hinweise.

Mittlerweile benutzen die meisten datenliefernden Einrichtungen das vom International Plant Genetic Resources Institute (IPGRI, Rom) vorgeschlagene Austauschformat für Passportdaten, die *Multicrop Passport Descriptor List* (LIPMAN ET AL 1997). Dies erleichtert erheblich die Kompilierungsarbeit und erhöht den Informationsgehalt der Datenbank durch die Benutzung von Standardkodierungen, die von der kompilierenden Stelle automatisch für den Nutzer dekodiert werden können.

Nutzeroberfläche

Die Nutzeroberfläche von PGRDEU wird analog zu den weiteren online-PGR-Datenbanken der ZADI gestaltet. Es gibt zwei Hauptformen der Abfrage, mittels eines Formulars oder gezielt unter Benutzung der Indices der wichtigsten Datenbankfelder. Das Formular ermöglicht das Kombinieren von Suchkriterien, z.B. Suche nach Akzessionen der Spezies *Hordeum vulgare*, deren Ursprungsland Äthiopien ist. Dagegen werden bei der Indexsuche gezielt einzelne Felderinhalt abgefragt (in der Regel ergibt die Indexsuche schnellere Antworten als die Formularabfrage), z.B. Akzessionen deren Name bestimmte Teilworte enthalten (siehe Abb. 1).

Für die nächste Zukunft stehen Ausbaurbeiten an der Nutzeroberfläche an – ihre Funktionalität soll erweitert werden. Gedacht ist, eine dritte Variante des Suchens anzubieten, das sogenannte Browsing, womit der Nutzer sich, ausgehend z.B. von einer Liste der Gattungen, immer tiefer in die Datenbank bis hin zur gewünschten Informationen durchklicken kann (zum Beispiel: Gattung ⇒ Spezies ⇒ Sortenname ⇒ Erhaltende Einrichtung).

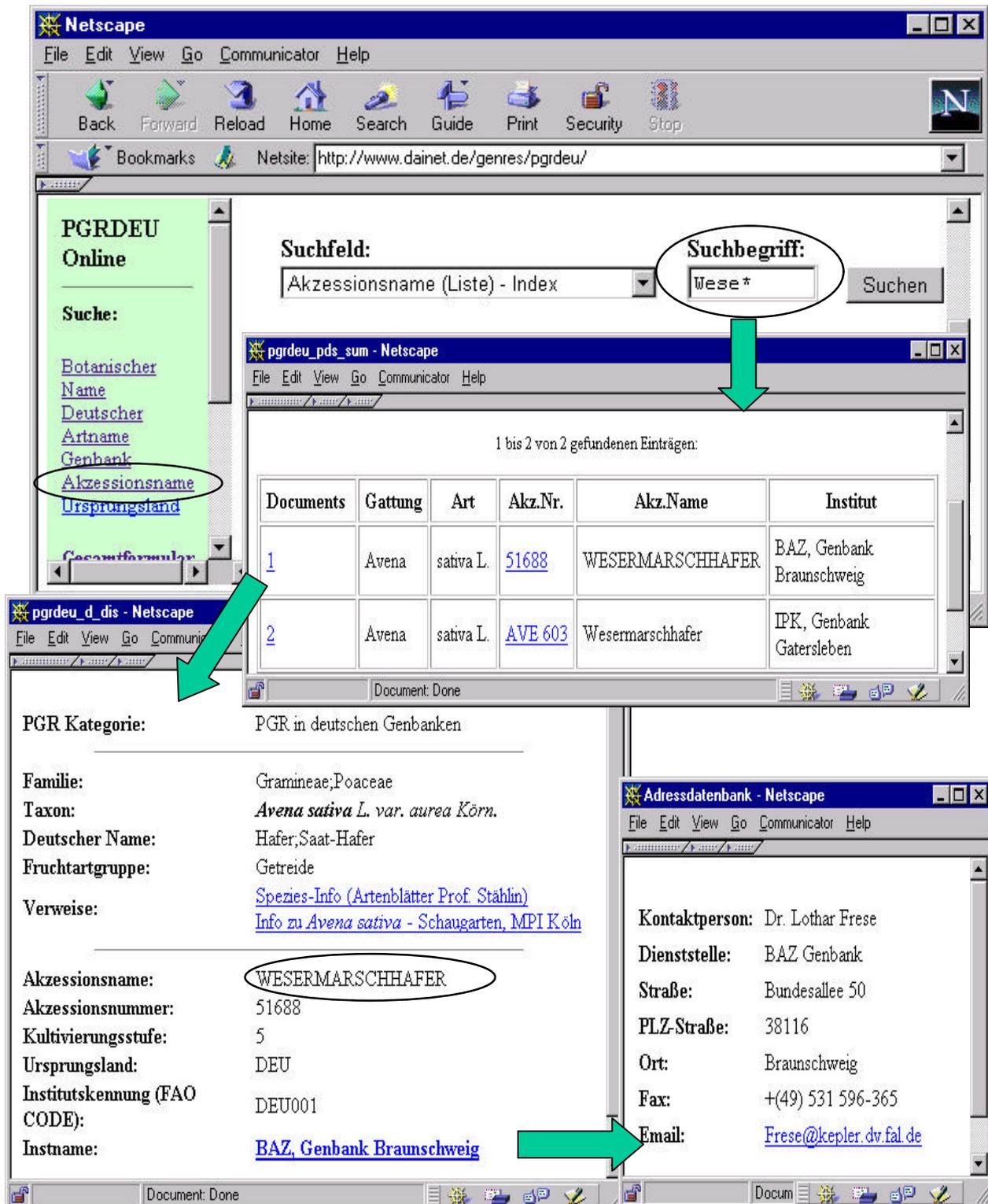


Abb. 1: Suche in PGRDEU

Fig. 1: Search for germplasm in PGRDEU

Verknüpfungen zu weiteren Informationen

Datensatzweise werden in PGRDEU Verweise zu weiteren online-Informationsquellen angeboten. Somit steht ein viel breiteres Spektrum an Informationen zur Verfügung, die einfach per Mausklick aufrufbar sind. Die Vernetzung durch Verweise auf unterschiedliche Informationssysteme definiert einen virtuellen Datenraum, der für den Nutzer einen deutlichen Mehrwert gegenüber der Summe der einzelnen Systeme hat.

Verweise auf der Taxonebene:

- online-Datenbank STAEHLIN (Artenblätter weltweit genutzter Pflanzenarten, siehe Beitrag HOHENESTER & STÄHLIN)
- "online-Schaugarten" des Max-Planck-Institutes für Züchtungsforschung

Verweise auf der Akzessionsebene:

- EVA-Datenbank (Evaluierungsdaten zu Kartoffel, Gerste und Apfel)
- Erreger-Datenbank (für Öl- und Futterpflanzen)

Das datensatzweise Einfügen weiterer Verweise auf externe Informationsquellen wird fortgesetzt. Demnächst werden solche Verweise auf die online-Datenbank der IPK-Genbank in Gatersleben (s. Beitrag KNÜPFER) sowie auf die Adressdatenbank der ZADI realisiert. Ersteres wird dem Nutzer ermöglichen, für Einträge in PGRDEU, die aus Gatersleben kommen, per Mausklick detailliertere Informationen aus der Originaldatenbank abzurufen.

PGRDEU und BIG

PGRDEU vertritt im Bundesinformationssystem Genetische Ressourcen (BIG) (s. Beitrag BECKER) u.a. auch die PGR-erhaltenden Einrichtungen, deren Passportdaten zum Teil in PGRDEU enthalten sind (siehe Tab. 1) aber selber ihre Daten nicht online verfügbar machen können. Das sind alle dort genannten mit Ausnahme der Genbank des IPK in Gatersleben, die über einen eigenen WWW- und online Datenbankserver verfügt. Somit wird gewährleistet, daß die Informationen zur *Ex-situ*- Erhaltung von PGR in Deutschland für BIG durch PGRDEU und IPK Gatersleben größtenteils abgedeckt sind. Zu den praktischen Aufgaben, die durch die Anbindung von PGRDEU an BIG entstehen, gehört u.a. die Abbildung der taxonomischen Information von PGRDEU in die taxonomische Datenbank von BIG.

PGRDEU beinhaltet die taxonomischen und Passport-Daten der Akzessionen, für die über EVA Evaluierungsdaten angeboten werden. Daher werden PGRDEU und EVA verknüpft zu einer virtuellen Datenbank in BIG teilnehmen.

Literatur

- BECKER, B. (1999): Bundesinformationssystem Genetische Ressourcen – Konzept und erste Ergebnisse. In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen. Schriften zu Genetischen Ressourcen, Bd. 12, ZADI, Bonn.
- BEGEMANN, F. (1999): Überblick über Konzepte und Entwicklungsschritte zur Dokumentation und Information pflanzengenetischer Ressourcen. In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ-KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen. Schriften zu Genetischen Ressourcen, Bd. 12, ZADI, Bonn.
- BÜCKEN, S. UND L. FRESE (1999): Das Informationsangebot der Genbank der BAZ. In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ-KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen. Schriften zu Genetischen Ressourcen, Bd. 12, ZADI, Bonn.
- HARRER, S. (1999): EVA-Aufbau eines Informationssystems für Evaluierungsdaten pflanzengenetischer Ressourcen in Deutschland. In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen, Schriften zu Genetischen Ressourcen Bd. 12, ZADI, Bonn.
- JIMENEZ KRAUSE, J.D. UND E. MÜNCH (1996): GENRES - Das deutsche Informationssystem für genetische Ressourcen im Internet. In: F. BEGEMANN, C. EHLING UND R. FALGE (Hrsg.): Vergleichende Aspekte der Nutzung und Erhaltung pflanzen- und tiergenetischer Ressourcen, Schriften zu Genetischen Ressourcen, Bd. 5, p. 258-268, ZADI, Bonn.
- KNÜPFER, H. (1999): Dokumentation pflanzengenetischer Ressourcen in der Genbank des IPK Gatersleben. In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ-KRAUSE (Hrsg.): Dokumentation und Informationssystem im Bereich der pflanzengenetischen Ressourcen, Schriften zu Genetischen Ressourcen, Bd. 12, ZADI, Bonn.
- LIPMAN, E., M.W.M. JONGEN, TH.J.L. VAN HINTUM, T. GASS AND L. MAGGIONI (eds.) (1997): Central Crop Databases: Tools for Plant Genetic Resources Management. 75-78. International Plant Genetic Resources Institute, Rome, Italy/CGN, Wageningen, The Netherlands.
- MÜNCH, E. (1999): Pflanzengenetische Ressourcen (PGR) im Informationssystem für Genetische Ressourcen (GENRES). In: F. BEGEMANN, S. HARRER UND J.D. JIMENEZ-KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich der pflanzengenetischen Ressourcen, Schriften zu Genetischen Ressourcen, Bd. 12, ZADI, Bonn.
- SCHMIDT, F. (1993): Dokumentation genetischer Ressourcen in der Bundesrepublik Deutschland – Ein Konzept für ein zentrales Informationssystem. Schriftenreihe des Bundesministeriums für Ernährung, Landwirtschaft und Forsten (BML), Angewandte Wissenschaft, Heft 422, p. 165-188, Bonn.
- SEIDEWITZ, L., L. FRESE UND M. DAMBROTH (1993): Datendokumentation und computergestützte Informationsbereitstellung der Sammlung pflanzengenetischer Ressourcen am Institut für Pflanzenbau. Schriftenreihe des Bundesministeriums für Ernährung, Landwirtschaft und Forsten (BML), Angewandte Wissenschaft, Heft 422, p. 131-143, Bonn.

EVA - Aufbau eines Informationssystems für Evaluierungsdaten pflanzengenetischer Ressourcen in der Bundesrepublik Deutschland

EVA – establishment of an information system for evaluation data of plant genetic resources in Germany

SIEGFRIED HARRER¹

Zusammenfassung

Ziel des Verbundprojekts EVA ist der Aufbau eines Informationssystems für Evaluierungsdaten pflanzengenetischer Ressourcen in der Bundesrepublik Deutschland. Im Rahmen von EVA wird dabei eine Online-Datenbank im Internet erstellt, wobei verschiedene Datenquellen zusammengefaßt werden. Hierfür wurden neue Datenstrukturen entwickelt. Zur Sekundärauswertung und notwendigen Aggregierung der verschiedenen Datenquellen wurden verschiedene statistische Verfahren geprüft bzw. modifiziert. Die Berechnung der Parameter für das zur Verfügung stehende Datenmaterial bei den EVA-Modellfruchtarten Gerste, Kartoffel und Obstarten (hauptsächlich Apfel) ist noch nicht abgeschlossen. Dennoch steht bereits ein Prototyp mit Teildatenbeständen für Online-Recherchen im Internet zur Verfügung (<http://www.genres.de/eva/>).

Abstract

The development of the Online-Information System for Evaluation Data (EVA, <http://www.genres.de/eva>) aims at establishing a central documentation system on the agronomic and breeding characteristics of germplasm. Different data sources have been available with different structures of the data. Thus, a standard data structure was developed. After having all data in a unique database, the next step is to develop methods for the proper interpretation and representation of the original data. At the moment methods for a combined statistical analysis are being developed. Currently, the EVA prototype database consists of evaluation data for barley, potato and fruit crops.

Einleitung

Ziel des 3-jährigen Verbundprojektes (1996-1999) ist der Aufbau eines Informationssystems zur Zusammenführung der in der Ressortforschung, universitären und außeruniversitären Forschung sowie in Sortenprüfungen anfallenden Evaluierungsdaten. Eine Nutzung der in diesen

¹ Zentralstelle für Agrardokumentation und –information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

Einrichtungen erfaßten Angaben zu den Werteigenschaften (Evaluierungsdaten) ist derzeit nur eingeschränkt möglich. So wird u.a. eine direkte Vergleichbarkeit der Angaben aus verschiedenen Quellen dadurch erschwert, daß die meisten Evaluierungsmerkmale in ihrer Ausprägung mittel bis stark durch die Umwelt beeinflußt werden (mittlere bis geringe Heritabilität der Merkmale). Des weiteren variieren die Methoden der Merkmalerfassung und Dokumentation. Durch die Integration dieser Informationen in ein einheitliches System und die damit verbundene Standardisierung der Datenstrukturen sowie die Möglichkeit zu Sekundärauswertungen solcher Daten soll der Zugriff auf die genetischen Ressourcen und damit ihre Nutzbarkeit erleichtert werden.

EVA ist dabei als Modellvorhaben zu sehen, in dem anhand verschiedener, heterogener Datenbestände bei Gerste, Kartoffel und Obstarten (hauptsächlich Apfel) die Realisierbarkeit sowie der mögliche Nutzen eines solchen Systems geprüft werden soll. Zur Gewährleistung eines schnellen und ständig verfügbaren Informationszugangs wird EVA als online-Informationssystem im Internet erstellt. Das Projekt wird durch das Bundesministerium für Ernährung, Landwirtschaft und Forsten (BML) gefördert.

Beteiligte Institutionen

Die **Genbank des Instituts für Pflanzengenetik und Kulturpflanzenforschung (IPK)** in Gatersleben beteiligt sich gemeinsam mit ihren **Außenstellen Groß Lüsewitz** (Kartoffel) und **Dresden-Pillnitz** (Obst) an dem Projekt. Am Standort Gatersleben werden in diesem Rahmen vorrangig Evaluierungsdaten von Gerste bearbeitet. Dem IPK und seinen beteiligten Außenstellen kommen im Rahmen dieses Projektes folgende Aufgaben zu:

- Rückwirkende Erfassung vorliegender langjähriger Evaluierungsdaten zu Gerste, Kartoffeln und Obstarten in computerlesbarer Form
- Entwicklung einer universellen Datenbankstruktur für Charakterisierungs- und Evaluierungsdaten im Rahmen des Genbank-Informationssystems des IPK, sowie Entwicklung von entsprechenden Datenerfassungsprogrammen
- Transformation bereits früher erfaßter Charakterisierungs- und Evaluierungsdaten der Gerste in diese Struktur
- Bereitstellung erfaßter Daten an den Projektpartner Martin-Luther-Universität Halle-Wittenberg (MLU) zur Erarbeitung von statistischen Auswertungsverfahren für genbankspezifische Evaluierungsdaten
- Bereitstellung erfaßter Daten an den Projektpartner ZADI zur Einbindung in das zentrale Informationssystem EVA

Am **Institut für Epidemiologie und Resistenz der Bundesanstalt für Züchtungsforschung an Kulturpflanzen (BAZ)** in Aschersleben werden in der dreijährigen Projektlaufzeit folgende Aufgaben bearbeitet:

- Rückwirkende Erfassung vorliegender langjähriger Resistenzevaluierungsdaten zu fast 8.000 Gerstenakzessionen aus der IPK-Genbank Gatersleben.
- Entwicklung von entsprechenden Datenstrukturen und Datenerfassungsprogrammen hierfür.
- Bereitstellung erfaßter Daten an den Projektpartner ZADI zur Einbindung in das zentrale Informationssystem EVA.

Die **Genbank** der **BAZ** in Braunschweig ist assoziierter Partner und stellt Daten von ca. 7.500 Gerstenakzessionen für das EVA-Projekt zur Verfügung.

Die **AG Datenspeicher** des Instituts für Acker- und Pflanzenbau der **Martin-Luther-Universität Halle-Wittenberg (MLU)** in Bad Lauchstädt ist innerhalb des Verbundprojektes für folgende Aufgaben verantwortlich:

- Bereitstellung von Daten zu Gerste und Kartoffel aus Landessortenversuchen (LSV) und Datenspeicher Feldversuche (DAVEP)
- Erarbeitung von Algorithmen zur Auswertung und Beschreibung der Daten aus Genbanken und Sortenversuchen
- Berechnung der entwickelten Parameter für das entsprechende Datenmaterial und Bereitstellung dieser Daten an den Projektpartner ZADI zur Einbindung in das zentrale Informationssystem EVA

Bei der **Zentralstelle für Agrardokumentation und -information (ZADI), Informationszentrum für Genetische Ressourcen (IGR)** in Bonn sind im wesentlichen folgende Aufgaben zu erledigen:

- Gesamtkoordination des Projektes.
- Entwicklung der Datenbanksysteme und Nutzeroberflächen für online-Recherchen.
- Erstellung und permanente Aktualisierung des Informationsangebotes im Internet.

Datenbasis

Für das EVA-Projekt stehen einerseits Daten zu Verfügung, welche erst noch EDV-gerecht erfaßt werden müssen, andererseits sind auch Daten bereits elektronisch verfügbar, welche dann teilweise noch entsprechend aufgearbeitet bzw. verfügbar gemacht werden müssen. Beim IPK Gatersleben lagen die Gerstendaten des Feldanbaues bis 1991 nur auf Karteikarten vor, hiervon wurden bisher über 60.000 Jahresdatensätze erfaßt worden. Die ab 1992 elektronisch verfügbaren Daten wurden bereits in die für EVA neu entwickelte Datenstruktur überführt und stehen seitdem ebenfalls für EVA zur Verfügung.

Die beim IPK Groß Lüsewitz auf Karteikarten vorliegenden und für das Projekt vorgesehenen Daten des Kulturkartoffelsortimentes (ca. 2.000 Sorten) wurden vollständig EDV-gerecht erfaßt. Sie wurden an die ZADI übermittelt und sind bereits online verfügbar. Eine Aufarbeitung dieser Daten durch den Projektpartner MLU war nicht notwendig, da diese Daten bereits durch das IPK „verdichtet“ worden sind.

Das IPK Dresden-Pillnitz hat Daten zu ca. 1.000 Apfelsorten vollständig elektronisch erfaßt. Innerhalb der Projektlaufzeit wurden in Dresden-Pillnitz noch die Daten zu Kirsche (ca. 200 Akzessionen Süßkirsche und ca. 120 Akzessionen Sauerkirsche) und Pflaume (ca. 200 Akzessionen) sowie Daten des Apfelwildsortimentes für EVA erfaßt. Die Daten wurden dort ebenfalls schon „verdichtet“, so daß auch hier eine Aufarbeitung durch den Projektpartner MLU nicht notwendig ist. Die Übergabe dieser Daten an ZADI ist bereits erfolgt und der größte Teil der Daten ist inzwischen online verfügbar.

Bei der BAZ Aschersleben konnte die Eingabe der Mehлтаudaten (ca. 8.000 Datensätze) abgeschlossen werden. Allerdings machten die Erfahrungen bei der Datenaufbereitung die Erstellung eines neuen Boniturschlüssels notwendig. Bereits eingegebene Daten mußten entsprechend dieses neuen Schlüssels konvertiert werden. Die Übergabe der Mehлтаudaten an ZADI ist bereits erfolgt, die Daten sind auszugsweise online verfügbar. In der restlichen Projektlaufzeit soll die bereits bestehende dBase-Datei mit Virusresistenzdaten für das EVA-Projekt umgearbeitet und angepaßt werden. Abschließend sollen noch Daten zu Rostkrankheiten und *Drechslera* eingegeben werden.

Die Gerstendaten der Genbank der BAZ in Braunschweig wurden an ZADI übergeben und in einer ersten Version online verfügbar gemacht. Eine Aufbereitung dieser Daten durch die MLU ist aufgrund der sehr heterogenen Datenstruktur nicht möglich.

Dem Projektpartner MLU Bad Lauchstädt wurden die LSV-Daten zu Wintergerste (zweizeilig und mehrzeilig), Sommergerste und Kartoffel vom Bundessortenamt (BSA) zur Verfügung gestellt. Die Daten umfassen in einzelnen Jahren allerdings nicht die Gesamtheit aller in der BRD durchgeführten LSV-Versuche zu diesen Kulturen. Nach erfolgreichem Test der Auswertungsalgorithmen wird MLU noch die für das Projekt vorgesehenen DAVEP-Daten entsprechend aufbereiten und über ZADI verfügbar machen.

Datenauswertung

Zur Aggregation der Daten aus Versuchen (LSV, DAVEP) und Erhebungen (Genbankdaten) ohne Wiederholungen werden durch MLU sowohl für quantitative (Erträge, Messungen etc.) als auch für qualitative (z.B. Bonituren) Merkmale Auswertungs- und Aggregierungsverfahren entwickelt. Da die für EVA bereitstehenden Daten aus unterschiedlichen Herkünften stammen und zudem aus Versuchen und Erhebungen, müssen die zu entwickelnden Lösungen eine weitgehend daten- und sortimentsunabhängige Einschätzung der Sorten/Genbankmuster ermöglichen. Deshalb ist es notwendig, mit Bezugsgrößen zu arbeiten. Für Daten aus Versuchen (LSV, DAVEP) wird hierzu für quantitative Merkmale der Sortimentsmittelwert der niedrigsten Intensitätsstufe verwendet. Für alle anderen Merkmale und Datengruppen (Genbankdaten) werden Minimum und Maximum je Ort und Jahr (wiederum in der niedrigsten Intensitätsstufe) als Bezugsgröße herangezogen.

Für quantitative Merkmale werden je Sorte und zugehörige Vergleichsbasis jahresweise arithmetischer Mittelwert, Minimum, Maximum, Variationskoeffizient und relative Spannweite berechnet. Zudem werden die Ökoregression (Bestimmtheitsmaß und Regressionskoeffizient, STEGEMANN ET AL. 1995, WRICKE 1965) und eine modifizierte Ökovalenz (WRICKE 1962, 1964, 1965) sowie die Differenz zwischen Sorte/Genbankmuster und der jeweiligen Vergleichsbasis mit zugehöriger Irrtumswahrscheinlichkeit P% ermittelt. Für den Nutzer von EVA besteht damit die Möglichkeit, durch Einbeziehung dieser Parameter die Ausprägung der entsprechenden Merkmale sehr viel detaillierter beurteilen zu können.

Bei qualitativen Merkmalen sind solche Auswertungen nur sehr eingeschränkt anwendbar. Statt dessen werden für diese Merkmale je Ort und Jahr ein Lageparameter berechnet und die Häufigkeitsverteilung dieses Parameters über die verschiedenen Umwelten eines Jahres ermittelt (HEINE UND WEBER 1982, SACHS 1974). Durch den Vergleich dieser Häufigkeitsverteilung mit der Häufigkeitsverteilung der Minima und Maxima der Vergleichsbasis ist eine bedingt umweltunabhängige Beurteilung der Eigenschaften einer Sorte/eines Genbankmusters möglich.

Nicht alle für EVA zur Verfügung stehenden Datengruppen eignen sich aber für derartige Auswertungen. Die Daten aus Aschersleben, Dresden-Pillnitz und Groß Lüsewitz wurden deshalb von den verantwortlichen Sortimentsbearbeitern aufbereitet, aggregiert und ohne weitere Auswertungen durch die MLU in EVA übernommen. Bei den Daten der Genbank der BAZ in Braunschweig waren, aufgrund der sehr heterogenen Datenstruktur, weitere Auswertungen ebenfalls nicht möglich.

EVA-Prototyp

Seit Ende 1997 steht eine Online-Version von EVA (<http://www.genres.de/eva/>) zur Verfügung. Dieser Prototyp wurde im November 1998 komplett überarbeitet und ergänzt. Derzeit beinhaltet der EVA-Prototyp auszugsweise die Gerstendaten der BAZ-Genbank sowie Daten des Instituts für Epidemiologie und Resistenz von mehrjährigen Mehlauresistenzuntersuchungen. Nach Abschluß der Teilprojekte in Groß Lüsewitz und Dresden-Pillnitz sind bei Kartoffel und Apfel umfangreiche Daten zu Qualität und Resistenz online recherchierbar. Ein weiteres Update (Resistenzdaten zu Gerste, weitere Obstdaten) steht unmittelbar bevor.

Der Zugang zum Informationssystem EVA (Abb.1) wurde völlig neu gestaltet. Interessierte Nutzer erhalten hier Informationen zum Projekt bzw. können sich als Testnutzer anmelden. Der Prototyp des EVA-Informationssystems ist derzeit frei nutzbar. Alle Informationen und Möglichkeiten des Systems stehen ohne Anmeldung zur Verfügung. Dennoch werden alle interessierten Nutzer gebeten, sich einmalig anzumelden und als „Beta-Tester“ des EVA-Informationssystems teilzunehmen. Daraus ergeben sich folgende Vorteile:

- Intensive Betreuung bei der Nutzung
- Sofortige Information über alle Neuerungen und Updates
- Die Möglichkeit, mit weiteren Interessierten den Endausbau des Informationssystems mitgestaltend zu begleiten

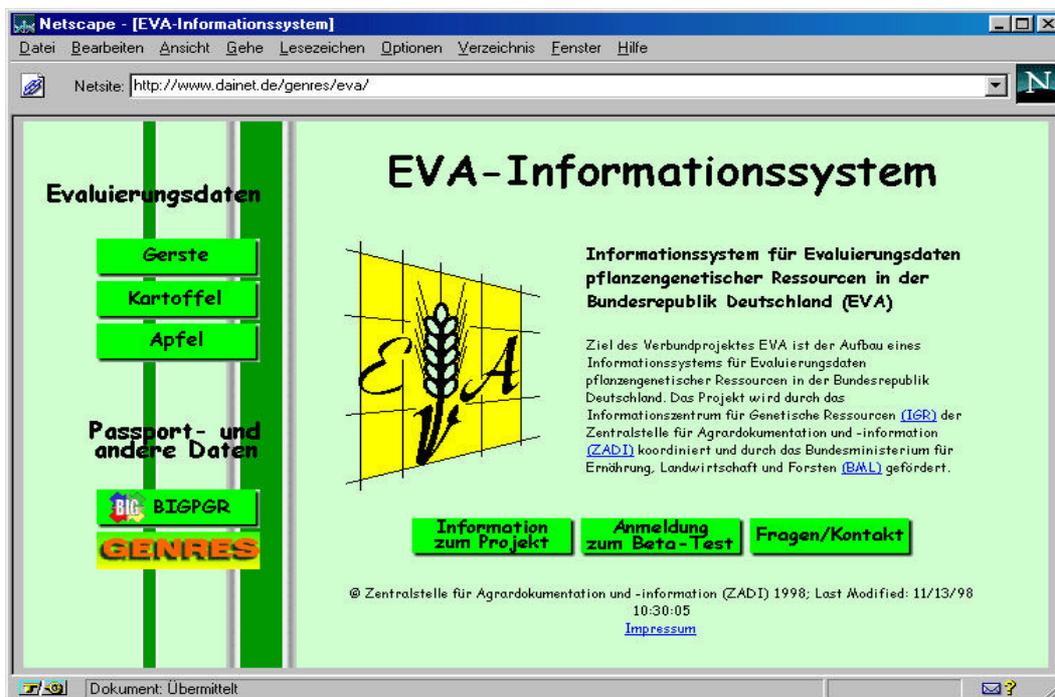


Abb. 1: Homepage des EVA-Informationssystems

Fig. 1: Homepage of the EVA Information System

Von der Homepage erfolgt direkt der Zugang zu den fruchtartenspezifischen Suchbildschirmen, welche einheitlich gestaltet sind und generell zwei verschiedene Zugänge zur Information anbieten:

- **Einfache Suche:** Aus einer Liste kann ein Merkmal ausgewählt werden, für das der Datenbestand recherchiert wird. Dies ermöglicht für wichtige Merkmale eine schnelle Übersicht über die Verteilung im Datenbestand, ferner können diese Merkmale natürlich auch anhand vorgegebener Suchkriterien recherchiert werden.
- **Kombinierte Suche:** Ein Suchformular mit den wichtigsten Merkmalen je Fruchtart ermöglicht eine kombinierte Suche nach mehreren Kriterien.

Als Ergebnis einer Suche erscheint zuerst eine zusammenfassende Anzeige. Sie bietet die Möglichkeit, das Suchergebnis weiter einzuschränken, bevor in der Endanzeige für jede Akzession/Sorte getrennt alle verfügbaren Informationen dargestellt werden (Abb. 2).

Resistenz mind. „7“ (Resistenz hoch) bei Phytophthora Blatt + Knolle, Blattrollvirus und Kartoffelvirus A und Y

Endanzeige (auszugsweise) für die Sorte „Roxy“

Dokument	Akzessionsnummer	Akzessionsname	Status der Akzession	Ursprungsland
23	1721	Paada		
22	1745	Hana		
22	1750	Polensky Bismar		
24	1777	Bibera (BBI)		
25	808	Eranta		
24	1812	Roxy		
22	1822	Saska		
28	1831	Sasut		
28	985	Sola		
28	1870	Sotka		
21	1889	Svetlana		
22	1890	Svetlana Krasny		
22	1892	Sagal		

Resistenz gegen Krankheiten und Stress:

Schorf: 8 BBD 81, 6 DDR-1

Dattelfleckenkrankheit: 1 CHE

Phytophthora inf. Blatt: 7 BBD 81

Phytophthora inf. Knolle: 7 BBD 81

Phytophthora R-gene: R?

Rhinizoma: 7 BBD 81

Kartoffelkrebs: R

Krebsrasen: 1

Erw. Schwarzeinigkeit: 7 BBD 81

Blattrollvirus: 8 BBD

Kartoffelvirus A: 9 BBD

Kartoffelvirus Y: 7 BBD

Kartoffelnematoden: R

Kartoffelnematoden-Rassen: Ro1

Sonstiger:

Gesundheitsstatus: Q1

Datenquelle: IFK, Genbank-Evaluierungsdaten, Groß Lüsewitz

Quellen:
 BRD Kartoffeln 1981, Nr. 4
 BRD 82 Kartoffeln 1982, Nr. 9
 BRD 81 Beschreibende Sortenliste 1981
 BRD 84 Beschreibende Sortenliste 1984
 CHE Hauptversuch auf Speisekartoffeln 1985 - 87
 DDR-1 Ergebnisse vom Groß Lüsewitz
 Kulturkartoffelzentrum

Abb. 2: Ergebnis (auszugsweise) einer Suche nach Kartoffelsorten mit hoher Resistenzprägung bei *Phytophthora* Blatt und Knolle, Blattrollvirus und Kartoffelvirus A und Y

Fig. 2: Result of a query for potato varieties with resistance against *Phytophthora* -leaf and tuber, leaf roll virus and potato virus A and Y

In der Endanzeige finden sich zudem für den Nutzer Angaben über die Datenquellen, und es besteht die Möglichkeit, zu der betreffenden Akzession/Sorte weitere Informationen aus der Datenbank zu Pflanzengenetischen Ressourcen in Deutschland PGRDEU abzurufen. Innerhalb der verschiedenen Informationsebenen kann jederzeit weiterführende Information z.B. über die Merkmalsdefinitionen abgerufen werden.

Datennutzung

Durch die intensive Information über das EVA-Projekt in den vergangenen Jahren sind bereits weite Kreise zukünftiger Nutzer aus Forschung und Privatwirtschaft an EVA interessiert. Die durchschnittlichen Zugriffszahlen auf den Prototypen liegen bereits jetzt bei ca. 300 Nutzerzugriffe je Monat. Nach dem Update und der damit verbundenen Erweiterung des Nutzerkreises ist eine deutlichen Steigerung zu erwarten. Die beteiligten Projektpartner haben mit Ausnahme von MLU Internetzugang (zumindest über Modem) und können das System zwischenzeitlich ebenfalls nutzen.

Weiterentwicklung der Datenstruktur für EVA

Das Informationssystem EVA ist Bestandteil der Dokumentation zu genetischen Ressourcen bei der ZADI. Es ist deshalb notwendig, daß die Datenstrukturen für das EVA-Informationssystem sowohl den augenblicklichen Gegebenheiten als auch zukünftigen Entwicklungen Rechnung tragen.

Dies erfordert, die Datenstrukturen entsprechend anzupassen und das EVA-System zu einem integralen Baustein der relationalen Datenstruktur der Dokumentation zu genetischen Ressourcen umzubauen. Dadurch können Redundanzen und Pflegeaufwand innerhalb der verschiedenen Dokumentationssysteme (PGRDEU, BIG, EVA etc.) stark reduziert werden, während die relationale Struktur direkten Informationszugriff von jedem System aus erlaubt. Weiterhin garantiert die gewählte Struktur auch künftig Kompatibilität zwischen den verschiedenen Systemen.

Ausblick

In der verbleibenden Projektlaufzeit liegt der Schwerpunkt der Arbeiten in der Weiterentwicklung des EVA-Prototypen. Dabei kann zum einen auf die Erfahrungen des mit guter Beteiligung angelaufenen Testbetriebes zurückgegriffen werden. Andererseits ist die internationale Entwicklung bei Dokumentationssystemen für Evaluierungsdaten (z.B. GRIN in USA) weiter zu verfolgen. Ziel ist es, den Zugang zu den verschiedenen Informationsebenen weiter zu vereinfachen und das Informationssystem EVA noch besser mit anderen Informationssystemen zu verknüpfen. Um den Nutzerkreis zu erweitern, ist auch eine englischsprachige

EVA-Version vorgesehen. Insgesamt wurden bisher gute Fortschritte bei den Arbeiten im Gesamtprojekt erzielt. Es zeichnet sich ab, daß sich das EVA-Informationssystem zu einem sinnvollen und nützlichen Werkzeug entwickeln läßt, für welches, nicht zuletzt auch aufgrund steigender Internetnutzung von Seiten der Wirtschaft, ein wachsendes Interesse zu verzeichnen ist. Hierzu ist es nötig, das Datenangebot zu erweitern und die Beteiligung weiterer Forschungseinrichtungen und der Privatwirtschaft sicherzustellen. Es sollten deshalb Überlegungen angestellt werden, EVA im Anschluß an die Pilotphase auch auf andere Fruchtarten und Werteeigenschaften auszudehnen sowie die Zugangsbedingungen zu den Evaluierungsdaten zu regeln. Dabei spielen u.a. wissenschaftlich-technische Umsetzungsmöglichkeiten sowie rechtliche Rahmenbedingungen eine Rolle.

Literatur

- HEINE, H. UND W.E. WEBER (1982): Die Aussagekraft statistischer Maßzahlen für die phänotypische Stabilität in amtlichen Sortenprüfungen bei Winterweizen und Körnermais. Z. Pflanzenzüchtung 89, S. 89-99.
- SACHS, L. (1974): Angewandte Statistik. Springer Verlag Berlin, Heidelberg, New York.
- STEGEMANN, K., H. DÖRFEL UND V. WEISE (1995): Methodische Untersuchungen zur Sekundärauswertung von Sortenversuchen bezüglich der Ertragsstabilität von Sorten mit Hilfe der Ökoregression. Arch. Acker-Pfl. Boden.39, S. 389-400.
- WRICKE, G. (1962): Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. Z. Pflanzenzüchtung 47, S. 92-96.
- WRICKE, G. (1964): Zur Berechnung der Ökovalenz bei Sommerweizen und Hafer. Z. Pflanzenzüchtung 52, S. 127-138.
- WRICKE, G. (1965): Die Erfassung der Wechselwirkung zwischen Genotyp und Umwelt bei quantitativen Eigenschaften. Z. Pflanzenzüchtung 53, S. 266-343.

Datenbanken mit molekulargenetischen Daten pflanzlicher Genome

Databases related to molecular genetics of plant genomes

WOLFGANG MICHALEK¹

Zusammenfassung

Fortschritte in den molekularbiologischen Methoden der Genomanalyse haben es ermöglicht, in kurzer Zeit große Mengen an Daten zu produzieren. Dazu zählen genetische Daten, gewonnen durch die Anwendung molekularer Marker sowie Nukleinsäure- und Proteinsequenzen. Datenbanken, welche mit molekularen Markern gewonnene Daten pflanzlicher Genome beinhalten, sind über AGIS (Agricultural Genome Information System) erreichbar. Hierbei handelt es sich in erster Linie um Nutzpflanzen und deren Schadorganismen.

Die Automatisierung der Sequenzanalyse von Nukleinsäuren hat vor allem im Bereich der Analyse von Säugetiergenomen und sogenannten Modellorganismen große Datenmengen entstehen lassen, eine Entwicklung, die auch bei der Analyse von Nutzpflanzengenomen eintritt. In erster Linie handelt es sich hierbei um die Erstellung von ESTs (expressed sequence tagged sites), die durch partielle Sequenzierung von cDNA Klonen gewonnen werden. Für die Verwaltung von Sequenzdaten sind die Datenbanken GenBank, EMBL und DDBJ von zentraler Bedeutung. Enge Kooperationen sorgen für die Abstimmung der jeweiligen Inhalte. Für den Nutzer werden über Internet Werkzeuge zur Datenbankabfrage und zu Sequenzvergleichen zur Verfügung gestellt.

Abstract

Advanced methods in molecular biology made it possible to create large amounts of data in the field of genome analysis. This are genetical data, produced by the utilization of molecular markers as well as nucleic acid- and protein-sequences. Databases including data of plant genomes analysed with molecular markers are available via AGIS (Agricultural Genome Information System). This databases mainly provide information, regarding crop plants and their pathogens.

Automation of nucleic acid sequence analysis resulted in large amounts of sequence data, in the first instance in the area of mammalian genomes and so-called model-organisms. This development has also been observed in the field of genome analysis of agriculturally important plant species. In the first place the production of ESTs (expressed sequence tagged sites), partially

¹ Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK)

sequenced cDNA clones, represents the biggest part of sequence data produced in the last few years. The most important databases for the management and storage of sequence data are GenBank, EMBL and DDBJ. All databases are available on the internet and users are provided with tools for data retrieval and sequence comparisons.

Einleitung

Die in den späten 70er Jahren erzielten Fortschritte bei der Klonierung und Sequenzierung von DNA brachte die Notwendigkeit mit sich, große Menge an Sequenzdaten zu speichern und zu verwalten. Die systematische Erfassung und Speicherung von Nukleotidsequenzen begann mit der 1980 von EMBL (European Molecular Biology Laboratory) etablierten Datenbank. Die rasant gestiegene Bedeutung der Molekularbiologie und der Einsatz von Sequenzierautomaten hat den Umfang an Daten sowohl in Menge als auch Vielfalt stark anwachsen lassen. Neben der kompletten Sequenzierung der Genome verschiedener Mikroorganismen wie *Saccharomyces cerevisiae* (GOFFEAU ET AL. 1996), *Escherichia coli* (BLATTNER ET AL. 1997) wird auch die Sequenzaufklärung des menschlichen Genoms mit Nachdruck verfolgt. Folgerichtig werden auch im Bereich der pflanzlichen Genome Anstrengungen zur Totalsequenzierung unternommen. Im Vordergrund stehen dabei die Aktivitäten bei *Arabidopsis thaliana* und *Oryza sativa*.

Während die Strukturaufklärung ganzer Genome auf Nukleinsäureebene aufgrund des erheblichen finanziellen Aufwands auf wenige Arten beschränkt bleiben wird, haben sich im Bereich der genetischen Kartierung und phylogenetischer Studien molekulare Marker durchgesetzt. Dabei steht die Bearbeitung landwirtschaftlich bedeutsamer Kulturpflanzen im Vordergrund. In diesem Beitrag soll ein Überblick über die wichtigsten Datenbanken gegeben werden, in welchen Daten der Nukleinsäure- und Proteinsequenzierung und Daten zu molekularen Markern bzw. deren genetischer Kartierung der Öffentlichkeit zugänglich gemacht werden.

Datenbanken des AGIS (Agricultural Genome Information System)

Der Server des AGIS (<http://probe.nalusda.gov:8000/index.html>) enthält Informationen über Genome landwirtschaftlich bedeutsamer Organismen. Sämtliche Datenbanken wurden mit ACEDB Software erstellt und können entsprechend abgefragt werden. Zur Zeit besteht der Datenbestand hauptsächlich aus Informationen zu Nutzpflanzen und -tieren. Darüber hinaus werden auch Daten gesammelt, die nicht direkt nutzbaren Spezies zuzuordnen sind (Nomenklatur von Genen, Information zu genetischen Ressourcen). Die in AGIS zusammengefassten Datenbanken bilden die zentrale Ressource für Daten, die im Zusammenhang mit der Anwendung molekularer Marker erzeugt wurden. Eine Übersicht über die Datenbanken zu verschiedenen pflanzlichen Genomen, welche zur Zeit verfügbar sind, gibt Tabelle 1.

Tab. 1: AGIS Datenbanken pflanzlicher Genome

Tab. 1: Databases of plant genomes in AGIS

Datenbank	Inhalt
AAAtDB	<i>Arabidopsis thaliana</i>
Alfagenes	Alfalfa (<i>Medicago sativa</i>)
BeanGenes	<i>Phaseolus</i> and <i>Vigna</i>
CabbagePatch	<i>Brassica</i>
ChlamyDB	<i>Chlamydomonas reinhardtii</i>
CoolGenes	Futterleguminosen
CottonGenes	<i>Gossypium hirsutum</i>
GrainGenes	Weizen, Gerste, Roggen und verwandte Arten
MaizeDB	Mais
MilletGenes	Hirse
PathoGenes	Pilzliche Krankheitserreger von Getreiden
Ricegenes	Reis
RoseDB	Rosaceae
SolGenes	Solanaceae
SorghumDB	<i>Sorghum bicolor</i>
SoyBase	Sojabohne
TreeGenes	Waldbäume
Mendel	Nomenklatur pflanzlicher Gene

Der Umfang der verschiedenen Datenbanken variiert in Abhängigkeit der Spezies, deren Daten gesammelt werden, stark. Als Beispiel soll auf die relativ umfangreiche Datenbank GrainGenes (<http://wheat.pw.usda.gov>), in welcher Daten zu Weizen, Gerste, Roggen, Hafer und Zuckerrohr gesammelt werden, näher eingegangen werden. Diese Datenbank umfaßt Informationen zu:

- genetischen und cytogenetischen Karten
- genomischen Sonden, Nukleotidsequenzen
- Genen, Allelen, Genprodukten
- Phänotypen, quantitativen Merkmalen, QTLs (Quantitative Trait Loci)
- Genotypen und Stammbäumen von Sorten etc.
- Krankheitsbildern und Pathogenen
- Taxonomie der Triticeae
- Adressen und Arbeitsgebieten von Wissenschaftlern
- relevanten bibliographischen Daten

Der Umfang der gespeicherten Informationen nimmt rasant zu. Tabelle 2 zeigt die Anzahl der Einträge in verschiedenen ausgewählten Datenklassen in GrainGenes. Dabei ist ersichtlich, daß die Einträge in verschiedenen Datenklassen unterschiedlich stark zunahmen. Dies spiegelt auch die Zielrichtungen wider, in welche sich in den letzten Jahren Arbeitstechniken entwickelten bzw. in welchen Themenkreisen Forschungsschwerpunkte angesiedelt waren. Beispielsweise verdoppelte sich die Anzahl molekularer Sonden lediglich, während in weit größerem Umfang Sequenzdaten erzeugt wurden. Die Ursache hierfür ist unter anderem die wachsende Bedeutung der PCR Technik bei der Anwendung molekularer Marker, deren Voraussetzung oftmals das Vorliegen von Sequenzinformation ist. Unter den Forschungszielen hat die genetische Kartierung von QTLs große Fortschritte gemacht, erkennbar an dem circa um das hundertfache gewachsenen Datenumfang zu diesem Thema. Diese Entwicklung konnte einsetzen, nachdem als Grundvoraussetzung die Entwicklung genetischer Karten mit ausreichender Markerdichte weit genug fortgeschritten war.

Tab. 2: Anzahl der Einträge in GrainGenes

Tab. 2: Number of entries in GrainGenes

Datenklasse (Auswahl)	Einträge 1995	Einträge 01/1998
Loci	7.400	14.755
Sonden	4.000	8.147
Polymorphismen	1.600	2.496
Sequenzen	200	2.967
Gene	660	2.448
QTL	4	404

Welche Arten von Daten in welcher Menge in Zukunft erhoben werden, läßt sich nur schwer vorhersagen. Mit hoher Wahrscheinlichkeit werden große Mengen an Sequenzdaten durch die partielle cDNA Sequenzierung zur Erzeugung von ESTs (expressed sequence tags) erzeugt werden. Diese Arbeiten wurden bereits an verschiedenen Pflanzenarten (*Arabidopsis*, Reis und andere) begonnen und die Sequenzen vor allem in Datenbanken wie GenBank gespeichert.

Internationale Sequenzdatenbanken

Ausgewählte Datenbanken für die Erfassung von Nukleotid- und Proteinsequenzen sind in Tabelle 3 aufgelistet. Zwischen diesen Datenbanken herrschen enge Kooperationen, so sind beispielsweise PIR (Protein Information Resource) und SwissProt auch über die Server des NCBI (National Center for Biotechnology Information, GenBank) bzw. EBI (European Bioinformatics Institute, EMBL) abzufragen. Zwischen den drei großen Nukleotiddatenbanken findet ein

täglicher Datenaustausch statt (Abb. 1), so daß sich für den Anwender das Abfragen der einzelnen Datenbanken erübrigt.

Tab. 3: Ausgewählte Nukleotid- und Proteinsequenzdatenbanken

Tab. 3: Selected databases for nucleotid and protein sequences

Datenbank	Inhalt	URL	Referenz
EMBL	Nukleotidsequenzen	http://www.ebi.ac.uk/ebi_docs/embl_db/ebi/topembl.html	STOESSER ET AL. 1998
GenBank	Nukleotid- und Proteinsequenzen	http://www.ncbi.nlm.nih.gov/Web/Genbank/index.html	BENSON ET AL. 1998
DDBJ	Nukleotidsequenzen	http://www.ddbj.nig.ac.jp/Welcome-e.html	
PIR	Proteinsequenzen	http://nbrfa.Georgetown.Edu/pir/	GEORGE ET AL. 1998
SwissProt	Proteinsequenzen	http://expasy.hcuge.ch/sprot/sprot-top.html	BAIROCH ET AL. 1998

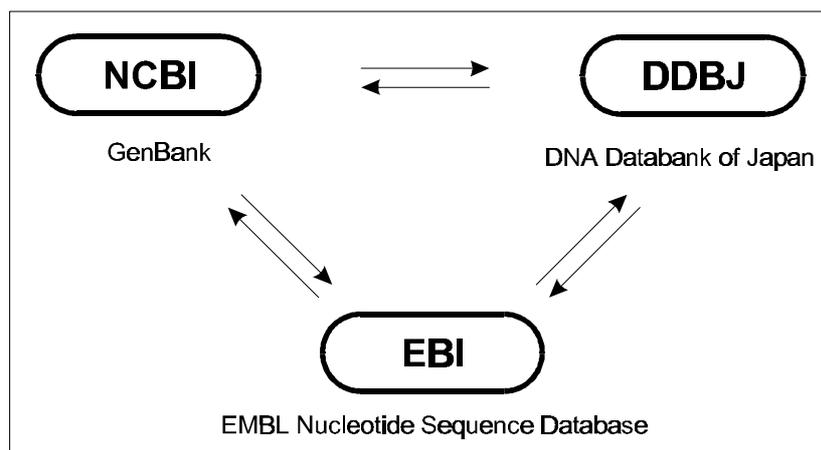


Abb. 1: Datenaustausch zwischen GenBank/EMBL/DDBJ

Fig. 1: Data exchange between GenBank/EMBL/DDBJ

Am Beispiel von GenBank sollen stellvertretend der Inhalt und Umfang einer Sequenzdatenbank dargestellt werden. Detaillierte Informationen finden sich unter <ftp://ncbi.nlm.nih.gov/genbank/gbrel.txt>. GenBank weist zur Zeit (01/98) circa 2,2 Mio. Einträge auf. Die Vielzahl der Einträge ist vor allem auf umfangreiche Aktivitäten zur Sequenzierung von ESTs zurückzuführen, die im Bereich Säugetiergenome ihren Anfang nahm.

Eine Auflistung der Einträge in Abhängigkeit vom betreffenden Organismus findet sich in Tabelle 4. Unter den am häufigsten vertretenen Organismen finden sich zwei Pflanzenarten, *Arabidopsis thaliana* und *Oryza sativa*.

Tab. 4: Anzahl der Einträge pro Organismus (01/98)

Tab. 4: Number of entries per organism (01/98)

Einträge	Basen	Spezies
1.038.373	551.650.819	<i>Homo sapiens</i>
281.308	133.404.600	<i>Mus musculus</i>
75.960	113.659.949	<i>Caenorhabditis elegans</i>
56.146	46.566.672	<i>Arabidopsis thaliana</i>
26.211	29.198.918	<i>Drosophila melanogaster</i>
10.474	28.567.944	<i>Saccharomyces cerevisiae</i>
4.743	17.345.841	<i>Escherichia coli</i>
10.517	13.826.989	<i>Rattus norvegicus</i>
1.058	9.715.721	<i>Bacillus subtilis</i>
21.006	9.198.129	<i>Human immunodeficiency virus type 1</i>
22.270	8.932.632	<i>Oryza sativa</i>

Zum weitaus größten Teil bestehen die Einträge aus EST-Sequenzen, welche man durch die partielle Sequenzierung von cDNA-Enden erhält. cDNA wird aus mRNA synthetisiert, woraus sich schließen läßt, daß exprimierte Teile eines Genoms bearbeitet werden. Derartige Sequenzen reflektieren den Expressionsstatus eines Gewebes zu einem spezifischen Zeitpunkt und bestimmen somit charakteristische Eigenschaften eines Organismus. Wenngleich der überwiegende Anteil von EST-Sequenzen in Säugetier- und sogenannten Modellgenomen wie *Caenorhabditis elegans* gewonnen werden, liegen auch zu pflanzlichen Genomen bereits mehr als 65.000 Einträge vor (Tab. 5).

Tab. 5: Anzahl EST pflanzlicher Genome in GenBank (Stand 01/98)

Tab. 5: Number of EST from plant genomes in GenBank (01/98)

Spezies	Anzahl EST	Spezies	Anzahl EST
<i>Arabidopsis thaliana</i>	36.173	<i>Sorghum bicolor</i>	93
<i>Oryza sativa</i>	20.708	<i>Citrus sinensis</i>	87
<i>Zea mays</i>	1.783	<i>Allium cepa</i>	78
<i>Brassica napus</i>	1.427	<i>Pinus radiata</i>	67
<i>Brassica campestris</i>	965	<i>Pisum sativum</i>	66
<i>Medicago truncatula</i>	899	<i>Solanum tuberosum</i>	49
<i>Pinus taeda</i>	856	<i>Nicotiana tabacum</i>	49
<i>Ricinus communis</i>	750	<i>Hordeum vulgare</i>	45
<i>Citrus unshiu</i>	426	<i>Daucus carota</i>	23
<i>Saccharum sp.</i>	293	<i>Triticum aestivum</i>	4
<i>Citrullus lanatus</i>	163	<i>Gossypium hirsutum</i>	2
<i>Avena sativa</i>	125	<i>Lycopersicon esculentum</i>	2
<i>Glycine max</i>	97	<i>Vigna radiata</i>	1

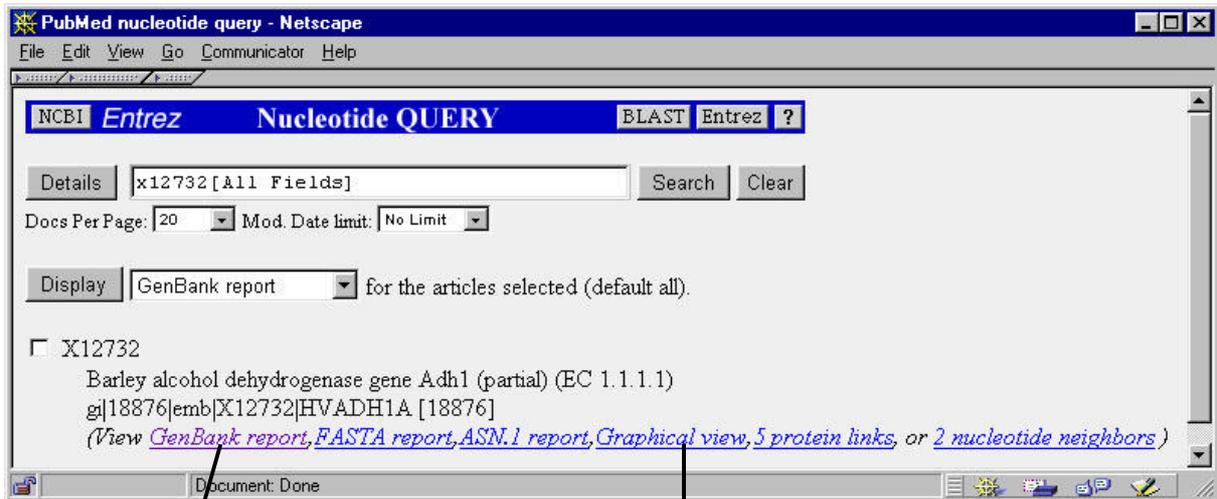
Es wird erwartet, daß diese Zahl in naher Zukunft stark anwachsen wird. Aktuelle Daten lassen sich unter http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html abrufen. Der zweitstark expandierende Bereich der Sequenzdatenbanken beinhaltet genomische Sequenzen, die in Projekten gewonnen werden, welche die Sequenzaufklärung ganzer Genome zum Ziel haben. Im Bereich pflanzlicher Genome liegen zur Zeit umfangreiche Daten nur zu *Arabidopsis thaliana* vor (<http://genome-www.stanford.edu/Arabidopsis/agi.html>), als erstes Genom einer Nutzpflanze wird angestrebt, das Reisgenom zu sequenzieren (<ftp://genome1.bio.bnl.-gov/pub/maize/RiceProject.html>).

Für die Nutzung von Sequenzdatenbanken sind leistungsfähige Abfragesysteme eine Grundvoraussetzung. Im Falle von GenBank steht über das Internet das *Entrez*-System zur Verfügung, über welches sowohl auf DNA- und Proteinsequenzdaten, relevante MEDLINE Einträge, als auch sonstige Ressourcen am NCBI zugegriffen werden kann. Abfrageergebnisse beinhalten Hyperlinks, die dem Nutzer Anschluß an weitere Informationsquellen ermöglichen. Abb. 2 skizziert die Möglichkeiten, zwischen verschiedenen Darstellungsformen und Informationsquellen zu wechseln. So existieren z. B. bei der dargestellten Sequenz Links zur entsprechenden AGIS-Datenbank, in diesem Falle GrainGenes, in welcher dann weitere Informationen über angefragte Sequenzen zur Verfügung stehen.

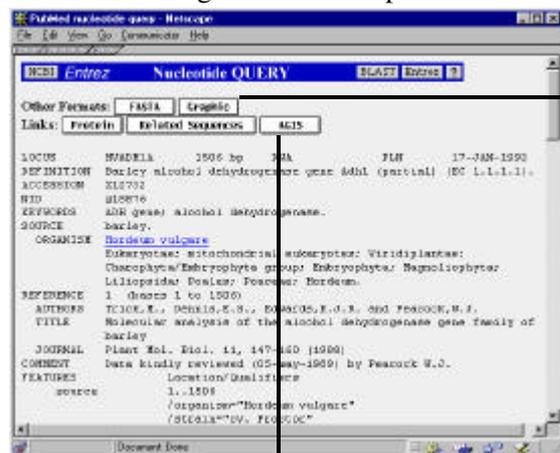
Ähnlich stellt sich die Situation bei der Abfrage von EBI-Datenbanken dar: Die Abfragemaske des hier verwendeten Sequence Retrieval Systems (SRS) ermöglicht detaillierte Abfragen der am EBI installierten Datenbanken (circa 50), von welchen aber nur einige für pflanzliche Genome relevante Daten enthalten.

Die wohl häufigste Art der Nutzung von Sequenzdatenbanken sind Sequenzvergleiche. Bei der Verwendung von GenBank stehen die Programme der BLAST (Basic Local Alignment Search Tool, ALTSCHUL ET AL. 1990) Familie zur Verfügung. Sie ermöglichen Sequenzvergleiche zwischen Aminosäuresequenzen und Proteindatenbanken (BLASTP), zwischen Nukleinsäuresequenzen und deren Datenbanken (BLASTN) sowie weitere Möglichkeiten, die jeweils die Übersetzung der sechs möglichen Leserahmen einer Nukleinsäuresequenz in eine Aminosäuresequenz, sei es seitens der Anfrage- oder seitens der Datenbanksequenz, beinhalten. Seit kurzem sind modifizierte BLAST-Programme verfügbar, die Verbesserungen der Suchalgorithmen bzw. des Ausgabeformats (graphisch) aufweisen. Zugangsmöglichkeiten bestehen sowohl über WWW als auch Email-Server. Die Nutzung der Email-Server erfordert im Gegensatz zu den Formaten, die mit WWW-Browsern zugänglich sind, die Angabe der Befehle in einem einfachen Textformat. Der Vorteil ist vor allem die einfachere Bearbeitung mehrerer Sequenzen in Folge und die Umgehung eventueller Wartezeiten bei starker Inanspruchnahme des Servers oder des Netzes.

Ergebnis einer GenBank Abfrage / Result of a GenBank query



GenBank Eintrag / GenBank Report



Graphische Ansicht / Graphical view



Grain Gene Server



Abb. 2: Der Entrez Browser: Mögliche Links und Darstellungsformen

Fig. 2: The Entrez Browser: Links and views

Parallel zur wachsenden Vielfalt an Informationen steigt der Bedarf an geeigneten Werkzeugen, die zum einen dazu dienen, Information zu finden und zum anderen, sich diese nutzbar zu machen. Unter <http://www-biol.univ-mrs.fr/english/logligne.html> findet sich eine Sammlung von Adressen, die die online-Bearbeitung von Sequenzdaten ermöglicht und unter http://www.biophys.uni-duesseldorf.de/bionet/research_tools.html ist eine weiter gefaßte Sammlung von Adressen zugänglich, die auch Zugangsmöglichkeiten zu Literatur, Protokollen und anderen molekulargenetisch relevanten Datenquellen auflistet.

Ausblick

Datenbanken, in welchen die verschiedensten Arten von Informationen gesammelt und über das WWW zugänglich gemacht werden, wachsen täglich. Genomprojekte für alle wichtigen Kulturpflanzenarten werden sowohl in den USA als auch in Europa geplant und werden in naher Zukunft vor allem Sequenzdaten exprimierter Bereiche (ESTs) in großer Menge liefern. Die multiparallele Analyse exprimierter Sequenzen kann neuartige Möglichkeiten der Genomanalyse schaffen. Während herkömmliche molekulare Marker weitgehend zufällig im Genom verteilt sind, erhält man auf der Basis von ESTs Zugriff auf die funktionellen Bestandteile eines Genoms. Zur Zeit werden zwar vor allem Genome sogenannter Modellorganismen und Säugetiergenome intensiv bearbeitet, es ist jedoch absehbar, daß auch bei der Analyse pflanzlicher Genome große Fortschritte gemacht werden.

Neben den beschriebenen großen Datenbanken etablieren sich Datenbanken mit spezialisierten Inhalten, beispielsweise für Taxonomie, Metabolismus oder Kohlehydrate (Übersicht: SOMERVILLE ET AL. 1997). In Verbindung mit Datenbanken genetischer Ressourcen, welche in GENRES gelistet sind, ergibt sich eine Vielfalt von Informationsquellen, was der Bioinformatik wachsende Bedeutung hinsichtlich ihrer Integration, Nutzung und Auswertung zukommen läßt. Gerade durch die Anwendung molekularer Techniken zur Evaluierung genetischer Ressourcen ergeben sich Zusammenhänge, die die Integration von Daten aus verschiedenen Bereichen notwendig machen. Nur durch die sinnvolle Vernetzung und die Entwicklung geeigneter Analysewerkzeuge kann eine effektive Nutzung der Fülle molekularbiologischer Daten ermöglicht werden, welche in stetig wachsendem Umfang produziert werden.

Literatur

- ALTSCHUL S.F, W. GISH, W. MILLER, E.W. MYERS AND D.J. LIPMAN (1990): Basic local alignment search tool. *J Mol Biol* 215: 403-10.
- BAIROCH A. AND R. APWEILER (1998): The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. *Nucleic Acids Res* 16: 38-42.
- BARKER WC, J.S. GARAVELLI, D.H. HAFT, L.T. HUNT, C.R. MARZEC, B.C. ORCUTT, G.Y. SRINI-VASARAO, L.L. YEH, R.S. LEDLEY, H.W. MEWES, F. PFEIFFER AND A. TSUGITA (1998): The PIR - International Protein Sequence Database. *Nucleic Acids Res* 16: 27-32.
- BENSON D.A, MS. BOGUSKI, D.J. LIPMAN, J. OSTELL AND B.F. OUELLETTE (1998): GenBank. *Nucleic Acids Res* 16: 1-7.
- BLATTNER F.R, G. PLUNKETT, C.A. BLOCH, N.T. PERNA, V. BURLAND, M. RILEY, J. COLLADO-VIDES, J.D. GLASNER, C.K. RODE AND G.F MAYHEW (1997): The complete sequence of *Escherichia coli* K-12. *Science* 277: 1453-1462.
- GOFFEAU A, B.G. BARRELL, H. BUSSEY, R.W. DAVIS, B. DUJON, H. FELDMANN, F. GALIBERT, J.D. HOHEISEL, C. JACQ, M. JOHNSTON, E.J. LOUIS, H.W. MEWES, Y. MURAKAMI, P. PHILIPPSEN, H. TETTELIN AND S.G. OLIVER (1996): Life with 6000 genes. *Science* 274: 546-567.
- SOMERVILLE C, D. FLANDERS AND J.M. CHERRY (1997): Plant biology in the Post-Gutenberg era. *Plant Physiol* 113: 1015-1022.
- STOESSER G, M.A. MOSELEY, J. SLEEP, M. MCGOWRAN, M. GARCIA-PASTOR AND P. STERK (1998): The EMBL Nucleotide Sequence Database. *Nucleic Acids Res* 16: 8-15.
- TATENO Y, K. FUKAMI-KOBAYASHI, S. MIYAZAKI, H. SUGAWARA AND T. GOJOBORI (1998): DNA Data Bank of Japan at work on genome sequence data. *Nucleic Acids Res* 16: 16-20.

Die Stählin-Datenbank weltweit genutzter Pflanzenarten der Gymnospermae und Angiospermae.

The „Stählin“ database of worldwide utilized plants of gymnosperms and angiosperms

ARMIN HOHENESTER¹ UND LIESELOTTE STÄHLIN²

Kurze Präsentation als Einführung

Die Datenbank (STAEHLIN-DB) beinhaltet ca. 18.200 Beschreibungen von Pflanzenarten, welche im weitesten Sinne menschlich genutzt werden. Die einzelnen Artenblätter sind einheitlich strukturiert und nach vielen Kriterien recherchierbar. Die Präsentation der einzelnen Artenblätter erfolgt generell mittels eines Browsers (z.B. Netscape oder Explorer) als Dokument variabler Länge in druckreifer Form. Der Umfang der Dokumente ist verschieden je nach Umfang der systematischen Abstufungen, der Synonyme und der Beschreibung der Kriterien der Art.

Geschichte der Datensammlung

Adolf Stählin begann in Jena eine Samensammlung, als er u.a. Abteilungsleiter der Samen- und Futtermittelkontrolle am Institut für Pflanzenbau der Universität Jena (1927 bis 1945) war. Die zu kontrollierenden Samen- und Fruchtpollen haben sich zum Grundstock einer eigenen Sammlung angeboten. Nach seiner Zeit in Jena hat er von 1946 bis 1956 am Institut für Pflanzenbau und Pflanzenzüchtung der Landwirtschaftlichen Hochschule Hohenheim gearbeitet. In dieser Zeit und danach von 1956 bis 1970 in Gießen als ord. Professor und Direktor des Institutes für Grünlandwirtschaft und Futterbau wie auch nach seiner Emeritierung von 1970 bis zu seinem Tod 1992 hat er die Sammlung stetig erweitert. Sie besteht aus selbstgesammelten Samenproben und Proben von Kollegen und Bekannten, besonders von Botanischen Gärten und Instituten weltweit.

In diesem Zusammenhang wird auf das 1955 von Professor Dr. Walther Brouwer und Professor Dr. Adolf Stählin herausgegebene Handbuch der Samenkunde für Landwirtschaft, Gartenbau und Forstwirtschaft im DLG-Verlag-GMBH, Frankfurt am Main hingewiesen. Dieses Handbuch enthält 2.512 Samenbeschreibungen mit 1.672 Feinstrichzeichnungen der Samen und Früchte. Adolf Stählin hat (besonders nach seiner Emeritierung) über die Literatur eine Basis erarbeitet zur gerichteten Sammlung von Samen weltweit genutzter Pflanzenarten. Dabei hat er über die Literatur wichtige Kriterien über die Arten gesammelt. Diese Arbeit ist von der

¹ ICS Marketing GmbH, Frankfurt

² Bergstr. 18, 35435 Wettenberg

Deutschen Forschungsgemeinschaft unterstützt worden, konnte aber durch seinen Tod nicht mehr vollendet werden.

Adolf Stählin hat sich vor seinem Tod noch Gedanken darüber gemacht, wie das gesammelte Material verwaltet werden kann. Seine Vorstellung ging in Richtung Datenbank, um eine umfassende Befragung nach den verschiedensten Kriterien zu ermöglichen. Eine Datenbank hat er für besonders zweckmäßig erachtet, weil derartige Arbeiten in einer solchen besser aufgehoben sind als in Büchern mit den bald erforderlichen Supplementbänden. Ergänzungen und Verbesserungen sind einfacher zu vollziehen. Auch kann der Index einer Datenbank sehr viel effizienter genutzt werden. Weiterhin kann eine notwendige Aktualisierung sukzessive erfolgen. Erste Gespräche über diese Möglichkeiten hat A. Stählin noch im Beisein seiner Frau und Mitarbeiterin, Dr. Lieselotte Stählin, mit Armin Hohenester geführt, woraus die Zusammenarbeit Letzterer entstanden ist.

Geschichte der Erfassung

Die Erfassung der Daten per Computer startete im Frühjahr 1993. Das zu bearbeitende Material gliedert sich wie folgt:

- Circa 22.000 per Hand beschriebene DIN A4-Blätter geordnet nach Familien, Gattungen und Arten. Für jede Art existiert ein Blatt und gegebenenfalls Folgeblätter. Auf den Blättern sind Angaben zur Art, Synonyme, Angaben zur Nutzung und Herkunft und Kürzel zur Literatur vermerkt.
- Circa 19.000 Samenproben, geordnet nach Familie, Gattung und Art. Die Samenproben sind verpackt in Zellophantüten und etikettiert mit Angabe des Artnamens und der Eingangsnummer.
- Mehrere Eingangsbücher mit fortlaufender Numerierung der Einträge. Eingetragen sind chronologisch die Samenproben mit Artnamen und Herkunft. Diese Angaben bilden die Basis der Etikettierung der Samenproben.
- Circa 18.000 kleine Karteikarten alphabetisch geordnet nach Familie, Gattung und Art. Vermerkt sind Angaben zur Art und die zugehörigen Eingangsnummern der Samenproben.
- Circa 2.000 vollständige Literaturangaben.
- Darüber hinaus existiert noch eine Fülle von Material mit Angaben zu weiteren Arten und Beschreibungen der Familien, welches zunächst von der Bearbeitung ausgenommen ist.

Technisches

Angesichts der Fülle des Materials galt es, zunächst eine sinnvolle Strategie für die Erfassung zu erarbeiten. Nach einigen Versuchen, mit einem Scanner wenigstens die maschinengeschriebenen Karteikarten einzulesen, war klar, daß alle Eingaben nur per Tastatur erfolgen konnten. Diesen Aufwand galt es zu minimieren. Dazu bedurfte es einer sinnvollen Reihenfolge der Erfassung und einer flexiblen Software.

Auswahl der Software für die Erfassung

Die wichtigsten Kriterien bei der Auswahl der Software waren die folgenden:

- Betrieb unter MS-DOS. Reine Tastatursteuerung. Keine Verwendung einer Maus.
- Höchste Flexibilität. Vollständige Kontrolle über alle Funktionen bei der Erfassung.

Die Wahl der Software fiel auf Advanced Revelation. Dies ist eine DOS-Datenbank mit dynamischem Daten-Dictionary und variablen Satz- und Feldlängen. Die Eingabe-Masken sind sehr flexibel und bieten insbesondere viele Möglichkeiten zur Prüfung der eingegebenen Daten für jedes Feld. Für Selektionen und Listen existiert eine Kommandosprache ähnlich SQL. Alle Funktionen können durch eigene Programmierung ergänzt oder ersetzt werden.

Definition und Programmierung der eigenen Datenbank

Nach der Beschaffung von PC und Software wurde sofort mit der Erfassung der Daten begonnen. Eine Systemanalyse wurde zunächst nur in geringem Umfang durchgeführt. Die Flexibilität und die Arbeit in einer 2er Gruppe, erlaubte es, diese bei Bedarf nachzuholen. Die ersten Programme dienten der Erfassung der Angaben zu den Samenproben aus den Eingangsbüchern. Parallel zu der anlaufenden Erfassung wurden bei Bedarf Prüfroutinen sowie weitere Hilfs-Datenbestände zur Erleichterung der Erfassung angelegt. Als Beispiele sind zu nennen eine Datenbank mit den Namen der Familien und eine weitere Datenbank mit den korrekten Bezeichnungen Botanischer Institute über Herkunft der Samenproben. Durch diese Maßnahmen konnte eine einheitliche Schreibweise erreicht, und insbesondere der Aufwand bei der Erfassung erheblich reduziert werden. Zur Erklärung: Innerhalb der Eingabe-Masken können nur „erlaubte“ Familien eingegeben werden. Für jede Samenprobe können nur "erlaubte" Herkunftsbezeichnungen verwendet werden. Bei der sukzessiven Erfassung werden die Eingaben des vorhergegangenen Datensatzes als „Vorschlag“ für den aktuellen Datensatz verwendet und dieser durch einfachen Tastendruck übernommen.

Bei allen folgenden Erweiterungen der Datenbank wurde konsequent auf eine Reduzierung des Aufwandes bei der Erfassung und damit eine Minimierung von Tippfehlern geachtet. Parallel dazu wurden ständig weitere Hilfs-Datenbestände zur Überprüfung der Eingaben angelegt.

Insgesamt besteht die Datenbank derzeit aus sechs Dateien für die Artenblätter und 18 Dateien mit den Hilfs-Datenbeständen. Diese Dateien sind vielfältig miteinander verknüpft. Das Spektrum der selbst erstellten Programme ist vielfältig: einfache Prüfroutinen, Listen zur Kontrolle der Erfassung bis zu komplexen Programmen zur Ermittlung der Synonyme einer Art aus der Kew-Datenbank (siehe unten). Insgesamt wurden circa 120 Programme erstellt. Die bisher erfaßten Daten belegen circa 100 MB auf der Festplatte. Die Kew-Datenbank belegt weitere circa 300 MB.

Besonderes

Zur Überprüfung der Artnamen mit Hilfe des Index Kewensis wurde dieser als „Index Kewensis on Compact Disc, 1993“ angeschafft. Für die Arbeit mit dem Index Kewensis wurde der komplette Inhalt der CD in eine eigene Datenbank transferiert und neu indiziert.

Dadurch konnte er mit den vorliegenden Programmen korrespondieren. Es war von vornherein klar, daß die Artnamen der Einsendungen und die in der Literatur gefundenen überprüft werden mußten. Dazu bot sich der Index Kewensis an als das umfassendste Werk von hohem Rang. Die Benutzung dieses Werkes geschah allerdings in dem Bewußtsein, den Stand der Forschung dieses Werkes zu akzeptieren. Es wäre wünschenswert, dieses auf einen international allgemein anerkannten Stand zu bringen. Dieser Wunsch wird z.B. bei *Triticum*-Arten wie auch bei *Brassica*-Arten deutlich.

Als Ergebnis waren einerseits Umbenennungen erforderlich, andererseits führte es zur Zusammenführungen verschiedener Samenproben.

Rechtschreibung und Suchbegriffe

Zur Vermeidung von Tippfehlern und zur Vereinheitlichung von Fachbegriffen wurde eine eigene Datenbank erstellt. Die wichtigsten Grundlagen bei der Ermittlung der korrekten Schreibweise sind:

- Duden - Rechtschreibung der deutschen Sprache, 21. Auflage 1996
- Brockhaus Kompaktwissen von A bis Z: in 5 Bd. 1983
- Der Fischer Weltalmanach 1998
- Duden - Das Wörterbuch medizinischer Fachausdrücke, 1992

- Botanisches Wörterbuch, 11. Auflage, Ulmer, 1993
- Internationaler Atlas, Kümmerley + Frey, Rand Mc Nally Westermann 1981

Diese Datenbank leistet im einzelnen folgendes:

- Die Überprüfung sämtlicher in den Artenblättern vorkommender Worte auf korrekte Schreibweise.
- Die individuelle Auswahl von Worten die zu Suchbegriffen bestimmt werden.
 - a) Bei den Suchbegriffen gibt es Basisworte, die in den Artenblättern vorkommen (z.B. Myanmar).
 - b) Weitere alternative, frühere, internationale oder veraltete Bezeichnungen, die nicht im Text der Artenblätter vorkommen dürfen, aber als Suchbegriffe dienen (z.B. Birma, Burma), und zum Basiswort führen (Myanmar).
 - c) Zur eigenen Orientierung können diese Suchbegriffe kommentiert werden (z.B. Staat in SO-Asien).
 - d) Zur eigenen Orientierung kann jeder Suchbegriff nach Fachgebieten klassifiziert werden (Staat/Region).
 - e) Über die Suchbegriffe, die einen Index darstellen, besteht die Suchmöglichkeit nach interessierenden Kriterien der einzelnen Arten.

Datenübergabe an ZADI

Zur Übergabe der Daten an ZADI wurde ein Programm entwickelt, welches aus den Dateien der eigenen Datenbank die Artenblätter zusammensetzt. Bei der Ausgabe werden dann die Datensätze (=Artenblätter) und deren einzelne Felder (=Inhalte) durch bestimmte vereinbarte Zeichen getrennt. Die zu übergebenden Artenblätter werden aus der eigenen Datenbank selektiert und mittels dieses Programmes in eine Textdatei geschrieben. Die so erzeugten Dateien sind je nach Anzahl und Umfang der Artenblätter mitunter recht groß, und können 20 und mehr MB Speicherplatz erfordern. Sie werden per FTP (File Transfer Protocol) an die Zentralstelle für Agrardokumentation und -information (ZADI) übermittelt und als Basis-Datenbank online im Internet zugänglich gemacht (<http://www.dainet.de/genres/staehlin>).

Datenübergabe für eine eigene CD-Rom

Zur Zeit wird eine Software zur Offline-Nutzung der STAEHLIN-DB entwickelt. Die Artenblätter und ein Programm zum Suchen der Artenblätter sollen auf einer CD-ROM verfügbar gemacht werden. Das Abfrage-Programm läuft unter Windows. Die Präsentation der Artenblätter erfolgt über einen beim Benutzer vorhandenen Browser. Die Datenbank basiert auf BTRIEVE. Die Artenblätter sind im HTML-Format in dieser Datenbank gespeichert. Die wichtigsten Gründe für die Wahl von BTRIEVE sind die kompakte Speicherung von Datensätzen variabler Länge und die Geschwindigkeit beim Lesen von Datensätzen. Nicht minder wichtig bei der Entscheidung war jedoch, daß nach dem Erwerb keine weiteren Lizenzgebühren bei der Verteilung anfallen.

Samensammlung

Das Ergebnis der Zusammenführung der Namen ist in den 4 Bänden der Samensammlung enthalten. Im 5. Band existiert ein Index aller gefundenen Namen. Geordnet sind die Arten jeweils alphabetisch nach Familien, Gattungen und Arten. Auch die Samentütchen sind, nachdem sie die neue Namensnennung erhalten haben, den 4 Bänden entsprechend geordnet. Dazu wurden alle Samentütchen erneuert, mit den endgültigen Namen etikettiert und die Samen umgefüllt. Diese Etiketten tragen auch die Eingangsnummer, damit der ursprüngliche Name gefunden werden kann. Der ursprüngliche Name mit Angabe des Herkunftsortes ist in den Bänden jeweils genannt. Der geprüfte Name ist mit "Kew" gekennzeichnet. Dadurch wurde die Vergabe von 2 Nummern notwendig, die Eg-Nr. = Eingangsnummer und die Artennummer, die in der gesamten Arbeit verwendet wird.

Die komplette Samensammlung mit den 5 Bänden ist, dem Wunsch von A. Stählin entsprechend, in Gatersleben dem Institut für Pflanzengenetik und Kulturpflanzenforschung zu Händen von Herrn Dozenten Dr. habil. Peter Hanelt bereits übergeben worden.

Inhalt der Artenblätter

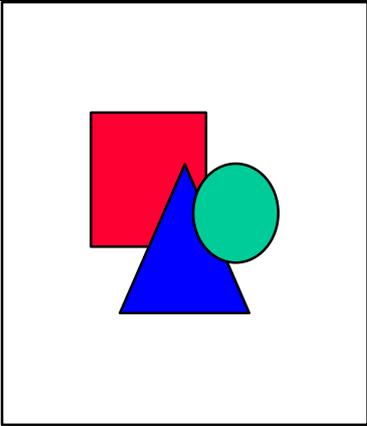
Die Artenblätter sind wie folgt strukturiert:

- **Artnamen nach Index Kewensis**
- **Artnamen nach ISTA 1988**
- **Familienname**
- **Systematische Abstufungen nach Index Kewensis:** Mit einem aufwendigen Programm sind die Systematischen Abstufungen wie ssp., var., subvar., f. ... aus dem Index Kewensis gewonnen worden und überprüft auf ihre Zugehörigkeit zu der jeweiligen Art

- **Synonyme:** Mit einem aufwendigen Programm sind die Synonyme aus dem Index Kewensis gewonnen worden und überprüft auf ihre Zugehörigkeit zu der jeweiligen Art
- **Eingangsdaten der Samenproben:** Eingangsnummer, Artname, Herkunft
- **Volksnamen:** Namen in den verschiedenen Ländern und Regionen
- **Wuchsform**
- **Frucht:** Beschreibung der Frucht nach vorgegebenen Kriterien und eventuellen Ergänzungen
- **Same:** Beschreibung des Samens nach vorgegebenen Kriterien und eventuellen Ergänzungen
- **Heimat, Verbreitung, in Kultur:** Nennung der Länder und Gebiete, eingeschleppt, eingebürgert, besondere Charakteristika der Fundorte, ob in Kultur im natürlichen Gebiet oder eingeführt, Aufzucht durch Samen oder Pflanzenteile
- **Nutzung:** Nahrung, Getränke, Genuß, Gewürz, Arznei, Droge, Kosmetik, Ritual und Zauber, Schmuck, Farbstoff, Gift, Schädlingsbekämpfung, Gewerbe, Forstpflanzen, Futter, Spezialanbau, Rasen, Zierpflanzen, Sonstiges

Bemerkungen zu den Artenblättern

Über den Artnamen direkt oder über seine Synonyme wird in den meisten Fällen die Möglichkeit bestehen, die gewünschte Art zu finden. Einschränkend muß gesagt werden, daß vorwiegend die Synonyme erster Ordnung, d.h. die Synonyme, die direkt auf die nach dem Index Kewensis benannte Basisart hinweisen, vorhanden sind. Synonyme höherer Ordnung sind nur dann genannt worden, wenn sie sich bei der Bearbeitung der Namen der Einsendung sowie der Literatur ergeben haben. Das Programm für die Synonyme erster Ordnung ist vorhanden, für die Synonyme höherer Ordnung wurde aufgrund erheblicher Schwierigkeiten noch kein Programm entwickelt. Diese Einschränkung ist zu vertreten, weil die meisten Fälle mit den Synonymen erster Ordnung zu lösen sind.

Artblatt:	9989
Artname nach Index Kewensis:	Phleum pratense Linn. Sp. Pl. 59. Reg. bor. temp.
Artname nach ISTA 1988:	Gramineae Phleum pratense L.
Systematische Abstufungen der Art nach Index Kewensis (Auswahl - für die gesamten Abstufungen klicken Sie hier):	Gramineae Phleum pratense L. subsp. abbreviatum (Boiss.) J. Molero Mesa & F. Perez Raya, (Fl. Sierra Nevada: 350 (1987):) PHLEUM pratense var. abbreviatum.; Gramineae Phleum pratense L. subsp. bertolonii (DC.) Serbanescu & E.L Nyarady (in Fl. Republ. Social. Roman., 12: 123 (1972):) PHLEUM bertolonii.; Gramineae Phleum pratense L. subsp. brachystachyum (Salis) Gamisans (in Candollea, 28(1): 48 (1973)) PHLEUM pratense var. brachystachyum.;
Synonyme nach Index Kewensis (Auswahl - für die gesamten Synonyme klicken Sie hier):	Gramineae Phleum arvense Pourr. (in Mem. Acad. Toul. iii. (1788) 324) =pratense. Gramineae Phleum bertolonii DC. (Cat. Hort. Monsp. 132) =pratense. Gramineae Phleum bulbosum Gouan, (Fl. Monsp. 117) =pratense. Gramineae Phleum ciliatum Gilib. (Exercit. ii. 518) =pratense. Gramineae Phleum deckeri Roem. ex Trin. (in Mem. Acad. Petersb. Ser. VI. v. II. (1840) 124) =pratense. Gramineae Phleum fallax Janka, (Koezlem. xii. (1876) 168; ex Just, Jahresb. iv. III. 1061) =pratense. Gramineae Phleum intermedium Jord. (in F. Schultz, Arch. Fl. Fr et Allem. 325) =pratense.
Volksname:	Deutsch Wiesen-Lieschgras, Timothe, Kaminkehrer; Katzenschweif; Englisch common timothy, cat's tail; Französisch fiéole des près, timothée, maselle; ivraie du vexin, fenasse à maillet; Italienisch coda di topo, flèolo dei prati, codolina; Spanisch fleo (pratense), cola de topo, topo timote; Portugiesisch timótio (capim), rabo de rato, rabo de gato.
Wuchsform:	Staude
	Fruchtart: Karyopse siehe Samen-Beschreibung Frucht: Ährchen (Spicula) in Büscheln spreizend, ohne Granne (Arista) (2-) 3 - 3.8 (-5.5) mm lang, länglich, abgeflacht, 1blütig, bei Reife über den Hüllspelzen (Glumae) zerbrechend; Hüllspelzen schmal-länglich, gestutzt, gekielt, an den häutigen Seiten sehr fein rauh, 3nervig, am Kiel mit steifen, weißen abstehenden Haaren, an der Spitze ausgezogen in 1/4 - 1/2 so lange, (0.2 -) 1 - 2.5 mm lange, steife, Granne (Arista), die untere Hüllspelze (Gluma inferior) an den Rändern weich behaart. Spelzfrüchte (Karyopse mit Deckspelze (Palea inferior) und Vorrspelze (Palea superior) 3/4 so lang wie Hüllspelzen, 1.5 - 2 x± 1 x ± 1 mm, dick-eiförmig, silbergrau, weißlich bis bräunlich, matt; Deckspelze breit, sehr stumpf, mit 5 (- 7) Nerven, kahl bis sehr fein behaart, häutig; Vorspelze etwas kürzer bis so lang wie Deckspelze, mit 2 deutlichen

	<p>Nerven, zart, durchscheinend.</p> <p>Samenform: Karyopse kugelig bis eiförmig und eilänglich, stielrund, zugespitzt, gelbbraun, runzelig, matt;</p> <p>Länge/Breite/Dicke in mm: 1.5 - 2 / &plusmn; 1 / &plusmn; 1</p> <p>Samen-TKG in g: 0.35 - 0.5</p> <p>Samen-Sonst.: Karyopsen von den Spelzen anliegend, aber lose umhüllt, leicht aus ihnen herausfallend.</p>
Geogr. Verbreitung:	Europa außer arktisches Rußland, Algerien, N-Asien; Grasland, offene Waldplätze, Grabenränder (0 - 1800 m); eingebürgert in Nordamerika, Indien; aus dem Anbau oft verwildert; kultiviert Areal, Australien, Neuseeland, tropisches Amerika, S-Afrika;
Nutzung:	<p>Futter: Sproß, Frucht Mäh- und Weidegras, Früchte geschrotet als Viehfutter.</p> <p>Spezialanbau: Ganze Pflanze. Bodenschutzpflanze im Westen von U.S.A.</p> <p>Rasen: Ganze Pflanze Rasengras, hier Phleum pratense Linn var. nodosum.Huds.</p>
Link auf Bibliographie:	Liste der bibliographischen Referenzen

Abb. 1: Artenblatt aus der online-Datenbank

Fig. 1: Species information sheet from the online database

Umfang der Datenbank

Die Stählin-Datenbank enthält zur Zeit 18.199 Arten aus 314 Familien der Gymnospermae und Angiospermae. Dazu kommen 94.250 Synonyme und 12.120 systematische Abstufungen. Die Anzahl der Arten mit Samenproben ist mit 10.540 kleiner, weil nicht zu allen ausgewählten Arten Samenproben zu erhalten gewesen sind. Die Anzahl aller Samenproben beträgt 18.991. Wie bereits erwähnt, ist es vorgesehen, den Umfang der Datenbank weiter zu erhöhen, dafür liegt weiteres Material vor. In der ersten Stufe der Erfassung wurden bereits die meisten der weltwirtschaftlich wichtigen Arten berücksichtigt. Die zweite Stufe sieht nun die Erfassung weiterer Arten vor.

Bisher verfügbare Daten

Für sämtliche Arten sind folgende Kriterien eingegeben und aufsuchbar:

- Artname nach Index Kewensis
- Artname nach ISTA
- Familienname
- Systematische Abstufungen nach Index Kewensis

- Synonyme nach Index Kewensis
- Eingangsdaten der Samenproben

Alle weiteren Kriterien wurden bereits für die folgenden Familien eingegeben und sind recherchierbar: Acanthaceae, Aceraceae, Achatocarpaceae, Aextoxicaceae, Agavaceae, Aizoaceae, Alangiaceae, Amaranthaceae, Amaryllidaceae.

Für einzelne Arten der folgenden Familien wurden alle verfügbaren Kriterien eingegeben: Annonaceae, Chenopodiaceae, Compositae, Cucurbitaceae, Gramineae, Liliaceae.

Die Arbeit wird laufend fortgesetzt, so daß sukzessive die gesamte Arbeit von A. Stählin verfügbar sein wird. Die Arbeit ist so angelegt, daß sie laufend ergänzt und erweitert werden kann. Zur Zeit ist es nicht möglich, die Samensammlung zu erweitern, weil die Suche nach Arten in den Katalogen von Botanischen Gärten oder Instituten sehr viel Zeit in Anspruch nimmt. Diese kann erst wieder aufgenommen werden, wenn die gesamte Arbeit verfügbar ist. An sich handelt es sich um eine Arbeit, die nie abgeschlossen sein wird. Einmal weil nur ein Teil der möglichen Kriterien ausgewählt wurde und beliebig ergänzt werden kann, zum anderen kommen laufend neue Erkenntnisse hinzu, die eingearbeitet werden sollten. Schließlich wäre es wünschenswert, den Umfang der Arbeit durch Kritik, Ergänzungen und Hinweise zu bereichern.

Danksagungen

Es ist uns ein Anliegen, unseren Dank an alle auszusprechen, die zur Verfügbarmachung der von Adolf Stählin hinterlassenen Arbeit beigetragen haben. Zu ihnen gehören in besonderer Weise Herr Professor Dr. Dr. h. c. mult. Dieter Bommer, der den Kontakt mit dem Institut für Pflanzengenetik und Kulturpflanzenforschung in Gatersleben hergestellt hat, und Herr Dozent Dr. habil. Peter Hanelt für seine Bereitschaft, die Samensammlung zu übernehmen. Ferner verdanken wir Herrn Professor Bommer den Kontakt mit der Zentralstelle für Agrardokumentation und -information (ZADI). Herrn Dr. Mangstl, Herrn Dr. Pohlmann, Herrn Dr. Begemann und Herrn J.D. Jiménez Krause verdanken wir, daß die Datenbank im Internet veröffentlicht und laufend ergänzt wird.

Aufbau einer WWW-Datenbank zu „Mansfeld's World Manual of Agricultural and Horticultural Crops“

Establishment of an online database of „Mansfeld's World Manual of Agricultural and Horticultural Crops“

JÖRG OCHSMANN¹, NORBERT BIERMANN², HELMUT KNÜPFER² UND KONRAD BACHMANN¹

Einleitung

Die zweite Auflage des Mansfeld-Verzeichnisses (SCHULTZE-MOTEL 1986), das eine Standardquelle für Informationen über Kulturpflanzen ist, behandelt ca. 4.800 Arten (ohne Zier- und Forstpflanzen). Inzwischen steht die dritte, erweiterte Auflage des Buches kurz vor der Fertigstellung. Die Neuauflage (HANELT, in Vorb.), an der rund 20 Bearbeiter mitwirken und die in englischer Sprache erscheinen wird, umfaßt rund 6.000 Kulturpflanzenarten sowie zahlreiche infraspezifische Taxa.

Parallel zur Buchausgabe wird eine im Internet verfügbare Datenbankversion erstellt. Hierzu werden die in elektronischer Form als Textdateien vorliegenden Manuskripte strukturiert und in Tabellen überführt, so daß sie in eine Datenbank übernommen werden können. Diese Arbeiten sind Teil des BIG-Projektes (**B**undes**I**nformationssystem **G**enetische Ressourcen), das durch das BMBF gefördert wird (s. Beitrag BECKER in diesem Band). Dabei wird auf die Erfahrungen des IPK beim Aufbau von kulturpflanzen-taxonomischen Datenbanken für ausgewählte Länder zurückgegriffen (KNÜPFER & HAMMER, im Druck).

Ziele

Das Hauptziel des Projektes ist die Bereitstellung der Informationen aus "Mansfeld's World Manual of Agricultural and Horticultural Crops" in einer über das Internet verfügbaren Datenbank. Dadurch ergibt sich einerseits die Möglichkeit zur gezielten Suche nach bestimmten Informationen, andererseits können die vorhandenen Daten kontinuierlich aktualisiert oder ergänzt werden.

Bei der Datenbankentwicklung wird besonderer Wert auf die Berücksichtigung kulturpflanzen-spezifischer Probleme gelegt. Während in der Taxonomie der Wildpflanzen heute üblicherweise davon ausgegangen wird, daß eine Hierarchie den verwandtschaftlichen Beziehungen ent-

¹ Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), AG Experimentelle Taxonomie

² Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), AG Genbank-Dokumentation

spricht, sind in der Kulturpflanzentaxonomie künstliche, auf praktische Überlegungen gestützte Systeme verbreitet. Hinzu kommt, daß für Kulturpflanzen neben dem Botanischen Code (ICBN, GREUTER et al. 1994) auch der Kulturpflanzencode (ICNCP, TREHANE et al. 1995) gilt, der in Anlehnung an das Culton-Konzept (HETTERSCHEID 1994, s.a. HETTER-SCHEID UND BRANDENBURG 1995) für Cultivars (Kultivare) auch nichthierarchische Beziehungen erlaubt. Ein Cultivar kann beispielsweise gleichzeitig mehreren Cultivar-groups (Kultivargruppen) angehören. Das bedeutet, daß eine Datenbank sowohl komplexe Hierarchien botanischer Namen als auch nicht vollständig hierarchisch aufgebaute Cultivar-Namen abbilden muß.

Weiterhin unterstützt das Projekt die Erstellung der druckfertigen Vorlage für die Buchversion unter Ausnutzung der technischen Möglichkeiten einer Datenbank. Als Beispiele sind die Erstellung verschiedener Register (wissenschaftliche Namen, Volksnamen, Nutzungsrichtungen) oder des Literaturverzeichnisses sowie Vereinheitlichungen von Schreibweisen (Ländernamen, Abkürzungen etc.) zu nennen.

Zusätzlich sollen im Rahmen des BIG-Projektes durch die Kombination mit anderen Datenbanken (u.a. Akzessionsdaten der Genbank des IPK, Verbreitungsdaten des Bundesamtes für Naturschutz (BfN), Akzessionsdaten der Zentralstelle für Agrardokumentation und -information (ZADI) und Daten aus SYSTAX, (vgl. HOPPE 1999)) weitere Informationen zugänglich gemacht und miteinander kombiniert werden.

Vorgehensweise

Die Arbeiten lassen sich in verschiedene Abschnitte gliedern. Am Anfang steht eine Überarbeitung der in elektronischer Form vorliegenden Rohmanuskripte der einzelnen Bearbeiter. Neben formalen Korrekturen und der Vereinheitlichung von Abkürzungen, die teilweise mit Hilfe interaktiver WordPerfect-Makros durchgeführt werden, erfolgt eine Strukturierung der Manuskripte durch Zuweisung von Formatvorlagen. Dabei beinhaltet jede Formatvorlage eine bestimmte Kategorie von Informationen:

- akzeptierter Name - wissenschaftliche Namen mit Autorenkürzel und Protolog
- Synonyme (optional) - Aufzählung der wissenschaftlichen Namen mit Autorenkürzel und Protolog
- taxonomische Anmerkungen (optional) - Angaben über alternative systematische Auffassungen etc.
- Typus (nur bei Gattungen)
- Volksnamen (optional) - Aufzählung der Volksnamen, soweit bekannt mit vorgestelltem Sprachkürzel, sonst mit Länder- oder Regionsangabe
- Wildareal (optional) - Angaben zur Verbreitung in der freien Natur
- weitere Informationen - Angaben zum Kulturareal, zu Nutzungen, Inhaltsstoffen, Resistenzen etc., zur Geschichte sowie zu weiteren, potentiell nutzbaren Sippen

- Referenzen

Jede Informationseinheit (im Sinne eines „potential taxon“, vgl. BERENDSOHN 1995) beginnt mit einem akzeptierten Namen (Kategorie 1) und enthält, abhängig vom Rang, eine unterschiedliche Anzahl von Informationen. Die größte Informationsdichte liegt bei Spezies und, falls vorhanden, bei infraspezifischen Taxa. Die Beschreibung von Familie, Gattung und Art ist obligatorisch. Die Angabe von Zwischenrangstufen wird im Einzelfall vorgenommen. Den drei Hauptrangstufen entsprechend wurden drei Sätze von Formatvorlagen formuliert. Jedem Absatz wird eine Formatvorlage zugewiesen. Bei dem o.g. Beispiel erhalten etwa die Synonyme von *Momordica* die Formatvorlage "Genus_Synonym". Alle Zwischenrangstufen werden wie die Hauptränge Familie, Gattung, Art behandelt.

Tab. 1: Beispiel für den Aufbau zweier Artikel für die Buchausgabe des „Mansfeld's World Manual of Agricultural and Horticultural Crops"

Tab. 1: Example of the structure of two articles in the printed version of "Mansfeld's World Manual of Agricultural and Horticultural Crops"

Kategorie	Text
1	Momordica L., Sp. Pl. (1753) 1009, Gen. Pl. ed. 5 (1754) 440.
2	<i>Zucca</i> Commers. ex Juss., Gen. (1789) 398; <i>Muricia</i> Lour., Fl. Cochinch. 2 (1790) 596; <i>Nevrosperma</i> Raf. in Amer. Monthly Mag. 4 (1815) 40.
4	Type: <i>Momordica balsamina</i> L.
1	Momordica cochinchinensis (Lour.) Spreng., Syst. Veg. 3 (1826) 14.
2	<i>Momordica trifolia</i> L., Herb. Amboin. (1754) 24 et Syst. Nat. ed. 10 (1758) 1278, et orth. mut.; <i>M. trifoliata</i> L., Sp. Pl. ed. 2 (1763) 1434, "trifoliata", orth. var., nom. rejic. propos.; <i>Muricia cochinchinensis</i> Lour., Fl. Cochinch. 2 (1790) 596; <i>Momordica mixta</i> Roxb., Hort. Beng. (1814) 70, nom. illeg.; <i>Zucca commersoniana</i> Ser. in DC., Prodr. 3 (1828) 315; <i>Momordica ovata</i> Cogn. in A. DC., Monogr. Phan. 3 (1881) 446; <i>M. macrophylla</i> Gage in Rec. Bot. Surv. India 3 (1908) 61; <i>M. meloniflora</i> Hand.-Mazz. in Anz. Akad. Wiss. Wien, Math.-Nat. 58 (1921) 94.
5	Spiny bitter cucumber; Chin. mu bie zi; Hindi gulkakra; teruach (Malaya); mak-kaao (Indochina); Jap. makubet sushi; kakkur (India).
6	E and S tropical Asia from Japan and India to Indonesia and New Guinea.
7	Locally cultivated (e.g., in S China). The unripe fruits are used as vegetables and in curries. The seeds yield a low-quality oil and have many medicinal uses in China, where the roots and stems are also used medicinally. The roots are rich in saponins and are also used in laundering.
8	Ref.: Herklots 1972, 525 pp.; Wealth of India 6, 1962; uravlev et al. 1992, 125.

Tab. 2: Generierte Tabelle zum vorgestellten Beispiel (s. Tab. 1)

Tab. 2: Generated table according to the previous example (see table 1)

Text	Kategorie	Art-ID	Node-ID	Inf-ID
Momordica L., Sp. Pl. (1753) 1009, Gen. Pl. ed. 5 (1754) 440	Genus_Name	216.2	216.1	
<i>Zucca</i> Commers. ex Juss., Gen. (1789) 398	Genus_Synonym			216.2
<i>Muricia</i> Lour., Fl. Cochinch. 2 (1790) 596	Genus_Synonym			216.2
<i>Nevrosperma</i> Raf. in Amer. Monthly Mag. 4 (1815) 40	Genus_Synonym			216.2
Type: <i>Momordica balsamina</i> L.	Genus_Typ			216.2
Momordica cochinchinensis (Lour.) Spreng., Syst. Veg. 3 (1826) 14	Species_Name	216.3	216.2	
<i>Momordica trifolia</i> L., Syst. Nat. ed. 10 (1758) 1278, nom. rejic. propos.	Species_Synonym			216.3
<i>M. trifoliata</i> L., Sp. Pl. ed. 2 (1763) 1434, orth. var.	Species_Synonym			216.3
<i>Muricia cochinchinensis</i> Lour., Fl. Cochinch. 2 (1790) 596	Species_Synonym			216.3
<i>Momordica mixta</i> Roxb., Hort. Beng. (1814) 70, nom. illeg.	Species_Synonym			216.3
<i>Zucca commersoniana</i> Ser. in DC., Prodr. 3 (1828) 315	Species_Synonym			216.3
<i>Momordica ovata</i> Cogn. in A. DC., Monogr. Phan. 3 (1881) 446	Species_Synonym			216.3
<i>M. macrophylla</i> Gage in Rec. Bot. Surv. India 3 (1908) 61	Species_Synonym			216.3
<i>M. meloniflora</i> Hand.-Mazz. in Anz. Akad. Wiss. Wien, Math.-Nat. 58 (1921) 94	Species_Synonym			216.3
Spiny bitter cucumber	Species_Volksn.			216.3
Chin. mu bie zi	Species_Volksn.			216.3
Hindi gulkakra	Species_Volksn.			216.3
teruach (Malaya)	Species_Volksn.			216.3
makkao (Indochina)	Species_Volksn.			216.3
Jap. makubet sushi	Species_Volksn.			216.3
kakkur (India)	Species_Volksn.			216.3
E and S tropical Asia from Japan and India to Indonesia and New Guinea.	Species_Wildareal			216.3
Locally cultivated (e.g., in S China). The unripe fruits are used as vegetables and in curries. The seeds yield a low-quality oil and have many medicinal uses in China, where the roots and stems are also used medicinally. The roots are rich in saponins and are also used in laundering.	Species_Info			216.3
Herklots 1972, 525 pp.	Species_Ref			216.3
Wealth of India 6, 1962	Species_Ref			216.3

URAVLEV et al. 1992, 125	Species_Ref			216.3
--------------------------	-------------	--	--	-------

Bei der automatischen Umsetzung der Texte in Tabellen wird die genaue Rangstufe durch Analyse des Präfix automatisch erkannt. Aufzählungen (Synonyme, Volksnamen, Referenzen) werden in Einzelinformationen aufgetrennt.

In diesem Stadium ist jeder Absatz des Textes automatisch identifizierbar hinsichtlich seines Informationsgehaltes. Beispielsweise ist „*Zucca* Commers. ex Juss., Gen. (1789) 398“ als Synonym eines Genus oder eines infragenerischen Taxons erkennbar. Die Tatsache, daß es sich um ein Synonym zu *Momordica* handelt, ist in der Buchform leicht zu ersehen; in einer Datenbank wird die Zuordnung mit Schlüsselnummern (IDs) vorgenommen. Bei der Umsetzung der Texte in Tabellen müssen diese Schlüsselnummern automatisch vergeben werden können. Hierfür wurde ein spezielles Makro entwickelt, welches die Umsetzung in eine Tabelle noch innerhalb des Textverarbeitungsprogramms vornehmen kann. Dadurch wird eine sehr flexible Bearbeitung ermöglicht: Wenn die Manuskripte verändert werden, kann die Umsetzung schnell aktualisiert werden. Da die Umsetzung parallel zur Erstellung der Manuskripte erfolgt, ist dies sehr häufig der Fall. Ein weiterer Vorteil ist, daß man noch nicht auf eine Datenbanksoftware festgelegt ist. Die entstehenden Tabellen können in jede relationale Datenbank importiert werden.

Mit der Schlüsselnummer für den Artikel (Art_ID) wird jede Informationseinheit durchnummeriert. Jede Einzelinformation (Inf_ID) bezieht sich hierauf. Gleichzeitig wird in der Node_ID das nächsthöhere Taxon zu jedem Artikel kodiert. So läßt sich eine beliebige hierarchische Struktur geeignet abbilden. Gerade bei den Kulturpflanzen ist oft eine komplexe Hierarchie anzutreffen, die sich in einer starren tabellarischen Einteilung nur schwer abbilden läßt. Cultivar-groups und Cultivare lassen sich nicht auf bestimmte Rangstufen festlegen, sind aber mit einer übergeordneten Anbindung genau beschrieben.

Das Beispiel zeigt, daß jede Einzelinformation einem Artikel (bzw. Taxon, akzept. Namen) zugeordnet werden kann. Damit ist die logische Voraussetzung für die Erstellung einer relationalen Datenbankstruktur gegeben. Diese könnte im einfachsten Fall eine verteilte Struktur mit zwei Tabellen und einer 1:n-Beziehung sein (Tab. 3).

Tab. 3: Verteilte Struktur der Datenbank (Erläuterungen im Text)

Tab. 3: Relational structure of the database

Art_ID	Text	Kategorie	Node_ID
216.2	Momordica L., Sp. Pl. (1753) 1009, Gen. Pl. ed. 5 (1754) 440	Genus_Name	216.1
216.3	<i>Momordica cochinchinensis</i> (Lour.) Spreng., Syst. Veg. 3 (1826) 14	Species_Name	216.2

Inf_ID	Text	Kategorie
216.2	<i>Zucca</i> Commers. ex Juss., Gen. (1789) 398	Genus_Synonym
216.2	<i>Muricia</i> Lour., Fl. Cochinch. 2 (1790) 596	Genus_Synonym
216.2	<i>Nevrosperma</i> Raf. in Amer. Monthly Mag. 4 (1815) 40	Genus_Synonym
216.2	Type: <i>Momordica balsamina</i> L.	Genus_Typ
216.3	<i>Momordica trifolia</i> L., Syst. Nat. ed. 10 (1758) 1278, nom. rejic. propos.	Species_Synonym
...
216.3	Spiny bitter cucumber	Species_Volksn.
...
216.3	E and S tropical Asia from Japan and India to Indonesia and New Guinea.	Species_Wildareal
216.3	Locally cultivated (e.g., in S China). The unripe fruits are used as vegetables and in curries. The seeds yield a low-quality oil and have many medicinal uses in China, where the roots and stems are also used medicinally. The roots are rich in saponins and are also used in laundering.	Species_Info
216.3	Herklots 1972, 525 pp.	Species_Ref
...

Ausblick

Die Datenbank zu "Mansfeld's World Manual of Agricultural and Horticultural Crops" wird parallel zur Buchausgabe erscheinen, beide Medien werden sich gegenseitig ergänzen. Die Verfügbarkeit von taxonomischen, nomenklatorischen, Verbreitungs- und Nutzungsinformationen in frei recherchierbarer Form im Rahmen des BIG-Projektes wird dazu beitragen, daß Nutzer sich umfassend über vielfältige Aspekte der jeweiligen Kulturpflanzenarten informieren können.

Literatur

- BECKER, B. (1999): Bundesinformationssystem Genetische Ressourcen – Konzept und erste Ergebnisse. In: F. BEGEMANN, S. HARRER UND J.D. KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. Schriften zu Genetischen Ressourcen, Bd 12, ZADI, Bonn.
- BERENDSOHN, W.G. (1995): The concept of "potential taxa" in databases. - *Taxon* 44: 207-212.
- GREUTER, W., F.R. BARRIE, H.M. BURDET, W.G. CHALONER, V. DEMOULIN, D.L. HAWKSWORTH, P.M. JØRGENSEN D.H. NICOLSON, P.C. SILVA, P. TREHANE AND J. McNEILL (eds.) (1994): International code of botanical nomenclature (Tokyo Code) adopted by the 15th International Botanical Congress, Yokohama, 8-9/1993. - *Regnum Veget.* 131: 389 S..
- HETTERSCHEID, W.L.A. (1994): The culton concept: recent developments in the systematics of cultivated plants. - *Acta Bot. Neerl.* 43: 78 [abstract].
- HETTERSCHEID, W.L.A. AND W.A. BRANDENBURG (1995): Culton versus taxon: conceptual issues in cultivated plant systematics. - *Taxon* 44: 161-175.
- HOPPE, J.R., E. BOOS UND TH. STÜTZEL (1999): SysTax – ein Datenbanksystem für Systematik und Taxonomie. In: F. BEGEMANN, S. HARRER UND J.D. KRAUSE (Hrsg.): Schriften zu Genetischen Ressourcen, Bd 12, ZADI, Bonn.
- KNÜPFER, H. AND HAMMER, K. (im Druck): Agricultural biodiversity: a database for checklists of cultivated plant species. in: Proc. Third International Symposium on the Taxonomy of Cultivated Plants. Edinburgh.
- SCHULTZE-MOTEL, J. (Hrsg.) 1986: Rudolf Mansfelds Verzeichnis Landwirtschaftlicher und Gärtnerischer Kulturpflanzen (ohne Zierpflanzen). (2. Neubearb. Aufl.) 4 Bde., 1998 S..
- TREHANE, P., C.D BRICKELL, B.R. BAUM, W.L.A. HETTERSCHEID, A.C. LESLIE, J. McNEILL, S.A. SPONGBERG and F. VRUGTMAN (Hrsg.) (1995): International Code of Nomenclature for Cultivated Plants - 1995. - *Regnum Veget.* 133: 175 S..

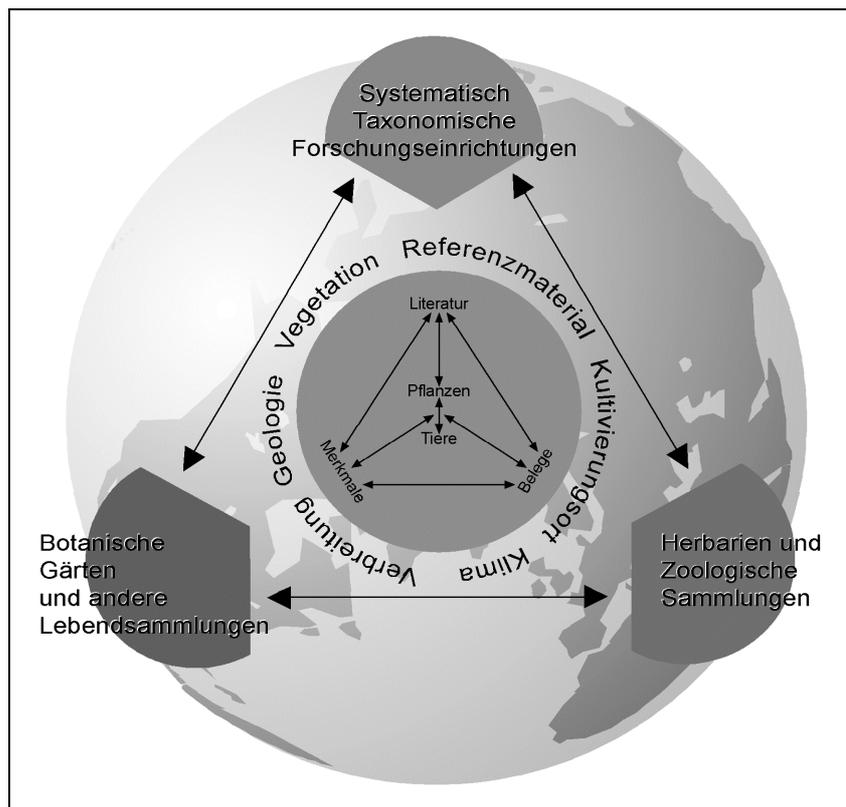
SysTax - ein Datenbanksystem für Systematik und Taxonomie

SysTax – a database system for systematics and taxonomy

JÜRGEN R. HOPPE¹, EVELIN BOOS^{1,2} UND THOMAS STÜTZEL²

Was wird gebraucht? - Der Grundgedanke

Informationen wurden und werden in der Wissenschaft nur höchst ineffektiv der Allgemeinheit verfügbar gemacht. Die sogenannten Publikationen, treffend auch als „paper“ bezeichnet, dienen mehr dem viel zu langsamen Informationsaustausch unter denen, die sich genau mit dem betreffenden Wissenschaftszweig beschäftigen. Ineffektiv deshalb, weil die Spanne zwischen Einreichen eines Manuskriptes und dem eigentlichen Veröffentlichungsdatum äußerst lang sein kann, und weil aus Kostengründen Zeitschriften nicht überall ausreichend verfügbar sein können. Dies trifft besonders zu auf dem Gebiet der Systematik und Taxonomie, der sog. „organismischen“ Wissenschaften. Mit klaren Regeln für die Nomenklatur wird hier in einer Art Gemeinschaftsarbeit der



Wissenschaftler das oder besser ein System der Organismen erarbeitet (siehe Abb. 1).

Abb. 1: Beteiligte Institutionen beim Aufbau des „System der Organismen“

¹ Universität Ulm, Abteilung Spezielle Botanik,

² Ruhr-Universität Bochum, Lehrstuhl für Spezielle Botanik,

Fig. 1: Institutions involved in the development of a „system of organisms“

Gerade hier ist es besonders unpraktisch, daß die Informationen von taxonomischen Erstbeschreibungen und Revisionen mehr oder weniger über die Literatur verstreut wird. Dies veranlaßt die wissenschaftlichen Einrichtungen, die im weitesten Sinne als Sammlungen bezeichnet werden können, Botanische und Zoologische Gärten und Herbarien oder Zoologische Sammlungen, parallel für ihre Zwecke ihr System der Organismen mit hohem Aufwand aus Daten aus systematischen Standardwerken, Floren und Revisionen zu erstellen und zu pflegen.

Das Suchen von speziellem Untersuchungsmaterial, z.B. für Inhaltsstoffanalysen oder molekularbiologische Arbeiten, stellt einen ebensolchen Aufwand dar. Im Extremfall bedarf es einer Anfrage bei den ca. 80 deutschen Botanischen Gärten, ob die gewünschte Pflanze irgendwo in Kultur ist, wenn man nicht beim langwierigen Durchsuchen der jeweiligen Samenkataloge fündig geworden ist.

Was liegt also heutzutage näher, als solche Informationen in Datenbanken zu speichern? Ein erster Schritt ist es, die Daten lokal zu speichern für den jeweiligen Zweck, der zweite und weiterführende Schritt ist es dann, die Daten des/eines Organismensystems in gemeinsamen, zentralen Datenbanken zusammenzufassen. Über das mittlerweile für diese Zwecke ausreichend verbreitete und schnelle Internet können Daten online verarbeitet werden, das sog. World Wide Web ermöglicht zusätzlich mit den einfach zu bedienenden Browsern eine unkomplizierte Datenbankabfrage.



Abb. 2: SysTax Hauptmaske

Fig. 2: SysTax main screen

Was ist zu tun? – Das Konzept

Das SysTax-System (KÜMMEL 1989, BOOS 1992, HOPPE 1993, HOPPE ET AL. 1996) entstand genau aus den genannten Anforderungen an eine biologische Datenbank (Abb. 2). Es sollte in der Abteilung Spezielle Botanik der Universität Ulm die systematische Forschung unterstützen und gleichzeitig die Bestände des Botanischen Gartens (GSCHNEIDNER ET AL. 1993) und des Herbariums verwalten helfen. Die Entwicklung des Systems begann 1989 in genau dieser Konstellation der Einrichtungen an der Universität Ulm im Rahmen eines DFG-Projektes zur taxonomischen Bearbeitung der Familie der Eriocaulaceae (STÜTZEL 1989, GRAMBIHLER 1991).

Zentraler Teil des SysTax-Datenbanksystems ist das System der Pflanzen (und Tiere) in insgesamt 23 Hierarchiestufen (siehe Abb. 3). Jedes Taxon ist in einer bestimmten Publikation beschrieben (Literatur), und es existiert hierzu im Falle von Arten und subspezifischen Taxa ein Herbarbeleg (Herbarium).

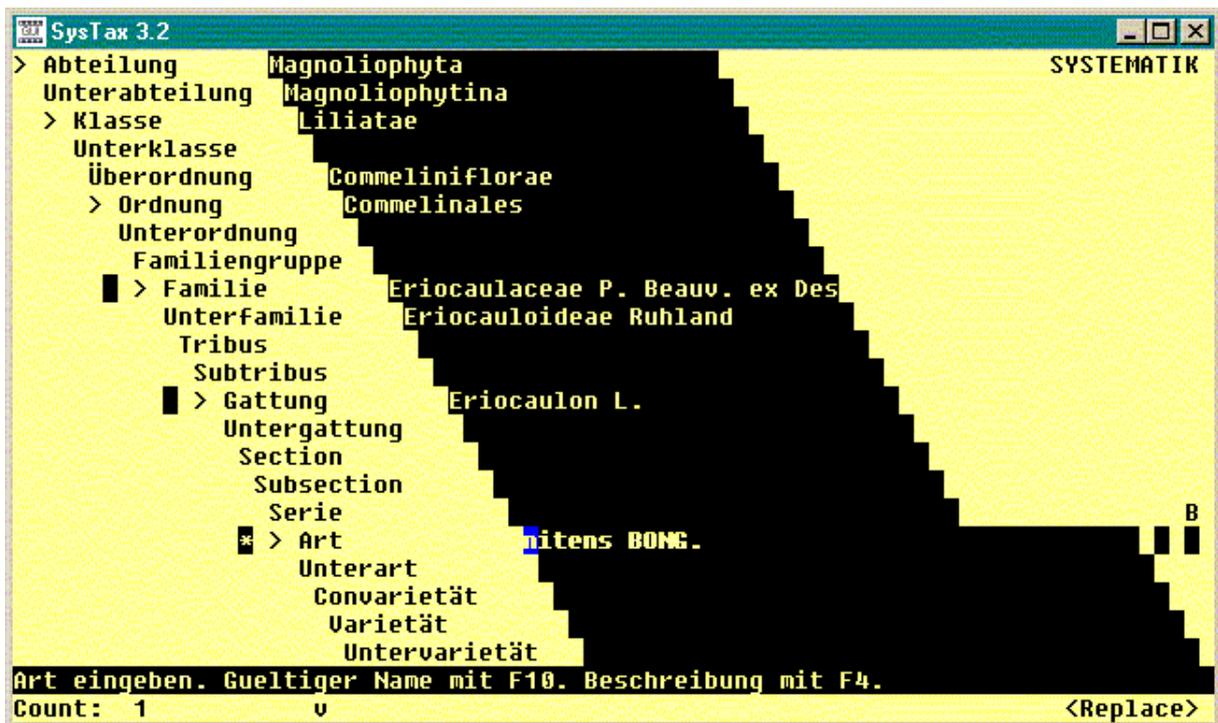


Abb. 3: SysTax Systematik-Maske

Fig. 3: SysTax systematic screen

Hat man ein solches Pflanzensystem gespeichert, kann es leicht für die Verwaltung der Pflanzenbestände eines Botanischen Gartens und der Belege eines Herbariums benutzt werden. Auf die Implementierung einer kompletten Bibliotheksverwaltung mit Leih- und Mahnverkehr wurde verzichtet, das Verwalten von Sonderdrucksammlungen ist aber möglich.

Von Beginn an wurde das System nicht nur als ein „Multiusersystem“ sondern auch als ein „multiinstitutionelles“ System entwickelt. So können mehrere Botanische Gärten und mehrere Herbarien parallel das System benutzen. Während diese Einrichtungen alle auf das gemeinsame System der Pflanzen zugreifen, werden die jeweiligen Bestandsdaten getrennt verwaltet. Dieses Konzept ermöglicht dann den Aufbau von übergreifenden Informationssystemen, wie z.B. dem „Informationssystem Botanischer Gärten“ (HOPPE & BOOS 1996a,b,c, HOPPE 1997) und dem „Bundesinformationssystem Genetische Ressourcen“ (BIG) (s. Beitrag BECKER).

Welche Funktionen sind implementiert? - Der Leistungsumfang

Die Leistungen umfassen derzeit (SysTax Version 3.2) die gesamte botanische Taxonomie, die botanische Sammlungserfassung (Herbarium und andere Sammlungen wie Holzproben oder Samen und Früchte) und Sammlungsverwaltung (Leih- und Mahnverkehr) und die Bestandserfassung und Verwaltung (Erstellung und Versand des Index Seminum, Samentausch) von Botanischen Gärten. Die Programmteile zur Speicherung von Daten der Ethnobiologie (BOOS & HOPPE 1996, HOPPE 1996), der Pflanze-Tier-Interaktionen (FERNER 1998, HOPPE 1998) und von Multimediatdaten liegen als Prototypen vor. Im einzelnen gibt es folgende Funktionen:

Die Masken

Systematik

- Speicherung des Systems in 23 taxonomischen Stufen (Abb. 3)
- mehrstufige, flexible Synonymieverwaltung auf allen Stufen inkl. *pro parte* Synonymien
- Verknüpfung der Taxa mit Literaturdaten
- Verknüpfung der Taxa mit Trivialnamen in beliebiger Anzahl und in beliebigen Sprachen
- Verknüpfung der Taxa mit Literaturdaten
- Verknüpfung der Taxa mit Herbarbelegen
- Zuordnung von Merkmalen und Merkmalsausprägungen zu Taxa
- Ausgabe der Daten im DELTA-Format
- Ausgabe der Daten in der für taxonomische Revisionen üblichen Form

Literatur

- Speicherung der Literaturdaten
- Verknüpfung von Zitat und Taxon
- Verwaltung von Sonderdrucksammlungen

Botanischer Garten

- Speicherung der Gartenakzessionen beliebig vieler Gärten
- Selektives Sperren „sensibler“ Daten für die WWW-Abfrage
- Ausgabe von Bestandslisten nach unterschiedlichen Suchkriterien

- Erstellung des Samenkataloges
- Ausdruck von Adressenetiketten für den Versand der Samenkataloge
- Ausdruck von Adressenetiketten und Etiketten der Samenproben für das Verschicken der Samenproben
- Erstellen einer Versandstatistik für den Index Seminum
- Ausdruck von Steck- und Schlaufenetiketten
- Erstellen von Daueretiketten

Herbarium

- Speicherung der Herbarbogendaten beliebig vieler Herbarien
- Selektives Sperren „sensibler“ Daten für die WWW-Abfrage
- Ausdruck von Herbaretiketten
- Ausdruck von Revisionsetiketten
- Erstellen von Bestandslisten nach unterschiedlichen Kriterien
- Abwicklung der Ausleihen, Ausdruck der hierfür vorgesehenen Leihschein
- Abwicklung von Leihfristverlängerungen und Mahnungen
- Verwaltung des Tauschverkehrs, Erstellen der Tauschstatistik mit benötigtem Schriftverkehr
- Ausdruck von Barcode-Etiketten

Ethnobiologie

- Speicherung ethnobiologischer Daten

Pflanze-Tier-Interaktionen

- Speicherung von Informationen zu Interaktionen zwischen/von Pflanzen und Tieren (Bestäubung, Ausbreitung, Fraß) - Programmteil liegt als Prototyp vor

Sammlungsdaten

- Speicherung beliebiger biologischer Sammlungsdaten zum Aufbau eines Informationssystems Biologischer Sammlungsobjekte

Adressenverwaltung

- Institutsadressen (für Versand von Index Seminum und Samenproben, Herbarleihverkehr)
- Adressen von Personen (mit Beziehungen zu Taxa zum Aufbau einer „Expertendatenbank“)

Verbreitung

- Speicherung von geographischen, biogeographischen und pflanzensoziologischen Verbreitungsdaten
- Zuordnung der Daten zu Taxa, Literatur und Herbarbelegen

WWW-Abfragen

- Die Daten können selektiv über den WWW-Server abgefragt werden. Das Angebot umfaßt derzeit taxonomische Daten (Taxa, Literatur, Synonymien, Typen), einheimische Pflanzen-

namen, Herbar- und Gartendaten und Adressen.

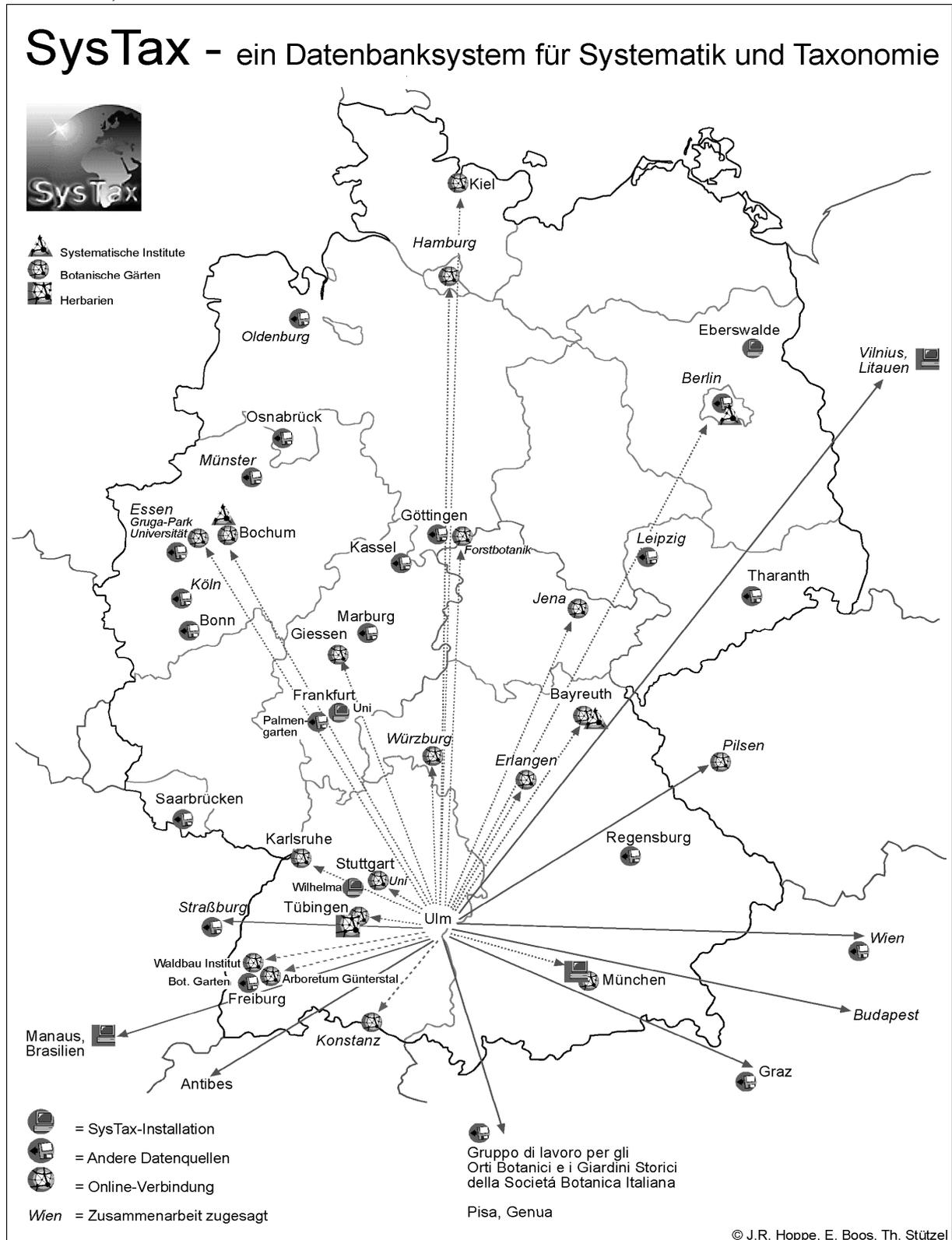


Abb. 4: Nutzer des SysTax-Systems in Ulm und externe Installationen

Fig. 4: Users of SysTax and external installations

Verwendete Standards

Das SysTax-System unterstützt alle wichtigen Standards für Datenstrukturen, Dateninhalte und Datenaustausch, insbesondere diejenigen der „International Working Group on Taxonomic Databases for Plant Sciences“ (TDWG). Bei der Konzeption und Implementierung des gesamten Systems wurde und wird stets drauf geachtet, daß eine Kompatibilität zu diesen Standards erreicht wird.

Für die „**strukturellen**“ Standards, die die Strukturierung von Daten und ihre Beziehungen zueinander beschreiben, gilt generell, daß sie fast nie 1:1 in eine konkrete Datenbank umgesetzt werden („There is no implication in this conceptual information model as to how it should be physically implemented“ (ASC 1993)). Berücksichtigt sind in diesem Sinne der ASC-Standard (ASC 1993), das CDEF-Modell (BEHRENDSONN ET AL. 1996), HISPID (CROFT 1992), DELTA (DALLWITZ ET AL. 1992 ff), POSS (GILLET ET AL. 1995) und einige andere (BEACH 1990, BEACH ET AL. 1990, BISBY 1995, FARR unpublished, FRENCH et al. 1990a und 1990b).

Die „**inhaltlichen**“ Standards, die also Dateninhalte verbindlich vorgeben, sind selbstverständlich in der Datenbank dieselben, die für die taxonomische Arbeit allgemein gelten. Hier sind u.a. zu nennen die Autorenabkürzungen (BRUMMITT AND POWELL 1992), Buchabkürzungen (STAFLEU AND COWAN 1976 - 1988), Zeitschriftenabkürzungen („BPH“, LAWRENCE 1968), die geographischen Bezeichnungen (HOLLIS AND BRUMMITT 1992), der Thesaurus zum Beschreiben von Pflanzennutzungen (COOK 1995), die Adressen und Akronyme von Herbarien (HOLMGREN ET AL. 1990) und Rote Liste Kategorien (z.B. IUCN 1994).

Ein Datenaustausch mit und Datenimport aus anderen Datenbanken erfolgt in einfachen Fällen über „flache“ Files z.B. als MS-EXCEL-Tabellen. Der Austausch komplexerer Daten zwischen zwei relationalen Datenbanken ist ausschließlich über solche Dateien möglich, die in sich die Relationalität der Daten abbilden. Hier werden als Datenaustauschformate das Exchange Data Format (XDF) (ALLKIN AND WHITE 1988) und das International Transfer Format for Botanical Garden Plant Records (ITF) (JACKSON 1996) unterstützt. Zu nennen ist hier ebenfalls das schon genannte HISPID (CROFT 1992).

Wie ist das realisiert? - Die Technik

Für die Realisierung standen zu Beginn der Systementwicklung 1989 die relationalen Datenbankmanagementsysteme INGRES und ORACLE zur Verfügung. Mit dem damaligen für PC viel genutzten DBASE war ein solches komplexes System nicht zu verwirklichen. Die Wahl fiel auf ORACLE, da mit diesem System vor Ort schon Erfahrungen vorlagen, und da es sowohl für MS-DOS als auch für UNIX-Betriebssysteme zu Verfügung stand.

Die Entwicklung des Systems erfolgt heute unter Windows-NT. Die ORACLE-Datenbank (ORACLE Server 7.3.2.2.0 Datenbankserver) läuft in Ulm auf einer HP 9800 E55 unter HP-UX 10.2. Die mit ORACLE Forms v.4.5 als Bestandteil des ORACLE Developers-2000 entwickelten Masken laufen für die einzelnen Nutzer ebenfalls auf dem zentralen Rechner. Der Zugriff auf die Masken und die Datenbank erfolgt über einen Telnet-Client („GUI“- Graphical User Interface, entwickelt und zur Verfügung gestellt von Dr. O. Boos und Dipl. Math. M. Juchem, Abteilung Unfallchirurgie, Klinikum der Universität Ulm), der für Windows-3.x, -95 und -NT zur Verfügung steht.

Die derzeit entstehende neue Windows-Version der SysTax-Benutzeroberfläche ist als Client-Server-Version ausgelegt. Die Applikation läuft dabei vollständig auf dem PC des jeweiligen Anwenders, was den Datenbankserver entlastet. Dies ermöglicht und vereinfacht auch die online-Nutzung der SysTax-Datenbank, da die gesamte Applikation lokal installiert ist und nur die Datenbankeingaben und -abfragen über die Netzverbindung geschickt werden müssen.

Daten von externen Quellen, also anderen Datenbanken und/oder in Listen oder MS-Excel vorliegende Daten, werden mittels externer C-Programme importiert.

Alle nicht über die Masken erfolgenden Ausgaben (taxonomische Reports, Adressenetiketten, Herbaretiketten, Briefe des Herbariumleihverkehrs u.a.) werden derzeit über externe C-Programme auf dem Datenbankrechner erzeugt. Der Ausdruck erfolgt lokal über den Windows-Druckmanager. Als Ausgabemedien dienen derzeit Nadeldrucker (Adressenetiketten, Etiketten für Samenproben), Laserdrucker (Herbaretiketten, Revisionszettel, Briefe und Formulare von Herbariumleih- und -tauschverkehr, Mahnbriefe) und Thermotransferdrucker (Plastikstecketiketten).

Längere taxonomische Reports und Listen (z.B. Bestandslisten des Botanischen Gartens, Indices Seminum) werden sinnvollerweise nicht direkt ausgedruckt sondern in eine Datei geschrieben. Diese Dateien können für eine weitere Formatierung in einem beliebigen Textverarbeitungsprogramm über einen FTP-Client oder einen WWW-Browser vom anonymen FTP-Server des SysTax-Rechners geholt werden.

In den Masken des Programmteiles Herbarium kann mit einem Barcode-Stift die Akzessionsnummer der Herbarbögen eingelesen werden, was den Herbarleihverkehr wesentlich effektiver macht.

Auf dem SysTax-Rechner ist ein eigener WWW-Server (Apache 1.2.0) installiert für die WWW-Seiten des SysTax-Systems selbst und die WWW-Masken mit Datenbankabfragen. Die Abfragen werden über die cgi-Schnittstelle des WWW-Servers und externe C-Programme an die ORACLE-Datenbank weitergeleitet und verarbeitet.

Wie kann man zusammenarbeiten? - Das Angebot

Die online-Nutzung des Ulmer SysTax-Systems ist für wissenschaftliche Einrichtungen kostenlos, ebenso wie natürlich die Abfrage via WWW. Dies gilt sowohl für die Benutzung des GUI-Telnet-Clients als auch für die spätere Client-Server-Version.

Zur Einarbeitung und für Tests steht parallel zur Arbeitsdatenbank eine völlig getrennte Testdatenbank zur Verfügung, für die bei Bedarf eine Zugangsberechtigung erteilt wird. Das Handbuch (HOPPE 1993 ff) steht als Datei auf dem FTP-Server oder als WWW-Dokument zur Verfügung.

Bei einer lokalen Installation des Gesamtsystems wird eine komplette Lizenz des ORACLE-Datenbankmanagementsystems samt Entwicklungsumgebung benötigt. Die SysTax-Applikation wird auch in diesem Fall kostenlos weitergegeben. Eine separate Installation ermöglicht natürlich nicht, die Merkmale der gemeinsamen Verwendung von Daten zu nutzen, für die ja das SysTax-System eigentlich konzipiert wurde.

Was wird derzeit gemacht? - Die Projekte und Nutzer

Im Idealfall ist das SysTax-System eine gemeinsam zu nutzende, zentrale taxonomische Datenbank, in der Spezialisten für verschiedene Pflanzen- und Tiergruppen die Daten eingeben und kontinuierlich verbessern (zum derzeitigen Datenbestand siehe Tab. 1). Das System ist hierbei in der Lage, die einzelnen Datenbereiche vor „unbefugtem“ oder zufälligem Update zu schützen.

Taxonomische Arbeiten

Das SysTax-System wird derzeit für einige taxonomische Bearbeitungen verwendet, oder es werden bestimmte Taxa von einzelnen Spezialisten betreut. Es sind dies die Eriocaulaceae (STÜTZEL, Bochum), Asclepiadaceae (insbesondere die Gattung *Cynanchum*; LIEDE, Bayreuth; s. LIEDE 1995), Gesneriaceae (FREIBERG, Ulm), Ascolichenes (TRIEBEL und RAMBOLD, München), Bromeliaceae (GSCHNEIDNER, Ulm), Lebermoose (BERNECKER, Ulm), Valerianaceae (WEBERLING, Ulm) und Boraginaceae (HILGER, FU-Berlin). Eine allgemeine Überarbeitung des Pflanzensystems erfolgt derzeit im Rahmen des BIG-Projektes. Die Zuordnung von Gattungen zu Familien folgt dem System von BRUMMITT (1992).

Tab. 1: Überarbeitung des Pflanzensystems

Tab. 1: Revision of plant systematics

Taxa	52.000 Arten	davon gültig	38.500
	14.800 Gattungen	davon gültig	14.600
	650 Familien	davon gültig	604
Herbarium	18.000 Bögen aus 57 Herbarien mit 5.000 Taxa		
Botanische Gärten	101.000 Gartenakzessionen mit 36.000 Taxa aus 22 Gärten		
Literatur	2.300 Zitate		
Adressen	922		

Botanische Gärten

Einige Botanische Gärten (Abb. 4) nutzen das SysTax-System online für Ihre Bestandsdaten oder haben eine solche Nutzung in Aussicht gestellt.

Das „**Informationssystem Botanischer Gärten**“ (Abb. 5) wurde an der Universität Ulm initiiert und ist ein Informationsangebot für den Verband Botanischer Gärten im WWW. Es bietet zum einen möglichst komplette Linklisten zu den Botanischen Gärten im In- und Ausland und zu für die Arbeit in den Gärten wichtigen Informationen. Zum anderen bietet es die Möglichkeit, Material aus den Samenkatalogen und Bestandslisten der Gärten abzufragen und direkt zu bestellen. Dieses Angebot richtet sich an die Gärten und an Einrichtungen, die Material für ihre wissenschaftliche Arbeit benötigen. Über eine einfach zu bedienende Abfragemaske (Abb. 6) kann derzeit in über 100.000 Gartenakzessionen von über 36.000 Arten gesucht werden. Die Daten stammen von 22 Botanischen Gärten. Weitere 15 Gärten haben Ihre Zusammenarbeit bereits in Aussicht gestellt.

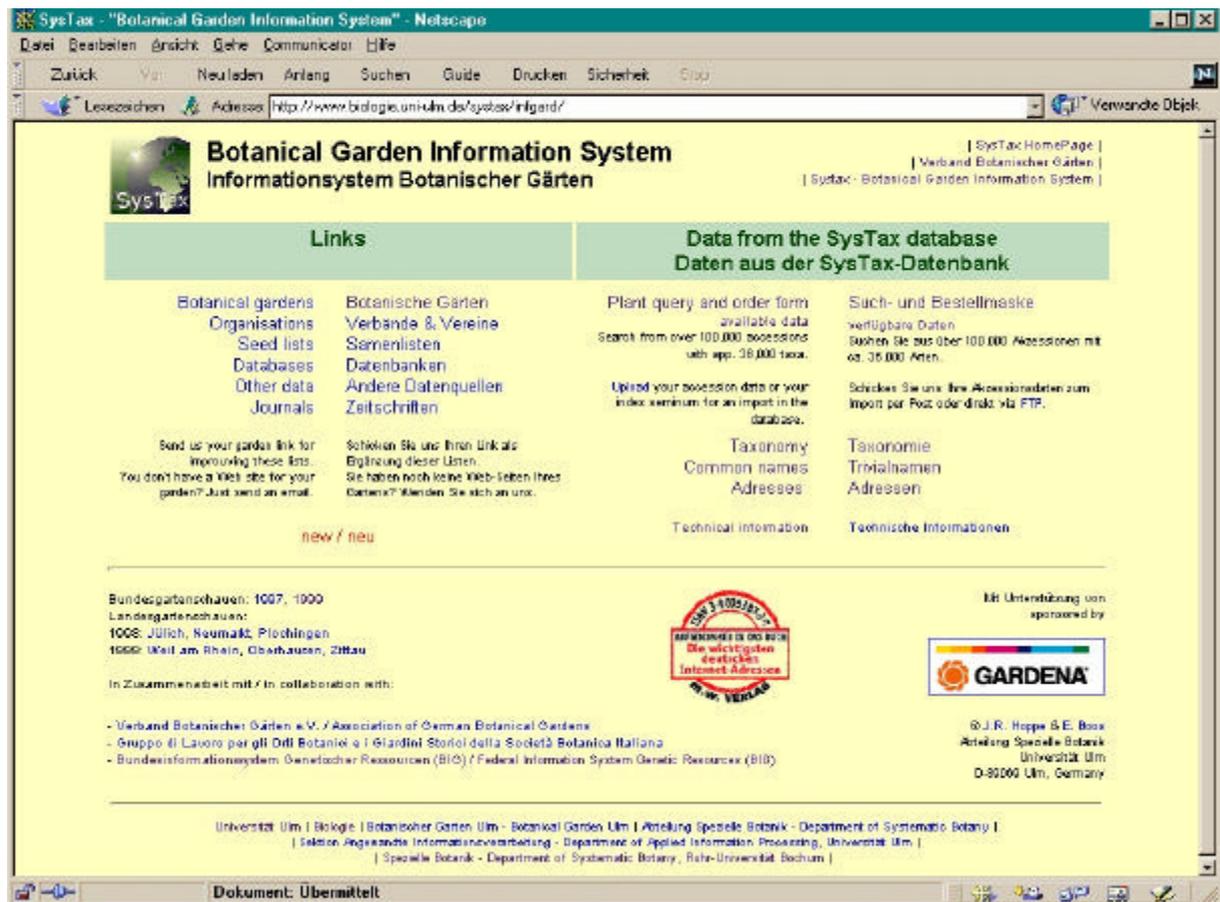


Abb. 5: Das Informationssystem Botanischer Gärten im Internet
 Fig. 5: The Information system of botanical gardens on the Internet

Die Daten werden entweder online erfaßt, oder es erfolgt ein Import der jeweiligen Bestands- oder Samenkatalogdaten. Dieses Datenangebot wird im „Bundesinformationssystem Genetische Ressourcen“ (BIG) erweitert.

Das **Bundesinformationssystem Genetische Ressourcen (BIG)**, finanziert durch das Bundesministerium für Bildung und Forschung (BMBF), ist ein Gemeinschaftsprojekt des Bundesamts für Naturschutz (BfN); der Ruhr-Universität Bochum (RUB), Lehrstuhl Spezielle Botanik und Botanischer Garten; dem Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Genbank und der Zentralstelle für Agrardokumentation und -information (ZADI), Informationszentrum für Genetische Ressourcen (IGR). Das SysTax-System (Teilprojekt RUB) liefert in diesem Zusammenhang das taxonomische Verweissystem zum Auffinden von Pflanzenmaterial in Botanischen Gärten und anderen Sammlungen. Eine Bilddatenbank von Pflanzen aus Botanischen Gärten soll eine Identifikation von kultivierten Pflanzen erleichtern.

Herbarien

Ebenso wie Botanische Gärten können auch beliebig viele Herbarien mit dem SysTax-System online arbeiten (Abb. 6).

Im **Herbarium der Botanischen Staatssammlung München (M)** ist das SysTax-System unter Windows-NT installiert. Es wird hier zur Verwaltung der Ausleihen und des Tauschverkehrs verwendet und in Zusammenarbeit mit den Kustoden in diesen Funktionen optimiert. Weitere Herbarien haben eine Nutzung des Systems geplant.

Finanziert durch die Gesellschaft für technische Zusammenarbeit (GTZ) wird am Herbarium der Universidade Federal do Amazonas (UFAM) in Manaus im Rahmen eines Projektes der Universität Ulm zur „**Biodiversitätserfassung in Herbarien des Amazonasgebieten**“ (HOPPE ET AL. 1996) ebenfalls ein System unter Windows-NT installiert. Es soll der Erfassung der dortigen Herbarbestände dienen, was eine Grundlage für eine spätere, umfassende Auswertung der Daten ist.

Die parallele Bearbeitung der Taxonomie und die Eingabe von Sammlungsdaten aus Herbarien und Botanischen Gärten legt es nahe, noch umfassendere Informationssysteme aufzubauen, welche ihre Daten ebenfalls durch einfache WWW-Masken der wissenschaftlichen Öffentlichkeit zur Verfügung stellen. Im Auftrag der EDV-Kommission der Direktorenkonferenz Naturwissenschaftlicher Forschungssammlungen (DNFS) wurde so ein Prototyp eines Informationssystems „**Biologische Sammlungsobjekte in Deutschland (BISAM)**“ entwickelt, welches nicht nur die Bestände von Botanischen Gärten und Herbarien, sondern von beliebigen biologischen Sammlungen (Botanik, Zoologie, Paläontologie, Mikroorganismen) ermöglicht.

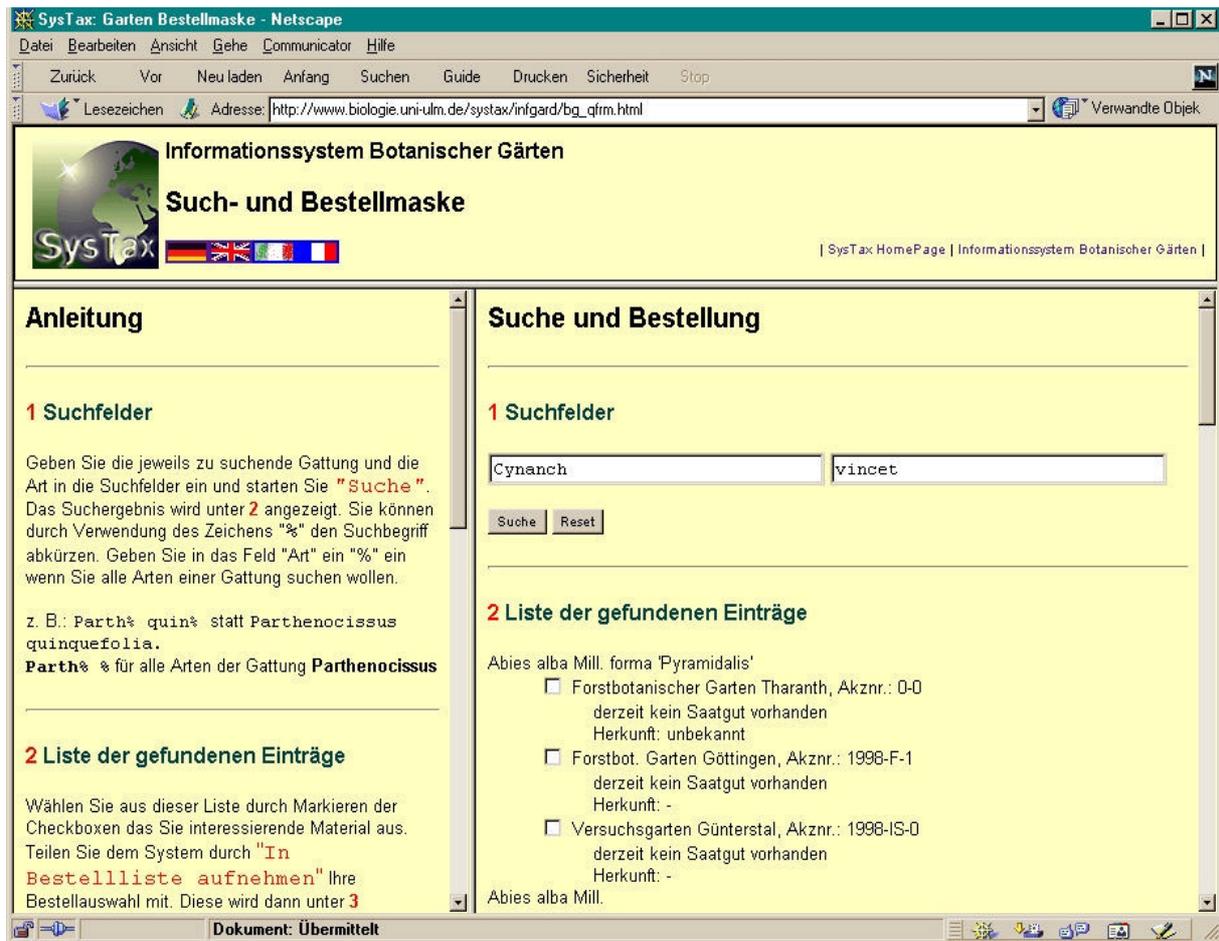


Abb. 6: Such- und Bestellmaske für Material aus Botanischen Gärten
 Fig. 6: WWW search and order screen

Welche Funktionen werden noch gebraucht? - Die Weiterentwicklung

Das SysTax-System wird in Zusammenarbeit mit den Nutzern ständig erweitert und verbessert. Derzeit liegt der Schwerpunkt neben den genannten Projekten besonders bei der Entwicklung einer neuen Windows-Benutzeroberfläche. Hierbei werden besonders die Speichermöglichkeit von Abbildungen, die Implementierung einer „alternativen Taxonomie“ (entsprechend etwa dem „Potential-Taxon-Konzept“ nach BEHRENDSONN (1995)) und die Taxonomie der Kulturpflanzen berücksichtigt. Das Angebot an Abfragemöglichkeiten im WWW wird hierbei ebenfalls ständig erweitert und verbessert.

Wo gibt es weitere Informationen? - Literatur und URLs

Standards (Strukturmodelle, Austauschformate, inhaltliche Standards)

- ALLKIN, R., R.J. WHITE (1988): XDF: a language for the description and exchange of biological data sets. An introductory document submitted to the UIBS Commission for Plant Taxonomic Databases at St. Louis, Missouri, USA.
- ASC (1993): An information model for Biological Collections (Draft). Report of the Biological Collections Data Standards Workshop, August 18-24, 1992. Association of Systematic Collections, Committee on Computerization and Networking.
- BEACH, J.H. (1990): Botanical Specimen Data Dictionary. - International Working Group on Taxonomic Databases in Plant Sciences.
- BEACH, J.H., S. PRAMANIK UND J.H. BEAMAN (1990): Hierarchical Structures for Taxonomic Database Systems.
- BEHRENDSONH, W. (1995): The concept of "potential taxa" in databases. *Taxon* 44: 207 - 212.
- BEHRENDSONH, W., A. ANAGNOSTOPOULOS, G. HAGEDORN, J. JAKUPOVIC, P.L. NIMIS, AND B. VALDÉS (1996): CDEFD - Information Model for Biological Collections. Proceedings of the European Science Foundation Workshop "Disseminating Biodiversity Information", Amsterdam, 25. - 27.3.1996.
- BISBY, F. (1995): Plant Names in Botanical Databases. Plant Taxonomic Database Standards No. 3, International Working Group on Taxonomic Databases for Plant Sciences (TDWG). Hunt Institute for Botanical Documentation, Pittsburgh.
- BRUMMITT, R.K. (1992): Vascular Plant Families and Genera. Royal Botanic Gardens, Kew.
- BRUMMITT, R.K. AND C.E. POWELL (1992): Authors of Plant Names. Royal Botanic Gardens, Kew.
- COOK, E. M. (1995): Economic Botany Data Collection Standard. Prepared for the International Working Group on Taxonomic Databases for Plant Sciences (TDWG). Royal Botanic Gardens, Kew.
- CROFT, J. R. (compiler) (1992): HISPID - Herbarium Information Standards and Protocols for Interchange of Data. Summary paper and data dictionary. Australian National Botanic Gardens (unpublished), available on the Internet via [gopher://muse.bio.cornell.edu/00/standards/hispid/hispid91.std](http://muse.bio.cornell.edu/00/standards/hispid/hispid91.std).
- DALLWITZ, M.J., T.A. PAINE AND E.J. ZURCHER (ab 1992): The DELTA System. <http://biodiversity.uno.edu/delta/>.
- FARR, E. (1991): Proposed new consensus structure for types databases. Working document, International Working Group on Taxonomic Databases for Plant Sciences (TDWG) (unpublished).
- FRENCH, J.C., A.K. JONES AND J.L. PFALTZ (eds.) (1990a): Scientific Database Management. - Report on the Invitational NSF Workshop on Scientific Database Management, Charlottesville, VA, USA.
- FRENCH, J.C., A.K. JONES AND J.L. PFALTZ (eds.) (1990b): Scientific Database Management (Panel Reports and Supporting Material. Report on the Invitational NSF Workshop on Scientific Database Management, Charlottesville, VA, USA.

- GILLET, H., C. LEON, D. MACKINDER, P. ROONEY, H. SYNGE AND K. WALTER (1995): Plant Occurrence and Status Scheme (POSS). - A standard for recording the relationship between a plant and a place. Plant Taxonomic Database Standards, V. 2.0, September 1995. Published for the International Working Group on Taxonomic Databases for Plant Sciences (TDWG) by the World Conservation Monitoring Centre.
- HOLLIS, S AND R.K. BRUMMITT (1992): World Geographic Scheme for Recording Plant Distributions. Plant Taxonomic Database Standard No. 2, Version 1.0, January 1992. Published for the International Working Group on Taxonomic Databases for Plant Sciences (TDWG) by Hunt Institute for Botanical Documentation, Carnegie Mellon University, Pittsburgh.
- HOLMGREN, P.K., N.H. HOLMGREN AND L.C. BARNETT (1990): Index Herbariorum, Pt. 1: The Herbaria of the World (ed. 8). Regnum Vegetabile 120.
- IUCN/WWF (1987): The International Transfer Format (ITF) for Botanic Garden Plant Records. Plant Taxonomic Database Standards No.1. Hunt Institute for Botanical Documentation, Pittsburgh.
- IUCN SPECIES SURVIVAL COMMISSION (1994): IUCN Red List Categories, as approved by the 40th meeting of the IUCN Council. Gland, Switzerland.
- JACKSON, D.W. (compiler) (1996): International Transfer Format for Botanical Garden Plant Records (Version 2.00 - Draft 3.1). Botanical Gardens Conservation International.
- LAWRENCE, G.M.H. ET AL. (1968): B-P-H, Botanico-Periodicum-Huntianum. Pittsburgh, Pa.: Hunt Botanical Library.
- STAFLEU, F.A. AND R.S. COWAN (1976 - 1988): Taxonomic Literature. 2nd ed. Vol.1 Utrecht: BOHN, SCHELTEMA AND HOLKEMA (1976); Vol.2-7 Utrecht: BOHN, SCHELTEMA AND HOLKEMA, The Hague: dr. W. Junk b.v., Publishers.

Originalarbeiten und Tagungsbeiträge SysTax:

- BOOS, E. (1992): Botanische Klassifikation und Taxonomie - Konzeption und Realisierung eines Informationssystems. Diss. Univ. Ulm, Fak. für Mathematik und Wirtschaftswissenschaften.
- BOOS, E. AND J.R. HOPPE (1996): Implementation of an Ethnobiological Information System within the systematic and taxonomical database system SysTax (Posterabstract No. 14.12). 5th International Congress of Ethnobiology, 2.-6.9.1996, Nairobi, Kenia. Abstracts p. 45.
- GRAMBIHLER, H.G. (1991): Systematisch-taxonomische Untersuchungen an Eriocaulaceae Südostasiens. Diplomarbeit Univ. Ulm.
- GSCHEIDNER, M., E. BOOS UND TH. STÜTZEL (1993): SysTax - Ein Datenbanksystem und dessen Nutzung in Botanischen Gärten. Gärtnerisch-Botanischer Brief **111**: 23 - 29.
- FERNER, W. (1998): Konzeption und Implementierung von "Plant-Animal-Interactions". Universität Ulm, Fakultät für Informatik, Diplomarbeit.
- HOPPE, J.R. (1993 ff): SysTax, ein Datenbanksystem für Systematik und Taxonomie - Online Handbuch v. 3.2. <http://www.biologie.uni-ulm.de/systax/handbuch/>.
- HOPPE, J.R. (1996): A Prototype for an Ethnobiological Information System on the basis of the systematic and taxonomical database system SysTax (Abstract No. 10.2). 5th International Congress of Ethnobiology, 2.-6.9.1996, Nairobi, Kenia. Abstracts p. 36.
- HOPPE, J.R. (1998): Practical suggestions for database implementations. In: THIÉRY, M., A.-D. STEVENS, J.R. HOPPE, P. CHARLES-DOMINIQUE AND K.-L. SCHUCHMANN: Angiosperm pollina-

- tion and seed dispersal, a review. *Ecotropica* 4: 69 - 91.
- HOPPE, J.R. UND E. BOOS (1996a): Informationssystem Botanischer Gärten (Posterabstract P-13.110). Jahrestagung der Deutschen Botanischen Gesellschaft, Düsseldorf, 25.-31.8.1996. Abstracts p. 350.
- HOPPE, J.R. UND E. BOOS (1996b): Ein "Informationssystem Botanischer Gärten" auf der Basis des Datenbanksystems SysTax. *Gärtnerisch-Botanischer Brief* 124: 10 - 15.
- HOPPE, J.R. UND E. BOOS (1996c): A "Botanical Garden Information System" on the basis of the systematic and taxonomic database system SysTax. (Posterabstract p. 34 - 35). Global Biodiversity Research in Europe. International Senckenberg Conference, Frankfurt a.M., Dec. 9-13, 1996.
- HOPPE, J.R., E. BOOS, G. BEDINI AND S. LIEDE (1997): INTERNET-ional Botanical Garden Information System (Abstract). Annual meeting of the Botanical Society of America, Montréal, 3.-7. August 1997. - *Amer. J. Bot.* (suppl.) 84 (6): 251.
- HOPPE, J.R., E. BOOS AND G. GOTTSBERGER (1996): The database system SysTax - an aid for systematics and taxonomy and the management of botanical gardens and herbaria. *Albertoa* 4 (9): 107 - 108.
- HOPPE, J.R., E. BOOS, A.C. WEBBER, M. GORDO, J. DA CRUZ, A.-D. STEVENS, K.V. MEHLTRETER AND G. GOTTSBERGER (1996): The implementation of the SysTax database system for the documentation of Amazonian biodiversity (Posterabstract). 3. Brazilian-German workshop for exchange in environmental science. Tagung "Umwelt und Chemie" der Gesellschaft Deutsche Chemiker. Ulm, 10.-12.10.1996.
- KÜMMEL, D. (1989): Software-Paket für Systematik und Taxonomie auf der Basis einer relationalen Datenbank. Diplomarbeit Univ. Ulm, Sektion Angewandte Informationsverarbeitung.
- LIEDE, S. (1995): Studien in den Cynanchinae (Asclepiadeae - Asclepiadaceae). Habilitationsschrift, Fakultät für Naturwissenschaften der Universität Ulm.
- STÜTZEL, TH. (1989): Eriocaulaceae. Fakultät für Naturwissenschaften und Mathematik, Universität Ulm.

URLs:

Biologische Sammlungsobjekte in Deutschland (BISAM); i. A. der Direktorenkonferenz Naturwissenschaftlicher Forschungssammlungen Deutschlands:

<http://www.biologie.uni-ulm.de/museum/>

Bundesinformationssystem Genetische Ressourcen (BIG): <http://www.dainet.de/genres/big/>

Informationssystem Botanischer Gärten: <http://www.biologie.uni-ulm.de/systax/infgard/>

SysTax - ein Datenbanksystem für Systematik und Taxonomie:

<http://www.biologie.uni-ulm.de/systax/>

Verband Botanischer Gärten: <http://www.biologie.uni-ulm.de/verband/>

Das SysTax-Projekt ist eine Gemeinschaftsentwicklung der Abt. Spezielle Botanik der Ruhr-Universität Bochum (Th. Stützel) und der Abt. Spezielle Botanik (J.R. Hoppe & E. Boos) und der Sektion Angewandte Informationsverarbeitung (F. Schweigert) der Universität Ulm. Das Projekt wurde und wird unterstützt von der Deutschen Forschungsgemeinschaft, der Universität Ulm, der

Deutschen Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH und dem Bundesministerium für Bildung und Forschung (BMBF).

Datengrundlagen des Bundesamtes für Naturschutz für das Bundesinformationssystem Genetische Ressourcen

Data basis of the Federal Agency for Nature Conservation (BfN) for the Federal Information System Genetic Resources (BIG)

HANS G. FINK¹ UND RUDOLF MAY¹

Unter dieser Überschrift wird über den im Bundesamt für Naturschutz (BfN) erreichten Entwicklungsstand der IT-basierten Informationsgrundlagen für Fragen des Artenschutzes bei Pflanzen berichtet, die für das Bundesinformationssystem Genetische Ressourcen (BIG) relevant und nutzbar erscheinen. BIG-Objekt sind derzeit ausschließlich die pflanzengenetischen Ressourcen.

Es wird vorausgesetzt, daß biologisch-ökologische Objektinformation (hier über Pflanzen) unabhängig von ihrer Herkunft in jeglichem Sachzusammenhang zugänglich sein muß, in dem sie benötigt wird. Unterschiedliche Rechtersysteme (so auch BIG) haben die Aufgabe, den Informationszugang aus speziellen Blickrichtungen problem- bzw. fragespezifisch zu optimieren und die Verbindung von Information aus verschiedenen Bereichen zu ermöglichen. Im Falle von BIG sind dieses die Botanischen Gärten, der Bereich Pflanzenzüchtung einschließlich Genbanken sowie der Naturschutz.

Die in sektoralen Datenbanken verfügbare Basisinformation (Faktengrundlagen, im BfN hier zum Thema Artenschutz Pflanzen) soll jedoch auch für andere (grundsätzlich alle) Informationssysteme zugänglich/offen sein. Basisdaten als solche dürften dann nicht kommerzialisierbar sein - allenfalls die Rechercheinstrumente und Expertensysteme, die auf sie zugreifen. (Sicherheitsgesichtspunkte: Objekt-, Daten- und Personenschutz ist natürlich Rechnung zu tragen, siehe unten)

Dementsprechend ist BIG, das geplante „Bundesinformationssystem Genetische Ressourcen“ - wenn auch das erste - nur eins der Informationssysteme, über die naturschutzrelevante biologisch-ökologische Objektinformation (Pflanzen) aus dem BfN zugänglich gemacht werden soll (s. Beitrag BECKER). Daneben befindet sich ein deutsches Umweltinformations-Netzwerk im Aufbau (Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit und Umweltbundesamt mit Beteiligung des BfN), über das die Information aus der „Datenbank Gefäßpflanzen“ des BfN ebenfalls zugänglich und Querverbindungen zu BIG möglich sein sollen. Vom Bereich Umwelt- und Naturschutz aus betrachtet ist BIG also ein wichtiger Baustein unter der größeren Klammer der Information über Naturressourcen und erschließt dem Naturschutz bisher noch nicht zugängliche Informationsbereiche.

¹ Bundesamt für Naturschutz

Fachübergreifender Bedarf an artenbezogener Objektinformation über pflanzengenetische Ressourcen

Das im Mittelpunkt des BIG-Projektes stehende Objekt Pflanze ist aus ökologisch-schützerischer Sicht (als lebende Ressource der Naturerneuerung und Teil der biologischen Vielfalt) von ebenso grundlegender Bedeutung, wie aus ökonomischer (als pflanzengenetische Züchtungs- und Produktions-Ressource). In einem früheren Beitrag (FINK UND MAY 1993) wurde unter Bezugnahme auf die Definitionen von BOMMER und BEESE (1990) bereits erläutert, daß sich die Interessen des Schutzes pflanzengenetischer Ressourcen mit jenen des Naturschutzes in maßgeblichen Punkten decken. Schon alleine daraus leitet sich für Nutzer und Schützer gemeinsamer Informationsbedarf bezüglich des Nutz- und Schutzobjektes ab. Wie einleitend bereits angesprochen, besteht jedoch auch auf einer noch breiteren Kommunikationsebene der Bedarf, biologische Objektinformation fachübergreifend zugänglich zu machen. Die mit den natürlichen und züchterisch bearbeiteten pflanzengenetischen Ressourcen befaßten Bereiche, einschließlich des Naturschutzes, müssen sich daran im eigenen Interesse beteiligen. Ansätze zur Metainformationserschließung sind dabei ebenso wichtig wie faktenbezogenes „Datamining“ und auswertende Expertensysteme.

Durch das Nachhaltigkeitskonzept des Übereinkommens über die Biologische Vielfalt der Konferenz der Vereinten Nationen für Umwelt und Entwicklung vom Juni 1992 in Rio de Janeiro (siehe auch Gesetz zum Übereinkommen über die biologische Vielfalt vom 30.8.1993) wird die enge Verbindung von Nutzung und Schutz der biologischen Ressourcen der Erde als grundlegende Maxime der Zukunftssicherung unterstrichen. Neben den Pflanzen- und Tierarten und ihrer innerartlichen Vielfalt (innerartliche genetische Diversität) sind auch die Lebensgemeinschaften und Lebensräume als Bestandteile der Biodiversität bzw. als Voraussetzung ihres Bestehens definiert. Die Unterzeichnerstaaten verpflichten sich, die notwendige Information über Bestand, Bestandsentwicklung und biologische/ökologische Existenzbedingungen der biologischen Vielfalt zu erfassen und bereitzustellen. Objektmonitoring und interdisziplinärer Informationszugang sind hierfür unabdingbar.

Im BMBF-finanzierten BIG-Projekt wird in Deutschland erstmalig versucht, den Informationsbedürfnissen bezüglich Nutzung und Schutz der Artenvielfalt der Pflanzen durch interdisziplinären Zugang zu Fachdaten zu genügen. Welche Datenkategorien hierfür besonders wichtig erscheinen, wurde in der Mitteilung von 1993 (FINK UND MAY) vorgestellt.

Allgemeiner Hintergrund der Beschaffung und Bereitstellung von Fachinformation für Naturschutz auf Bundesebene am Beispiel Pflanzen

Im Naturschutz gibt es noch kein strukturiertes Datenmeldesystem (vergleichbar den Statistiken anderer Branchen). Die Verwaltungsvereinbarung zwischen Bund und Ländern über den Datenaustausch im Umweltbereich (1993) hat angesichts ungleicher Bearbeitungsstände und

Entwicklungsperspektiven in den Bundesländern eher deklaratorischen Wert und kann bestenfalls als ein erster Ansatz in Richtung einer tatsächlichen, zentralen Datenverfügbarkeit Länder-Bund bewertet werden. Hingegen drängt die internationale Entwicklung stetig in Richtung einer solchen Verfügbarkeit im Rahmen von Meldepflichten (z.B. FFH-Richtlinie, OECD-Statistik) und Informationserfordernissen (Europäische Umweltagentur etc.). Durch das oben bereits erwähnte Übereinkommen über die biologische Vielfalt werden die einzelnen Staaten dazu verpflichtet, naturschutzrelevante Information bereitzustellen. Zunehmend wird der Bedarf an länderübergreifendem Bestands- und Trendmonitoring formuliert. In den Gesetzesentwürfen zur Novellierung des Bundesnaturschutzgesetzes (BnatSchG) findet sich der Auftrag an Bund und Länder, gemeinsam eine repräsentative Umweltbeobachtung durchzuführen.

Vor diesem Hintergrund eines allgemein zunehmenden Problembewußtseins bei in der Breite noch nicht vorhandenen oder funktionierenden Melde- und Abstimmungsmechanismen laufen (teilweise bereits seit den 70er Jahren) eine Menge staatlicher (Bundesländer) und privater (Forschungseinrichtungen und Verbände) Datenerhebungen (zumeist Geländeerhebungen), bei denen insgesamt sehr große Datenmengen erzeugt werden (Artenkartierungen, Populationserhebungen, Biotopkartierungen), die mittlerweile i.d.R. auch über elektronische Medien zugänglich sind. Die meisten dieser Daten werden im Rahmen privater wissenschaftlicher Arbeiten unentgeltlich bzw. ehrenamtlich erhoben.

Zentrale Datenzusammenführung auf Bundesebene beinhaltet unter den genannten Voraussetzungen demnach:

- Zusammenarbeit mit den Landesämtern für Naturschutz
- Zusammenarbeit mit wissenschaftlichen Institutionen (Universitäten, sonstige Forschungseinrichtungen)
- Zusammenarbeit mit Daten erhebenden Fachorganisationen
- Zusammenarbeit mit Fachexperten
- Auswertung wissenschaftlicher Datenquellen
- Projektauswertung (z.B. Biotopkartierungen u. Populationserhebungen der Bundesländer)
- Entwicklung und Abstimmung von Erhebungsmethoden
- Organisation und Aufrechterhaltung des Datenflusses
- Informationstechnische Gewährleistung des Datenflusses und der anwendungsorientierten Bereitstellung von Daten
- Bereitstellung und gezielte Vergabe von Fördermitteln
- Berichterstattung

Datenbank „Gefäßpflanzen“ und andere Datenbestände und -quellen

In Anbetracht der maßgeblichen Rolle der Gefäßpflanzen für das Ökosystem sowie der relativ guten Verfügbarkeit von Daten über diese Artengruppe wurde im BfN modellhaft der Komplex „Datenbank Gefäßpflanzen“ aufgebaut (siehe Beschreibung bei MAY 1994a,b und Methoden unter BERGMEIER 1992). Hierfür war nicht zuletzt ausschlaggebend, daß es durch die Aktivitäten der Zentralstelle für die Floristische Kartierung Deutschlands (Universitäten Bochum, Halle, Regensburg) eine bundesweit vergleichsweise gut funktionierende Organisation auf ehrenamtlicher Beteiligungsbasis und einen bestehenden Datenfluß gab. Dieser konnte durch bundesseitige Förderung organisatorischer und informationstechnischer Maßnahmen aktiviert werden. Darüber hinaus hatten bereits im Vorlauf auch etliche Bundesländer die Zusammenführung der Daten bei den Landesämtern eingerichtet. Dennoch werden auch heute noch nicht alle in den Bundesländern erhobenen Daten dort auch zentralisiert, so daß bei länderübergreifenden Bundes-Vorhaben auch direkt auf die Regionalstellen der Florenkartierung zurückgegriffen werden muß. Damit treffen alle bereits genannten Aspekte der zentralen Datenzusammenführung für die Datenbank Gefäßpflanzen zu. (Das FE-Projekt „Datenbank Gefäßpflanzen“ ist von 1988 bis 1997 erfolgreich gelaufen und hat einen umfassenden Datenbestand erbracht (s. unten).

Mit der laufenden, zentralen Zusammenführung der Daten aus der bundesweiten Florenkartierung werden im wesentlichen Daten über Arteninventar, Vorkommensstatus, Verbreitung und Bestandswandel der Farn- und Blütenpflanzen verfügbar gemacht. Weitere Daten werden im Rahmen anderer Amtsaufgaben intern und in Kooperation mit der Fachwelt erarbeitet (Rote Listen), gehen aus geförderten Forschungs- und Entwicklungsvorhaben hervor oder werden der Fachliteratur entnommen bzw. aus Datenbeständen(-banken) außerhalb des Hauses bezogen. An veröffentlichten Ergebnissen seien neben der Roten Liste (siehe auch Gesamtdarstellung aller Roten Listen auf CDROM, V.I.M.-Verlag 1999) die Standardliste der Farn- und Blütenpflanzen Deutschlands (WISSKIRCHEN UND HAEUPLER 1998), sowie die Auswertung der Roten Liste Gefäßpflanzen im Hinblick auf Gefährdungsursachen (KORNECK ET AL. 1998) genannt.

Weitere Datenbestände über Gefäßpflanzen werden im BfN für international handelsrelevante Arten (Datenbank WISIA) sowie für die Arten der FFH-Richtlinie aufgebaut.

Zu den Moosen, Algen, Flechten und Pilzen gibt es bundesweit zwar auch bereits etliche Erhebungsprojekte und Daten, jedoch noch in keinem, mit den Gefäßpflanzen vergleichbaren Umfang (siehe auch Rote Liste 1996). Dieses Defizit ist nur in Schritten behebbar und müßte bei zukünftigen Förderschwerpunkten von Bund und Ländern stärker berücksichtigt werden. Förderansätze des BfN gab es bisher nur für die Pilze (Rote Liste, Artenliste) und Moose (Artenliste).

Gesichtspunkte des Datenschutzes sowie urheberrechtliche Gesichtspunkte

Aus Gründen des Schutzes bestandsbedrohter Pflanzenpopulationen vor mißbräuchlichen, schädigenden Zugriffen werden vom BfN keine punktscharfen Vorkommensdaten weitergegeben, jedoch die Kontaktmöglichkeiten bzw. Datenquellen genannt (z.B. Landesamt für Naturschutz), über die unter entsprechenden Voraussetzungen der Zugang zu mehr Information möglich ist.

Bei der Vielzahl der privaten Datenquellen für Geländedaten (vorwiegend ehrenamtliche Mitarbeiter der Florenkartierung) ist es darüber hinaus notwendig, einer möglichen Vielzahl regionaler Bedingungen für die Weitergabe von Daten Rechnung zu tragen. Dieses ist eine überaus wichtige Grundlage der vertrauensvollen Zusammenarbeit von Datenerhebern, Zentralstelle und BfN. Die mit den Datenbeständen einzelner Regionalstellen verknüpften Bedingungen sind in die Datenbank eingearbeitet. Aufgrund der Vereinbarungen zwischen Regionalstellen (teils die Bundesländer), Zentralstelle und BfN ist nur der auf Meßtischblatt- oder Meßtischblattquadrantenebene aggregierte, sogenannte „Datenstand Atlas“ uneingeschränkt nutzbar. Dieses ist der aktualisierte, inhaltlich abgestimmte Datenstand, der statistischen Auswertungen zugrunde gelegt werden kann und in den Verbreitungsatlanten von Bund und Ländern veröffentlicht ist.

Urheberrechte kommen teilweise auch bei der Bereitstellung und Nutzung von Daten aus veröffentlichten Datenquellen zum Tragen.

All diese Gesichtspunkte, denen vornehmlich auch bei einem kommerziellen Nutzungskonzept für BIG Rechnung zu tragen ist, müssen berücksichtigt und noch geklärt werden.

IT-basierte Fachverfahren des BfN, in denen BIG-relevante Objektinformation über Pflanzen für interne und externe Informationssysteme verfügbar gemacht wird

Datenbank FLORKART (Datenbank Gefäßpflanzen)

Die Datenbank FLORKART beinhaltet Informationen über Vorkommen, Verbreitung und Bestandssituation der Gefäßpflanzen in Deutschland. Sie hat derzeit einen Datenbestand von ca. 14 Millionen Einzelnachweisen.

Im einfachsten Fall besteht ein Einzelnachweis aus der Angabe des Vorkommens einer Art in einem Rasterfeld zu einem Nachweis-Datum und von einer Datenquelle (Geländerheber / Literatur / Sammlung) stammend. Die Rasterfelder sind abgeleitet aus dem Blattschnitt der Topographischen Karte im Maßstab 1:25000 (TK25, Meßtischblatt). Mittlerweile liegen nahezu flächendeckend für Deutschland Daten für Viertelmeßtischblätter (= Quadranten) vor. Bei vielen regionalen Projekten liegen der Datenerhebung noch feinere Unterteilungen von Sechzehnteln (= Viertelquadranten) bis zu Vierundsechzigsteln (= Sechzehntelquadranten) des

TK25-Blattschnittes zugrunde.

Hinzu kommen zunehmend ergänzende Daten wie die Beschreibungen der Fundorte innerhalb eines Rasterfelds, die Zuordnung zu administrativen und naturräumlichen Einheiten, Angaben zur Populationsgröße am Fundort etc.

Koordinatenscharfe Fundangaben werden von den Bundesländern im Rahmen von Populationserhebungen gespeichert. Sie sind bisher jedoch in nur geringem Umfang an die Datenbank FLORKART weitergeleitet worden.

In Abb. 1 sind in Form einer Übersicht Inhalte und Anwendungsbereiche der Datenbank FLORKART dargestellt. Hier werden zum funktionalen Inhalt der Datenbank zusätzlich zu den eigentlichen Verbreitungsdaten (s.o.) auch die taxonomisch-nomenklatorische Referenz, geographische Referenzdaten zur Auflösung des Ortsbezugs sowie biologische und ökologische Charakterisierungsdaten der Gefäßpflanzen-Taxa hinzugezählt.

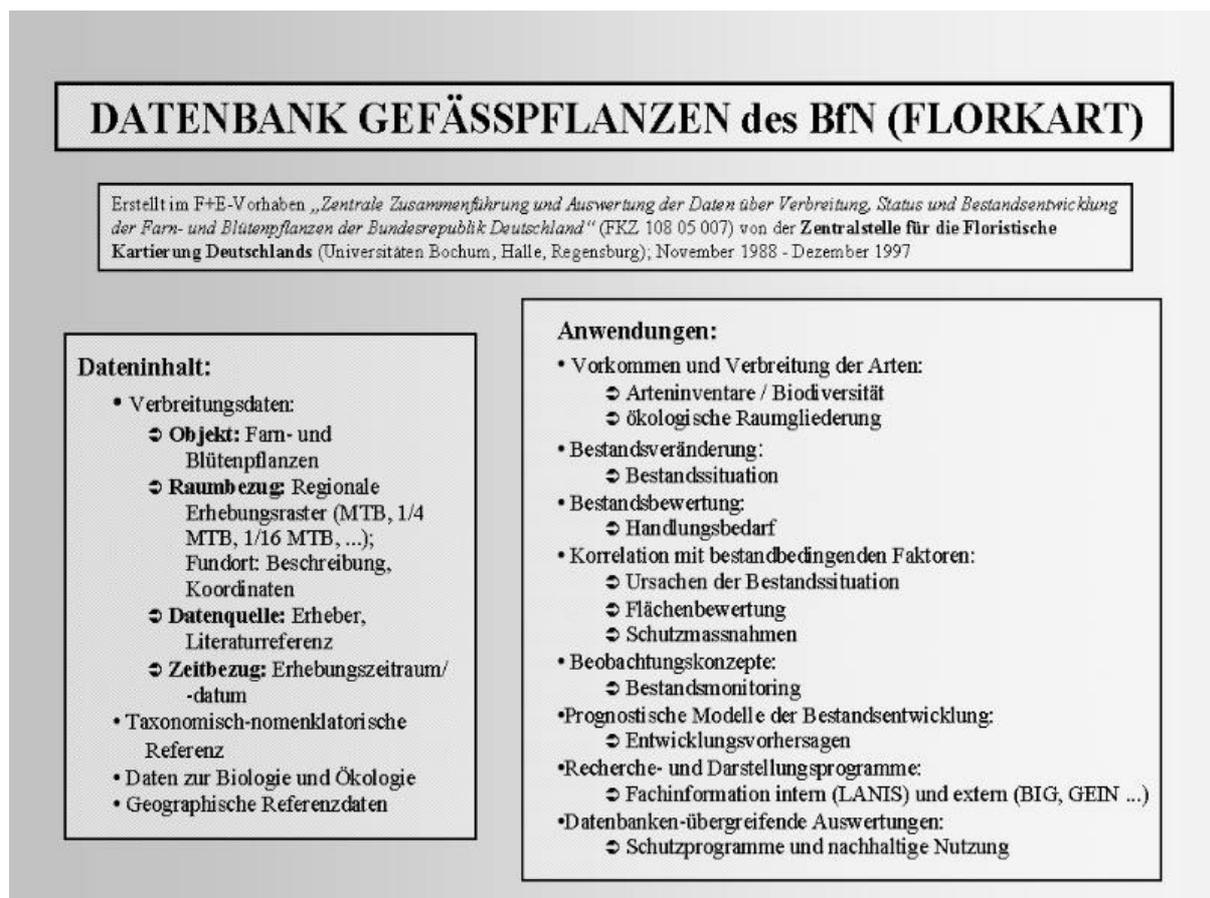


Abb. 1: Inhalte und Anwendungsgebiete der Datenbank FLORKART

Fig. 1: Content and potential of applications of the database FLORKART

Wichtige Anwendungen der Datenbank sind im benachbarten Kasten aufgelistet. Dialogorientierte Anwendungen, die dem Anwender im Intranet des BfN in Form von Abfrage- und Auswertungsmasken zur Verfügung stehen, ermöglichen Standardabfragen wie die Auflistung aller Nachweise in einem Rasterfeld, die Ausgabe von Verbreitungskarten oder Erzeugung von Reports auf 'Knopfdruck'.

Daneben gibt es spezielle Auswertungsverfahren, die unter Einsatz von GIS-Systemen und/oder Statistikpaketen komplexe Auswertungen realisieren. Als Beispiel für den ersten Typ sind in Abb. 2 ein Beispiel für eine Abfragemaske sowie in Abb. 3 eine Standard-Verbreitungskarte dargestellt.

Gat Epithet - Kürzel **Namensumfang der aktuellen Sippe im linken Block:**

ACERAS ANTHROPOPHORU 00017 -> Aceras anthropophorum (LINNAEUS) W. T. AITON FILIUS
 = Ophrys anthropophora LINNAEUS

MTB	Quad	U	Status	Zeitraum	AdB	Erheber	Projekt	alte Quellenangabe	L	T	D
4621	24		*	1985-1985	G	LucVo1	KASSEL	Nitsche GL		X	
4622	4		*	1975-1975	L		LITBOC		X	X	
4624	4		*	1989-1989	G	BaiEr1	HESSCH	Geländedaten Schnedler		X	
4625			*	1700-1899	G	EASEQ	ULMDAT	AHO 80			
4634	4		V_W+O	1815-1815	L	I1197	HALLHL	Herbar/Literaturdaten	X	X	X
4634	4		*	1842-1842	K	I1197	HALLHL	Herbar/Literaturdaten		X	X
4634	4		*	1832-1832	K	I1197	HALLHL	Herbar/Literaturdaten		X	X
4634	4		*	1815-1815	K	I1197	HALLHL	Herbar/Literaturdaten		X	X
4634	4		*	1857-1857	L	I1197	HALLHL	Herbar/Literaturdaten	X	X	X
4634	4		*	1842-1842	L	I1197	HALLHL	Herbar/Literaturdaten	X	X	X
4634	4		*	1200-1949	G	HerFI1	HALDAT			X	

Abfrage erstellen Karten (drucken, Layout, Symbole) Bericht (erstellen, ansehen) FlorEin Export Beenden

Messtischblatt-Nummer
Count: 42

Abb. 2: Beispiel einer Abfragemaske zur Datenbank FLORKART

Fig. 2: Sample search mask of the database FLORKART

Bundesamt für Naturschutz · Datenbank Gefäßpflanzen - Stand 1/1997

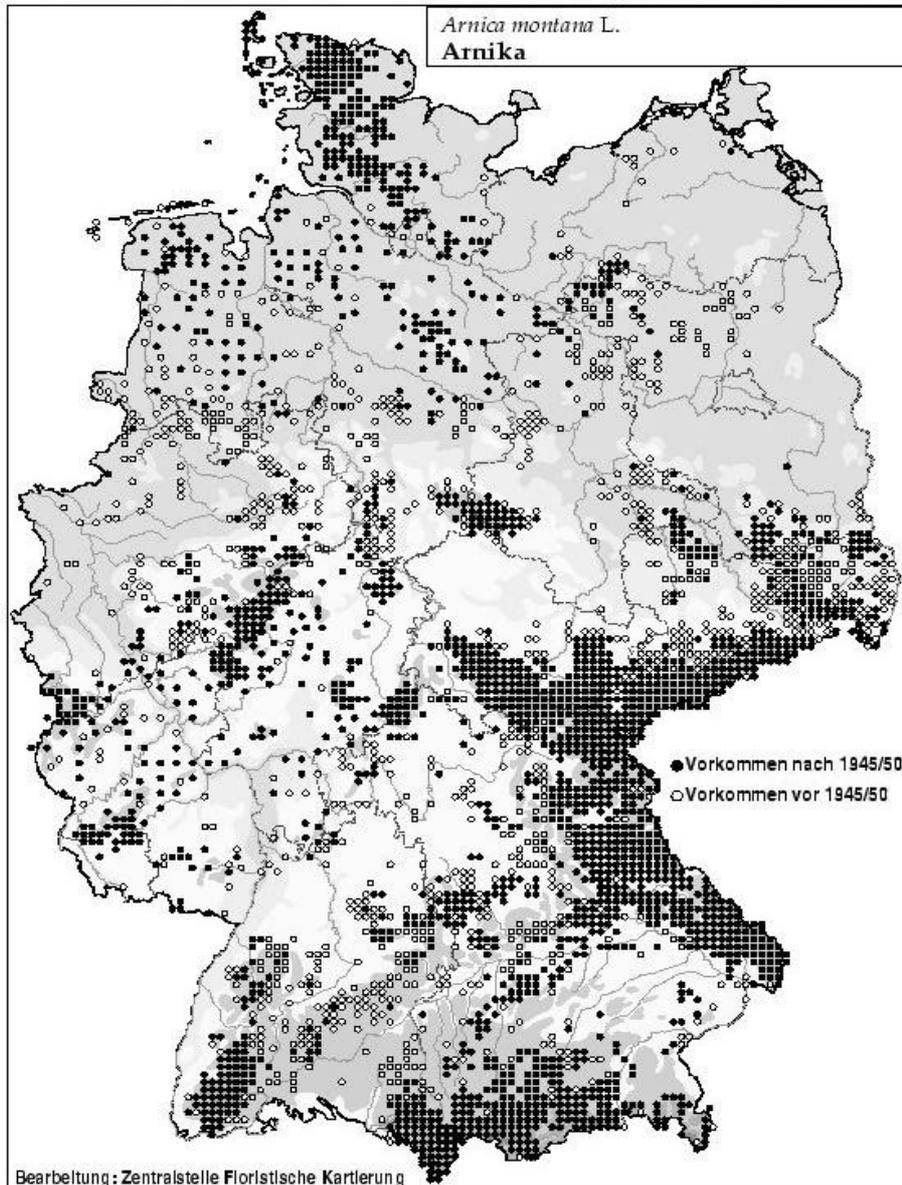


Abb. 3: Beispiel einer aus FLORKART generierten Standard-Verbreitungskarte

Fig. 3: Distribution map generated from the database FLORKART

Die Datenbank FLORKART lebt vom Datenfluß aus regionalen Erhebungsprojekten.

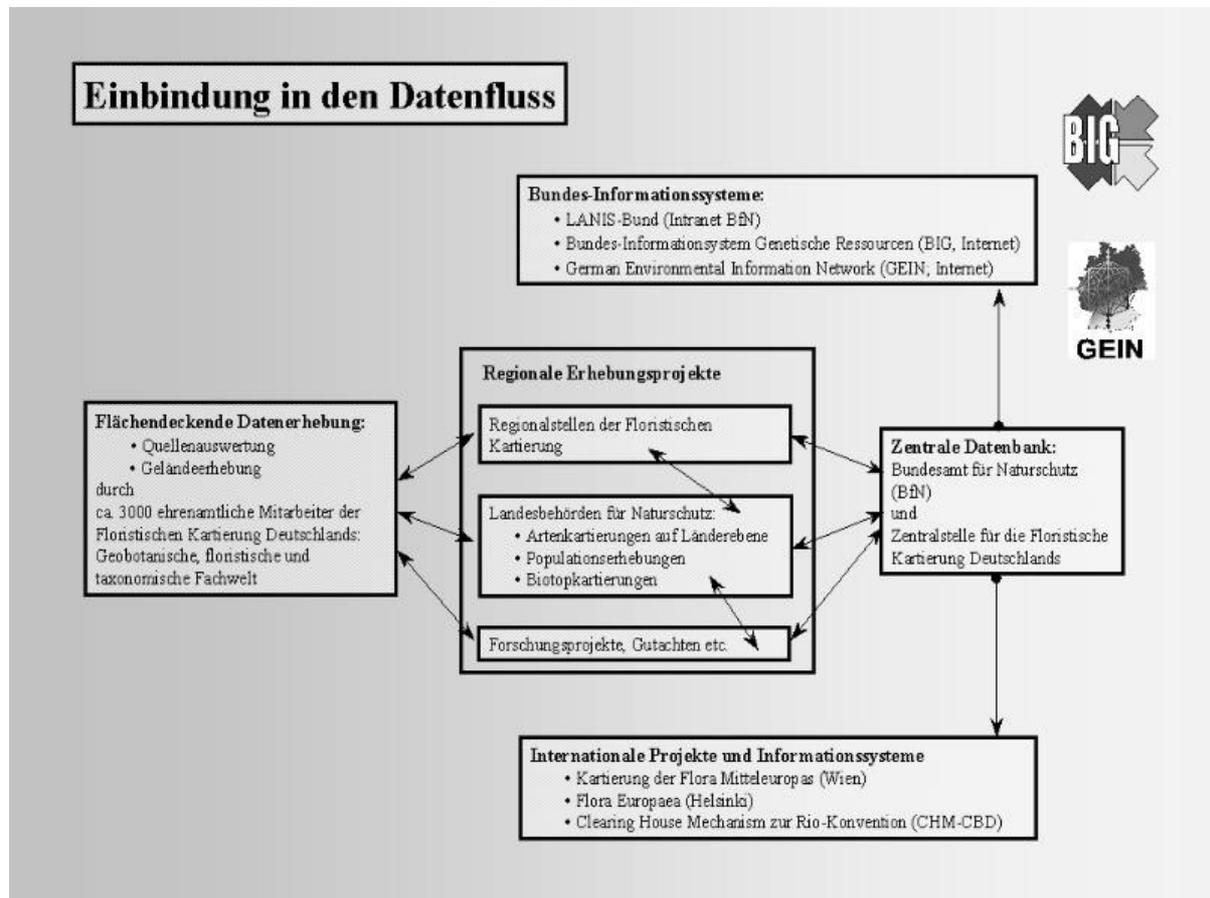


Abb. 4: Einbindung der Datenbank FLORKART am BfN in den bundesweiten Datenfluß und in nationale und internationale Auskunftssysteme

Fig. 4: Integration of FLORKART in national and international information systems

Für die technische Unterstützung des Datenflusses wurde zu Beginn der 90er Jahre das PC-Programm FLOREIN entwickelt. Es wurde im FE-Projekt „Datenbank Gefäßpflanzen“ als eigenständig laufende Datenbankanwendung auf der Basis von XBASE realisiert und allen regionalen Erhebungsprojekten kostenlos zur Verfügung gestellt. Das Programm sorgte für eine weitgehende Standardisierung sowohl im Bereich der Erhebungsmethodik als auch vor allem bei der Datenerfassung. Mit FLOREIN wurden die regionalen Projekte in die Lage versetzt, eigenständig ihre regionalen Kartierungsdatenbanken aufzubauen, zu pflegen und auszuwerten. Durch Import/Export-Funktionen können Daten zwischen einzelnen FLOREIN-Datenbanken sowie mit der zentralen ORACLE-Datenbank FLORKART ausgetauscht werden. Bis zum jetzigen Zeitpunkt hat sich FLOREIN als weit verbreiteter Standard bei nahezu allen großen regionalen Erhebungsprojekten etabliert. Die zukünftige Entwicklung sieht als Schwerpunkte die Portierung auf eine moderne 32 Bit Benutzeroberfläche (WINDOWS 98/NT) sowie die Vernetzung der regionalen Datenbanken und der zentralen Datenbank FLORKART via

Internet vor.

Datenbank FLORROT (Datenbank Rote Listen)

Die Datenbank FLORROT enthält die Daten der in der Roten Liste (BfN 1996) publizierten Einschätzung der Gefährdung aller in Deutschland nachgewiesenen Gefäßpflanzenarten. Die Gefährdung wird in einem in der Publikation näher erläuterten Bewertungsschema angegeben, welches verschiedene Gefährdungskategorien (0: ausgestorben, 1: vom Aussterben bedroht, 2: stark gefährdet, 3: gefährdet, 4: potentiell gefährdet) sowie die Kategorien G: Gefährdung anzunehmen und R: extrem selten beinhaltet.

Neben der Gefährdungseinstufung für das Bundesgebiet (Bundes-Rote-Liste) enthält die Datenbank auch die Gefährdungseinstufungen der einzelnen Bundesländer. Dazu kommen Informationen über den gesetzlichen Schutz nach CITES, Bundesartenschutzverordnung, Berner Konvention und FFH-Richtlinie.

Beim jetzigen Stand der Datenbank wurden für die synoptische Darstellung der Gefährdungseinschätzungen für Bund und Länder gelegentlich auftretende Unterschiede in der taxonomischen Auffassung der bewerteten Taxa durch Markierungszeichen dokumentiert. Die Weiterentwicklung dieser Datenbank soll die Verwaltung verschiedener taxonomischer Auffassungen im taxonomischen Kernsystem der taxonbezogenen Faktendatenbanken des BfN mit einschließen.

Datenbank WISIA (Informationssystem Internationaler Artenschutz)

In der Datenbank WISIA werden zur wissenschaftlichen Begleitung und Unterstützung der Tätigkeit des Bundesamtes bei der Umsetzung der internationalen Artenschutzabkommen Informationen über Nomenklatur, Systematik und Taxonomie, den Schutzstatus, die weltweite Verbreitung und die Ursprungsländer von handelsrelevanten Pflanzenarten bereitgehalten.

Referenzdaten

(1) Taxonomisch-nomenklatorische Referenz

In einem eigenen Forschungsprojekt des BfN wurde parallel zum Projekt 'Datenbank Gefäßpflanzen' eine neue Standardliste der Farn- und Blütenpflanzen Deutschlands (WISSKIRCHEN & HAEUPLER 1998) erstellt. Der Inhalt der publizierten Standardliste bildet das taxonomische Rückgrat aller taxonbezogenen Datenbestände im Bereich der Gefäßpflanzen. Die Weiterentwicklung dieser Datenbank wird auf der Basis von strukturellen Entwicklungen zur Verwaltung paralleler taxonomischer Auffassungen erfolgen, welche in einem Pilotprojekt am Beispiel der Standardliste der Moose Deutschlands erarbeitet werden.

(2) Geographische Referenzdaten

Als geographische Referenzen sind in Anwendungen der Datenbank FLORKART eingebunden:

- Namen der TK25-Kartenblätter
- Namen der administrativen Einheiten (Orts- Gemeinde- Kreis- Bundeslandnamen)

- Namen der Naturräume (Gliederung nach MEYNEN & SCHMIDTHÜSEN)

(3) Weitere Referenzlisten zur Charakterisierung von Fundorten

- Liste der Biotoptypen Deutschlands nach BOHN & PETERS (unpubl.)
- Liste der Pflanzengesellschaften nach OBERDORFER (1983)

Biologische und ökologische Charakterisierungsdaten

Durch die Bereitstellung von biologischen und ökologischen Charakterisierungsdaten von Gefäßpflanzentaxa wird die Möglichkeit eröffnet, kombinierte Auswertungen durchzuführen. Zum Beispiel lassen sich Artengruppen definieren, deren gemeinsames Vorkommen Aussagen zur ökologischen Raumbewertung oder die Darstellung der synthetischen Verbreitung von Pflanzengesellschaften ermöglicht. In die Datenbank integriert sind bislang folgende Datenbestände:

- Zeigerwerte nach ELLENBERG (1992)
- Biologische und ökologische Charakteristika nach FRANK und KLOTZ (1989)
- Pflanzensoziologische Bindung der Arten nach OBERDORFER (1983)

GIS - Daten

Sowohl für Auswertungs- als auch für Darstellungszwecke wird auf digitale Geodaten im Geographischen Informationssystem des BfN zugegriffen. Wesentliche Grundlagen für die Gestaltung von Verbreitungskarten bilden folgende Layer:

- TK25-Grundraster und TK25-Quadrantenraster
- Administrative Grenzen (Staatsgrenzen, Bundeslandgrenzen)
- Gewässer (Flüsse, Seen, Meeresflächen)
- Höhenschichten

Weitere Geodatensätze, die teilweise für Darstellungs- meist jedoch in der Verschneidung mit dem TK25-Raster für Auswertungszwecke genutzt werden, sind:

- Naturräumliche Gliederung nach MEYNEN & SCHMIDTHÜSEN
- CORINE Landcover (aktuelle Landnutzung als Auswertung von LANDSAT-Aufnahmen)

Klimadaten des Deutschen Wetterdienstes (DWD):

- Mittlere Jahrestemperaturen im Zeitraum 1960-1980
- Mittlere Januar-temperatur im Zeitraum 1960-1980
- Mittlere Julitemperatur im Zeitraum 1960-1980
- Mittlere Niederschlagsmenge im Zeitraum 1960-1980

Informationstechnische Basis der BfN-Datenbanken und Entwicklungsperspektiven

Das BfN betreibt seinen IT-Einsatz derzeit in einem NOVELL basierten Netzwerk. Zum Einsatz kommen Server mit dem Netzwerk-Betriebssystem NETWARE 4.12, denen als Massenspeicher ein nach RAID Level 5 konfiguriertes Array von Festplatten zur Verfügung steht.

Als Client-Rechner werden PCs mit WINDOWS 3.11 eingesetzt, die derzeit durch WINDOWS NT 4.0 abgelöst werden.

Netzwerk-Protokoll für die Anbindung der Arbeitsplatzrechner ist das NETWARE-Protokoll SPX/IPX. Neben den Arbeitsplatzrechnern werden zwei SUN Workstations über TCP/IP-Protokoll in das Netz integriert. Die Anbindung der Arbeitsplatzrechner zu den Workstations sowie an das Internet wird über einen Kommunikationsserver mit SPX-TCP Gateway vermittelt, auf dem via NOVELL Bordermanager auch der Zugang zu Internet-Adressen administriert wird. Unter NETWARE 4.12 ist als Datenbankmanagementsystem ORACLE version 7.3.1 installiert. Über das Netzprotokoll SQL*NET von ORACLE kann von Arbeitsplatzrechnern und Workstations aus auf die ORACLE Datenbank zugegriffen werden. Als Werkzeuge für die Entwicklung von Datenbankanwendungen werden ORACLE FORMS 4.5, DELPHI von INPRISE, VISUAL BASIC und VISUAL C++ von MICROSOFT eingesetzt.

Auf den Workstations ist als Geographisches Informationssystem ARC-INFO der Firma ESRI installiert. Von ARC-INFO kann per Datenbankschnittstelle über SQL*NET auf ORACLE zugegriffen werden. Als Peripheriegeräte im GIS-Bereich stehen ein DIN A0-Plotter und ein DIN A0 Digitalisiertisch zur Verfügung.

Zur Zeit unterhält das BfN noch keinen eigenen Internet-Server. Die Homepage des BfN (www.bfn.de) ist auf dem Server eines Dienstleisters aufgelegt. Für die Partizipation des BfN in internetbasierten Informationssystemen (GEIN, BIG) soll mittelfristig ein eigener Webserver betreiben werden.

Literatur

BUNDESAMT FÜR NATURSCHUTZ (BFN) (Hrsg.) (1996): Rote Liste gefährdeter Pflanzen Deutschlands. - Münster-Hiltrup (Landwirtschaftsverlag) - Schriftenreihe für Vegetationskunde 28, 744 S.

BERGMEIER, E. (1992): Grundlagen und Methoden Floristischer Kartierungen in Deutschland. - Floristische Rundbriefe, Beiheft 2, 146 S.

BOMMER, D.F.R. UND K. BEESE (1990): Pflanzengenetische Ressourcen - ein Konzept zur Erhaltung und Nutzung für die Bundesrepublik Deutschland. - Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Reihe A: Angewandte Wissenschaft, Heft 388, Teil I: 190 S., Teil II: 13 S.

- ELLENBERG, H., H.E. WEBER, R. DÜLL, V. WIRTH, D. WERNER UND D. PAULISSEN (1992): Zeigerwerte von Pflanzen in Mitteleuropa. 2. Aufl. Scripta Geobotanica 18: 1-258.
- FINK, H.G. UND R. MAY (1993): Dokumentation Genetischer Ressourcen und Naturschutz - Diskussionsbeitrag aus der Bundesforschungsanstalt für Naturschutz und Landschaftsökologie. In: Pflanzengenetische Ressourcen - Situationsanalyse und Dokumentationssysteme; Schriftenreihe des Bundesministers für Ernährung, Landwirtschaft und Forsten, Reihe A: Angewandte Wissenschaft, Heft 422, S.144-161.
- FRANK, D. UND S. KLOTZ (1990): Biologisch-Ökologische Daten für Flora der DDR. 2.Aufl. Wiss. Beitr. Univ. Halle 32: 1-167.
- HAEUPLER, H. UND P. SCHÖNFELDER (Hrsg.) (1989): Atlas der Farn- und Blütenpflanzen der Bundesrepublik Deutschland. 2. Aufl. - Stuttgart (Ulmer), 768 S.
- KORNECK, D., M. SCHNITTLER, F. KLINGENSTEIN, G. LUDWIG, M. TAKLA, U. BOHN, UND R. MAY (1998): Warum verarmt unsere Flora? Auswertung der Roten Liste der Farn- und Blütenpflanzen Deutschlands. - Schriftenreihe für Vegetationskunde 29:299-444.
- MAY, R. (1990): Die Datenbank der Florenkartierung: Weiterentwicklung der Datenstrukturen - Einbindung von PC's in ein integriertes EDV-Konzept. - Floristische Rundbriefe 23(2): 160-169.
- MAY, R. (1994a): Datenbank FLORKART. Entwicklungsstand, Import- und Ausgabefunktionen. Floristische Rundbriefe 28(2): 84-94.
- MAY, R. (1994b): Die Datenbank der Floristischen Kartierung: Ein Beispiel für die dezentrale Erhebung und zentrale Zusammenführung von raumbezogenen, naturschutzrelevanten Informationen. - In: KREMERS, H. (Hrsg.) (1994): Praxis der Umweltinformatik. Band 5 Umweltdatenbanken. Marburg. S. 155-175.
- MAY, R. UND W. SUBAL (1992): Einsatz von EDV für Kartierungsprojekte. - Floristische Rundbriefe Beiheft 2, 87-93.
- WISSKIRCHEN, R. UND H. HAEUPLER (1998): Standardliste der Farn- und Blütenpflanzen Deutschlands. - Stuttgart (Ulmer) 764 S.
- ZENTRALSTELLE FÜR DIE FLORISTISCHE KARTIERUNG DEUTSCHLANDS (Hrsg.) (1997): FLOREIN - Interaktives Programm zur Bearbeitung floristischer Daten. Version 5.0. Benutzerhandbuch. Bochum, Halle, Regensburg 1997.
- ZENTRALSTELLE FÜR DIE FLORISTISCHE KARTIERUNG DEUTSCHLANDS (Hrsg.) (1997): LITFAS - Bibliographischer Teil. Version 2.5. Benutzerhandbuch. Bonn 1997.

Bundesinformationssystem Genetische Ressourcen - Konzept und erste Ergebnisse

Federal Information System Genetic Resources – concept and first results

BARBARA BECKER¹

Abstract

In April 1998 a network project was started in Germany to establish the **Federal Information System Genetic Resources (BundesInformationssystem Genetische Ressourcen, BIG, <http://www.genres.de/big/>)** as an online information system on the Internet. In this project four German institutions pool their extensive databases on the wild flora of Germany, collections of botanical gardens, accessions of the German genebank, and other relevant databases, as well as their expertise in database management.

Taking advantage of synergy effects, the BIG project aims at developing an integrated information system on plant genetic resources that covers a wide range of taxonomic, genetic, biological, ecological, economic and geographical information. It will permit complex searches in heterogeneous, decentralized databases, and thus facilitate access to the actual germplasm. Technically this will be realized by a central view for accessing the on-line information system of various inter-linked, but independent databases. A user-friendly navigation system will be developed to search by keywords like plant names or attributes (e.g. breeding traits) or factual data like geographical references. In this way, BIG will provide a model online system for genetic resources with a high potential for application to other taxonomic databases.

The paper presents the conceptual approach of BIG, its underlying data model, the problem of hierarchical and non-hierarchical presentation of complex taxonomies of wild and cultivated taxa and their synonymies, the approach to establish an appropriate thesaurus, and the required technical solutions for integrating the decentralized, heterogeneous databases. Furthermore, some perspectives on assessing user needs, the market potential of BIG, and on ways of dealing with data property rights are given.

¹ Zentralstelle für Agrardokumentation und –information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

Einleitung

Vier Partnerinstitutionen mit umfangreichen Datenbanken zu Wild- und Kulturpflanzen und mit entsprechendem Know-How erstellen gemeinsam das BundesInformationssystem Genetische Ressourcen (BIG) als Online-System im Internet (<http://www.genres.de/big/>). BIG wird aus Mitteln des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie (BMBF) finanziert. Die Arbeit wurde im Frühjahr 1998 aufgenommen.

Beim **Bundesamt für Naturschutz (BfN)** in Bonn werden Datenbanken über Arteninventar (*in situ*), Verbreitung, Bestandssituation und Ökologie der einheimischen Wildpflanzen geführt sowie über schutz- und handelsrelevante Informationen zu den nach nationalem und internationalem Recht geschützten Arten.

Pflanzenbestände der Botanischen Gärten in Deutschland (*ex situ*) werden vom Verband Botanischer Gärten über den **Botanischen Garten der Ruhr-Universität Bochum (RUB)** erfaßt und dokumentiert.

Das **Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK)** in Gatersleben unterhält die Datenbank der fast 100 000 Kulturpflanzen-Akzessionen der Genbank (*ex situ*), erstellt eine Datenbank zu „Mansfeld`s World Manual of Agricultural and Horticultural Crops“ und bearbeitet grundlegende Fragen der Kulturpflanzen-taxonomie.

Die **Zentralstelle für Agrardokumentation und -information (ZADI)** in Bonn hat langjährige Erfahrung mit internet-gestützten Informationssystemen und bietet neben einer Reihe von fruchtartenspezifischen Datenbanken die zentrale Datenbank deutscher pflanzengenetischer Ressourcen im Internet an.

Abb. 1: Die BIG-Verbundpartner und ihr Beitrag zum gemeinsamen Informationssystem BIG

Fig. 1: The BIG network partners and their contributions to the integrated information system BIG

Ziel des Projektes ist die Entwicklung eines Informationssystems, das biologische, genetische, ökologische, ökonomische und geographische Informationen in einem System abbildet, um dadurch Synergieeffekte zu erzeugen und vernetzte Abfragen zu ermöglichen. Dabei sollen Techniken entwickelt werden, die heterogene, dezentrale Datenbanken durch einen zentralen Zugang in einem effizienten Online-Informationssystem verbinden, um sie dezentral abrufbar anzubieten. Es wird ein nutzerfreundliches Navigationssystem entwickelt, das den Einstieg über verschiedenartige Kriterien wie Pflanzennamen, züchterisch relevante Merkmale (Werteigenschaften), geographische Referenzen u.ä. (Faktendaten) erlaubt (Abb. 2).

Abb. 2: Zugang zu genetischen Ressourcen über das Informationssystem BIG

Fig. 2: Access to genetic resources via the information system BIG

BIG soll die in der Ressortforschung der Ministerien sowie der universitären und außeruniversitären Forschung vorhandenen Daten zu genetischen Ressourcen zusammenführen und zugänglich machen. Damit sollen sie Vertretern zuständiger Ämter (z.B. Naturschutzbehörden), Wissenschaftlern an Universitäten und in anderen Forschungseinrichtungen sowie der interessierten Öffentlichkeit wie Nichtregierungsorganisationen zur Verfügung stehen. Gleichmaßen wird BIG für die Privatwirtschaft von Bedeutung sein, insbesondere für Züchtungsfirmen sowie die Naturstoffindustrie, die mit pflanzlichen Inhaltsstoffen arbeitet.

BIG ist so konzipiert, daß es taxon-bezogene Informationen generell erschließen kann und abrufbar macht. Das Konzept und die eingesetzten Methoden und Werkzeuge sind daher auch auf andere Organismen übertragbar. Damit wird modellhaft ein Verfahren zur Online-Information mit einem hohen Übertragungspotential auf weitere Datenbestände bereitgestellt.

Vorgehensweise

Die Umsetzung dieser Ziele und Aufgaben wurde im ersten Projektjahr auf zwei Ebenen in Angriff genommen: (i) durch den Aufbau von Kommunikationsstrukturen und gemeinsame Aktivitäten der Verbundpartner sowie (ii) durch die BIG-spezifische Weiterentwicklung der Datenbanken in den Institutionen des Projektverbundes.

Zur Etablierung von BIG als Verbund wurde in dieser ersten Phase besonderes Gewicht auf den Aufbau effizienter Kommunikation untereinander und zielorientierte Teamarbeit an dem gemeinsamen Produkt BIG gelegt sowie auf die Bildung einer „Corporate Identity“ und die einheitliche Außendarstellung.

In allen Partnerinstitutionen wurden die Einzeldatenbanken weiter gepflegt und die technische Implementierung der Datenbankvernetzung vorbereitet. Insgesamt werden hierfür in den vier beteiligten Institutionen zwölf Mitarbeiter/innen aus BIG-Mitteln finanziert, ca. zur Hälfte für die taxonomisch-inhaltliche Datenpflege und zur anderen Hälfte für die informationstechnologische Umsetzung. In allen Institutionen sind weitere nicht aus BIG-Mitteln finanzierte Mitarbeiter/innen mit dem Vorhaben befaßt.

Ergebnisse

Die wichtigsten Ergebnisse zur Entwicklung von BIG sind der Entwurf eines zentralen Datenmodells, Strategien zum Aufbau einer zentralen Metadatenbank zur Taxonomie (BIGTAX) und eines Thesaurus, die Einigung auf ausgewählte Tabellen zur Entwicklung des ersten BIG-Prototyps sowie ein IT-Konzept zur Datenbankvernetzung.

Datenmodell

Das Datenmodell (Abb. 3) liegt in Form eines Entity-Relationship (ER)-Diagramms vor, d.h. als konzeptuelles Modell. Es baut auf dem zunächst noch theoretischen Anforderungsprofil auf, wie es sich aus dem Projektantrag bzw. aus der Sicht der BIG-Mitarbeiter/innen darstellt. Das Anforderungsprofil als zentrale Komponente des externen Modells wird im zweiten Projektjahr durch eine Zielgruppenerhebung zum Nutzerbedarf präzisiert.

Die Grundannahme für das Anforderungsprofil ist der Einstieg in BIG über Pflanzennamen im weitesten Sinne (Taxonomie) und / oder über Suchbegriffe zu Faktendaten über Pflanzen (Thesaurus). In dem konzeptuellen Modell wird dies dadurch umgesetzt, daß die „Information“ drei Teilkomponenten verknüpft: „Pflanzename“, „Thesaurus“ und „Source of Information“. Diese letzte Kategorie trägt der Tatsache Rechnung, daß erstens die „Information“ aus den verschiedenen Teildatenbanken gespeist wird und dort die Details abzufragen sind, und daß damit zweitens eine eindeutige Quelle und die dort verwendete taxonomische Umgrenzung zu dem jeweiligen Pflanzennamen zugeordnet werden kann.

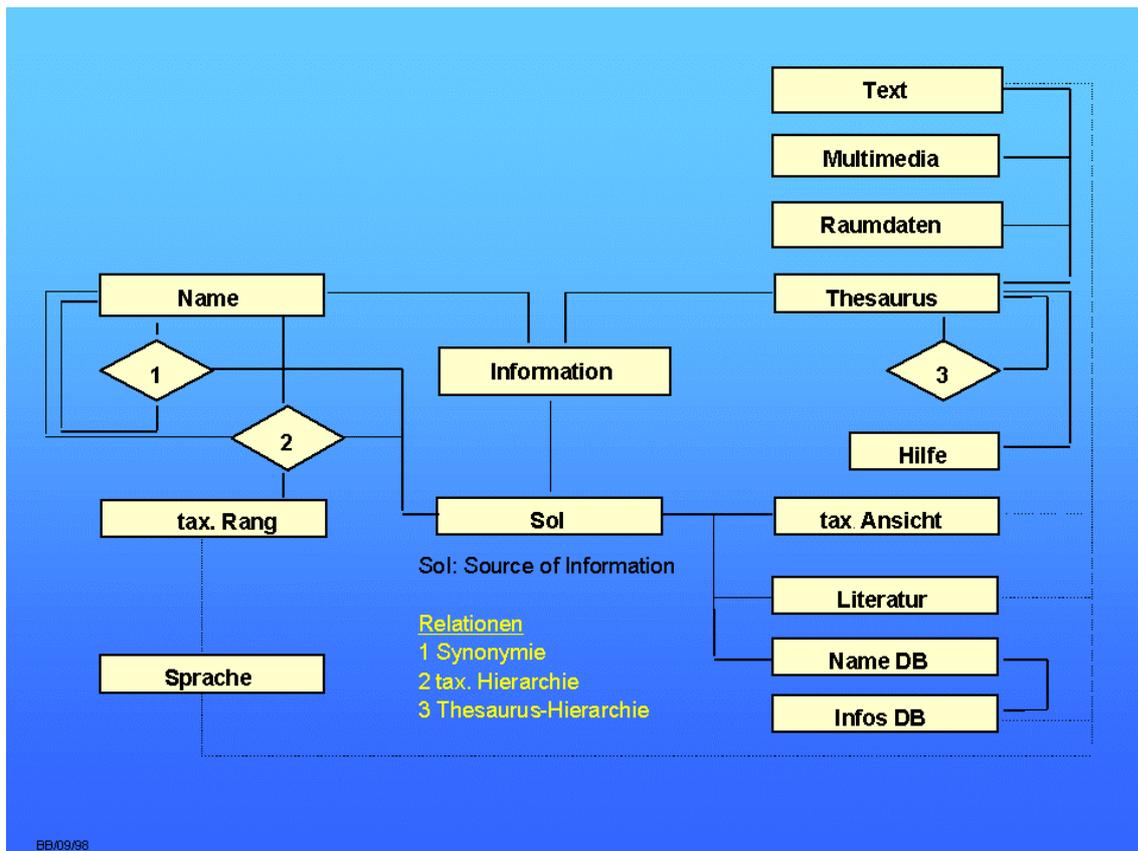


Abb. 3: Das konzeptuelle Datenmodell von BIG (BOOS 1998)

Fig. 3: The conceptual data model of BIG

Zur Übersetzung des konzeptuellen Modells in ein internes Modell, d.h. die detaillierte, programm-abhängige Definition von Tabellen und Verknüpfungsvorschriften wurden für die Entwicklung des BIG-Prototyps zunächst die Kategorien Gattung, Art und Trivialname ausgewählt und vereinbart, mit den Familien Poaceae (Gramineae) und Brassicaceae (Cruciferae) zu beginnen, zumindest bis zur Ebene der Art. Dem Datenmodell entsprechend muß die hierfür aufzubauende zentrale Metadatenbank (BIGTAX) die Angaben zur taxonomischen Hierarchie, zur Synonymie sowie für die „Source of Information“ zunächst den Namen der Quelldatenbank enthalten und wiedergeben. Aus der Kategorie der Faktendaten sollen in dem BIG-Prototyp die Verbreitungsdaten abgebildet werden, zunächst nur als Ergebnis (noch nicht als Suchbegriff) in Form von Links zu vorhandenen Verbreitungskarten und als Listen der (Bundes)Länder sowie dem Raster der Topographischen Karten im Maßstab 1:25.000, wo das gesuchte Taxon georeferenziert dokumentiert ist.

Taxonomie

Zur angemessenen Abbildung taxonomischer Angaben in dem Informationssystem BIG müssen insbesondere zwei Probleme gelöst werden: Da die taxonomische Umgrenzung gleichnamiger botanischer Taxa von verschiedenen Autoren unterschiedlich gehandhabt werden kann, ist ein Verweis auf den Autor der Information zur eindeutigen Interpretation erforderlich („potential taxa“, BERENDSOHN 1995). Dies wird in dem o.a. Datenmodell durch den Bezug zur „Source of Information“ gelöst. Dieses Modell erlaubt damit die Abbildung von Synonymien innerhalb der Partnerdatenbanken, die keine zentrale Datenbank mit einer vollständig vereinheitlichten Synonymie erfordert.

Die eindeutige Abbildung von Taxa (taxonomische Umgrenzung durch Darstellung der Synonymik und systematisch untergeordneter Taxa) und ihrer Synonyme in einer Datenbank setzt in der Regel eindeutige hierarchische Beziehungen voraus. Dies ist für den Bereich der Wildpflanzen auch weitgehend der Fall und in dem „International Code of Botanical Nomenclature“ (ICBN) klar definiert. Für den Bereich der Kulturpflanzen hingegen sind hierarchische Verwandtschaftsbeziehungen zwischen Taxa, insbesondere unterhalb der Artebene, oft nicht gegeben (vgl. HETTERSCHEID UND BRANDENBURG 1995). Eine Überarbeitung des „International Code of Nomenclature for Cultivated Plants“ (ICNCP) wird gegenwärtig vorbereitet. Die für BIG angestrebte Lösung zielt auf ein integriertes Modell für Wild- und Kulturpflanzen anstelle einer künstlichen Dichotomie zwischen beiden Gruppen, die nicht biologisch, sondern durch die Züchtungspraxis, insbesondere im Zierpflanzen Sektor, begründet ist (OCHSMANN 1998). Die taxonomische Hierarchie und Synonymie für die in BIG vernetzten Datenbanken muß insbesondere die Genbank-Akzessionen abbilden können, die eine Art Zwischenstellung zwischen den beiden Codes einnehmen, da sie u.a. Wildformen von Kulturpflanzen und Landsorten umfassen, die in keinem der beiden Codes hinreichend berücksichtigt sind. Daneben müssen langfristig neben den Volksnamen auch die Handelsnamen abgebildet werden können, für die nicht immer eindeutige Beziehungen zu den Sortennamen bestehen.

Die Entwicklung angemessener Konzepte zur Abbildung taxonomischer Verwandtschaftsbeziehungen und Synonymien in Datenbanken stellt international eine Herausforderung dar. Die verschiedenen Modelle und Konzepte national und international werden im Rahmen von BIG aufmerksam verfolgt (vgl. BERENDSOHN ET AL 1998), und der Austausch mit ähnlichen Initiativen wird gesucht und gepflegt, damit BIG einen eigenständigen Beitrag zu dieser Entwicklung leisten kann. Die bislang etablierten Systeme bei den Projektpartnern im BIG sind taxonorientiert. Die Etablierung konzeptbasierter Systeme steht noch am Anfang und wird im BIG-Partner BfN im Projekt TaxLink am Beispiel der Moose prototypisch realisiert. Der Bedarf an konzeptbasierten Systemen mit der Möglichkeit der Verwaltung mehrerer taxonomischer Auffassungen zum gleichen Organismennamen wurde in der BIG-TAX Arbeitsgruppe bereits diskutiert und als notwendige Voraussetzung für die Interoperabilität verteilter, heterogener Organismendatenbanken beschrieben. Dies ist bei der Entwicklung der BIG-Schnittstellen zu berücksichtigen. Bereits etablierte Standards wie Pflanzennamen, Autorennamen, Literatur oder geografische Daten sollten übernommen werden (BIERMANN UND MAY 1998).

Thesaurus

Zur Erschließung der Faktendaten ist die Erstellung eines Thesaurus erforderlich, der in einem hierarchischen System die Begriffe abbildet, nach denen in BIG gesucht werden soll. Der BIG-Thesaurus spiegelt die Semantik des BIG wieder. Er wird von einer internen Arbeitsgruppe festgelegt bzw. ergänzt. Er erfüllt zwei wichtige Funktionen. Erstens dient er als Grundlage zur Definition der BIG-Begriffe und somit der Definition der BIG-Syntax.

Zweitens ist er der Kern der zentralen BIG-Oberfläche und dessen automatischen Abfragegenerators (JIMÉNEZ KRAUSE ET AL 1999). Nach dem gegenwärtigen Stand der Diskussion wird die Festlegung der Begriffe für den BIG-Thesaurus pragmatisch erfolgen auf der Grundlage der bereits in den Einzeldatenbanken verwendeten Felder. Dabei soll so weit wie möglich auf bestehende Thesauri zurückgegriffen werden und diese den BIG-Erfordernissen angepaßt werden. Um die Breite der in BIG erfaßten Daten abzudecken und um größtmögliche nationale und internationale Kompatibilität zu gewährleisten, müssen verschiedene deutsch- und englischsprachige fachspezifische Thesauri geprüft und kombiniert werden. Dies sind in erster Linie der Umweltdatenkatalog, für dessen Entwicklung Deutschland das Mandat der Europäischen Umweltagentur innehat, der ZADI-eigene Agrokat sowie AGROVOC und ENVOC, und für kulturpflanzen-spezifische Begriffe die Deskriptorenlisten von FAO und IPGRI sowie deren Verschlagwortung von Passportdaten. Nach einer ersten Überprüfung hat sich bereits gezeigt, daß die in diesen Thesauri verfügbaren Begriffe die Belange von BIG nur unzulänglich abbilden und entsprechend ergänzt und modifiziert werden müssen.

Neben der taxonomischen Metadatenbank BIGTAX wird dieser Thesaurus den Kern der zentralen BIG-Suchmaschine bilden.

Raum- und Bild-Daten

Zur Abbildung von Verbreitungs- und Herkunftsdaten der in BIG erfaßten Taxa ist die Verknüpfung mit einem internet-fähigen Geographischen Informationssystem (GIS) vorgesehen. Deshalb wurde zunächst entsprechende Software getestet, um für BIG ein geeignetes Produkt zu finden. Nach entsprechenden Recherchen wurde das Produkt ArcView und die Internetkomponente IMS von der Firma ESRI unter dem Betriebssystem UNIX als Testlizenz beschafft. Ein Prototyp für die Abfrage und Darstellung von Verbreitungsdaten der Floristischen Kartierung via Internet-Protokoll-basierter Abfrage wurde entwickelt. Der Test hat ergeben, daß der eingeschlagene Lösungsweg über Internet Protokoll-basierte Abfragen gut geeignet ist. In der konkreten Anwendung stellte sich jedoch heraus, daß die getestete GIS-Software zu langsam ist. Als weitere Option wurde die bei der Gesellschaft für mathematische Datenverarbeitung (GMD) entwickelte Software DESCARTES geprüft, die jedoch bisher nur zum Handling von 500 Objekten ausgelegt ist, und damit ebenfalls nicht für die Anwendung in BIG in Frage kommt. Weitere Optionen werden geprüft. Die Darstellung von

Abbildungen zu in BIG erfaßten Taxa ist vorgesehen und durch die Bereitstellung digitalisierter Bilder in den Einzeldatenbanken bereits initiiert.

IT-Konzept

Das IT-Konzept zur Datenbankvernetzung (FRIEDRICH UND HEYM 1998) hat die Entwicklung einer datenbank- und rechnerübergreifenden, fachspezifischen Suchoberfläche zum Ziel. Der Einstieg soll über eine zentrale Seite erfolgen, die ein am Endnutzer orientiertes, strukturiertes und detailliertes Retrieval ermöglicht. Dabei soll auf die dezentralen Angebote der BIG-Partnerinstitutionen zugegriffen werden und deren Eigenständigkeit gewährleistet werden. Diese Anforderungen können weder durch zentrale Datenbanken noch durch die Nutzung von herkömmlichen Suchmaschinen erfüllt werden.

Die angestrebte IP-Lösung setzt voraus, daß alle teilnehmenden Datenbanken via Internet-Protokoll (IP) kommunizieren (d.h. sie können via WWW sowohl abgefragt werden als auch Daten zurückliefern), und daß die dezentralen Datenbanken lediglich transparent und zentral abgefragt werden. Diese Transparenz der Datenbanken ohne Interoperabilität ist für das Ziel von BIG ausreichend. Bei der IP-basierten Lösung werden die dezentralen Datenbanken über die logische Ebene des WWW-Gateway angesprochen, dessen Installation bei den teilnehmenden Datenbanken vorausgesetzt wird. Aus diesem Grund ist das Kommunikationsprotokoll unterhalb des WWW-Gateway (Z39.50 o.ä.) für BIG unerheblich.

Die Abfrage der BIG Datenbanken erfolgt über eine zentrale Suchoberfläche, die den Nutzer bei der Eingabe von Suchkriterien unterstützt. Die zentrale Suchoberfläche basiert auf einer Metadatenbank, die Informationen darüber enthält, in welcher Datenbank es Daten gibt zu spezifischen Taxa, inkl. Synonymie (BIGTAX) und Deskriptoren (Thesaurus).

Für die IP-basierte Lösung benötigt man ferner eine festdefinierte Schnittstelle, die die Kommunikation in beiden Richtungen zwischen der zentralen Suchoberfläche und den dezentralen Datenbanken regelt. Die Festlegung auf eine für BIG angemessene Schnittstelle wird gegenwärtig analysiert (Abb. 4).

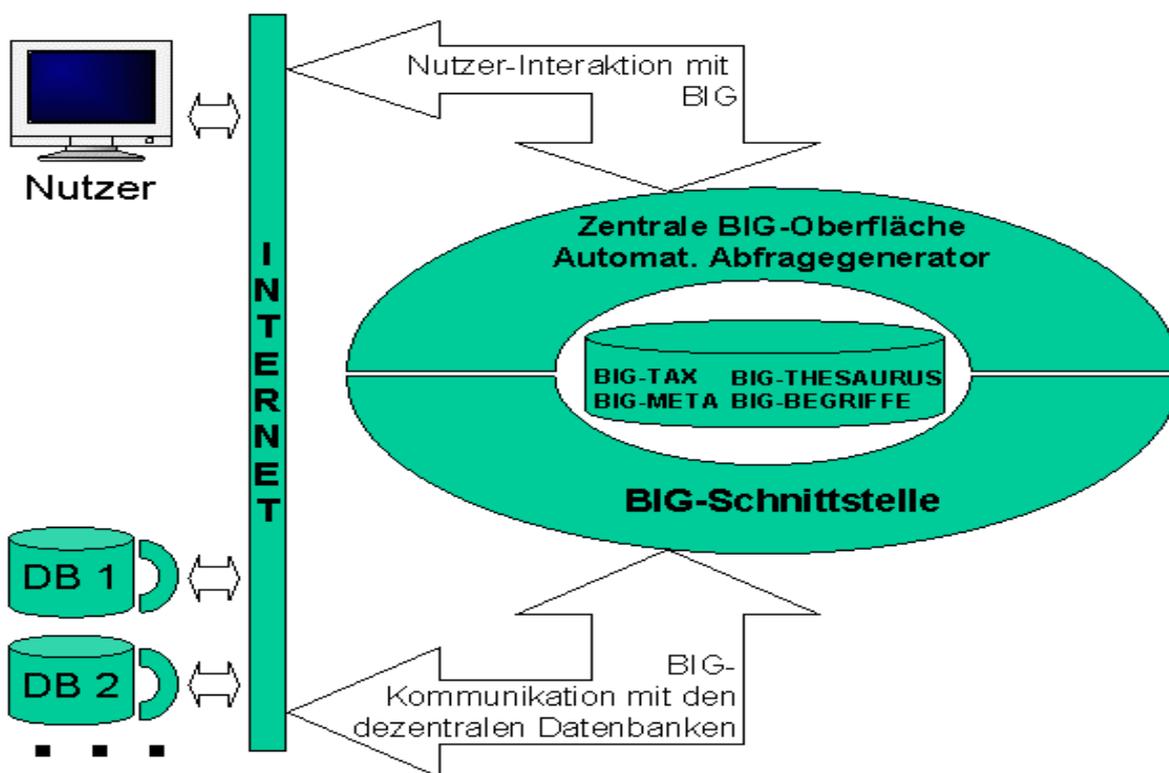


Abb. 4: Konzept zur informationstechnologischen Umsetzung der Suche in den dezentralen heterogenen Datenbanken des BIG-Verbundes (JIMÉNEZ KRAUSE ET AL 1999)

Fig. 4: Internet retrieval concept in the decentralized, heterogenous databases of the BIG network

Perspektiven

BIG hat zunächst eine Laufzeit von drei Jahren (1998 - 2001), danach ist im Falle einer positiven Evaluierung der Aufbauphase eine zweijährige Implementierungsphase vorgesehen. Um darüber hinaus das Angebot von BIG langfristig zu sichern, hat BIG den Auftrag des BMBF, einen Businessplan zu erstellen, der einerseits das Nutzerprofil klar umschreibt und andererseits Strategien zur potentiellen (Teil)Kommerzialisierung aufzeigt.

Die Erhebung zum Nutzerprofil wurde durch die ausführliche Beratung dieses Themas im projektbegleitenden Beirat² begonnen. Sie soll durch einen Workshop mit den potentiellen Zielgruppen im Juni 1999 in Gatersleben soweit zum Abschluß gebracht werden, daß daraus das Nutzerprofil für den Businessplan hervorgeht und für das externe Modell das Anforderungsprofil klar beschrieben werden kann. Die weitere Beobachtung und Präzisierung des

² Der von BMBF berufene BIG-Beirat setzt sich aus je drei Vertretern aus Wissenschaft und Wirtschaft zusammen.

Nutzerprofils wird anhand der Zugriffe auf den BIG-Prototypen und entsprechende Rückmeldungen erfolgen.

Zur Entwicklung von Kommerzialisierungsstrategien ist zuerst sorgfältig die Entwicklung der Internet-Ökonomie zu verfolgen, bei der bisher nicht eindeutig erkennbar ist, ob die Tendenz der Kommerzialisierung von im Internet angebotenen Informationen eher zunimmt oder abnimmt (vgl. SHAPIRO UND VARIAN 1998).

Im Hinblick auf die in BIG erfaßten Daten sind bei ihrer potentiellen Kommerzialisierung zwei Gesichtspunkte zu berücksichtigen. Erstens steht BIG im Spannungsfeld zwischen einer Dienstleistung als öffentliche Aufgabe im Mandatsbereich der Ressorts, denen die BIG-Partnerinstitutionen angehören und andererseits dem Informationsservice als privatwirtschaftlich handelbares Produkt. Zwischen beiden berechtigten Interessen wird unter Berücksichtigung der Eigentumsrechte an den Daten und den Kosten für die Informationsbereitstellung sorgfältig abgewogen werden müssen. Zweitens steht BIG international im Kontext dynamischer Veränderungsprozesse im Hinblick auf die ökonomische Bewertung sowohl von geistigem Eigentum als auch von genetischen Ressourcen bzw. genetisch kodierter Information (VON BRAUN UND VIRCHOW 1998). Hier müssen die Entwicklungen bei der Welthandelsorganisation WTO, insbesondere im Hinblick auf das TRIPS-Abkommen (Agreement on Trade-related Aspects of Intellectual Property Rights), bei der World Intellectual Property Organization (WIPO) und bei UPOV (Union for the Protection of Plant Varieties) sorgfältig beobachtet und beachtet werden. Darüber hinaus wird zur Anpassung an die Bestimmungen der Konvention zur Biologischen Vielfalt (CBD) gegenwärtig das International Undertaking for Genetic Resources for Food and Agriculture der FAO neu verhandelt. Mit einem Abschluß der Verhandlungen wird bis Mitte 1999 gerechnet, auch wenn der Erfolg bzw. das Ergebnis dieser Verhandlungen bisher noch nicht klar absehbar ist.

Als Informationssystem über physisch vorhandene genetische Ressourcen in der Bundesrepublik Deutschland (*ex situ* und *in situ*) steht BIG an der Schnittstelle zwischen dem Zugang zu der Information über genetisches Material und dem Zugang zu dem Material selbst. Für den Zugang zu dem Material in *Ex-situ*-Sammlungen werden gegenwärtig zur Umsetzung der Konvention zur Biologischen Vielfalt verschiedene Material Transfer Agreements (MTAs) ausgehandelt. Modellhaft sei hier der zur International Centres Week im Oktober 1998 verabschiedete MTA-Entwurf der Consultative Group for International Agricultural Research (CGIAR) genannt. Im Rahmen von BIG ist zu prüfen, wie weit parallel zu solchen Material Transfer Agreements entsprechende „Information Transfer Agreements“ (ITAs) ausgehandelt werden müssen. Dies betrifft insbesondere Informationen über genetisches Material in *Ex-situ*-Sammlungen, dessen Herkunft aus anderen Ländern ist, die damit seit Inkrafttreten der Konvention zur Biologischen Vielfalt (CBD) nationale Eigentumsrechte daran geltend machen können.

In diesem Zusammenhang stellt sich ebenfalls die Frage, wie weit die zunehmend an Bedeutung gewinnende Information über Patente an in BIG erfaßten Pflanzen in diesem Informationssystem abgebildet werden kann und soll. Erste Recherchen dazu haben ergeben, daß seitens der Patentämter (Deutsches und Europäisches Patentamt) sowie durch eigens damit beauftragte Institutionen (z.B. FIZ Karlsruhe) umfangreiche Informationen bereitgestellt werden, diese aber in der Regel kostenpflichtig sind. Spezifische Recherchen für die Pflanzenzüchtung werden seit kurzem durch die Patentstelle der GVS (Gesellschaft für Erwerb und Verwertung landwirtschaftlicher Sorten) durchgeführt, insbesondere im Hinblick auf die potentielle Anmeldung neuer Patente. Vorläufiges Fazit dieser Recherchen ist, daß die Abbildung von Patentinformation in BIG nicht vordringlich ist, da es spezialisierte Informationsanbieter gibt und die Erhebung dieser Information gegenwärtig den Rahmen der BIG-Partner übersteigt. Die Entwicklungen auf diesem Gebiet werden jedoch weiter beobachtet und bei dem weiteren Aufbau von BIG berücksichtigt.

Zusammenfassend läßt sich feststellen, daß sich BIG sowohl im Hinblick auf die notwendige Entwicklung informationstechnologischer Lösungen als auch bezüglich der Inhalte (Taxonomie und Thesaurus) anspruchsvolle Ziele gesetzt hat. Um diese Ziele zu erreichen, wird bewußt der Austausch mit verwandten Initiativen gesucht und mit zunehmender Entwicklung des BIG-Prototyp gezielt weiter ausgebaut. Wenn die in BIG gesetzten Ziele erreicht werden, ist damit ein Modell entwickelt, das international beispielhaft sein wird. Ebenso werden hohe Erwartungen an die zu entwickelnden Kommerzialisierungsstrategien gestellt, da es hierfür bisher keine Vorbilder an der Schnittstelle von öffentlichem und privatwirtschaftlichem Gut gibt, die den komplexen nationalen und internationalen politischen Rahmenbedingungen Rechnung tragen.

Literatur

- BERENDSOHN, W. G. (1995): The concept of "potential taxa" in databases. *Taxon* 44: 207-212.
- BERENDSOHN, W. G., C.L. HÄUSER UND K.-H. LAMPE, (1998). Biodiversitätsinformatik. Nationale und internationale Strukturen. Studie im Auftrag des BMBF, DLR-Projekträger Umweltsystemforschung, 41 S.
- BIERMANN, N. UND R. MAY, (1998): Dienstreise-Bericht über die Teilnahme am „Short Course“ über „Species Diversity Information Systems“ vom 7. - 11. 9. 1998.
<http://www.dainet.de/genres/big/interna/daten/bericht.htm>
- BOOS, E., (1998): Vorschlag für ein ER-Modell zur Lösung der BIG-Anforderungen.
<http://www.dainet.de/genres/big/interna/daten/boos-modell.htm>
- BRAUN, J.V, UND D. VIRCHOW (1998): Pflanzengenetische Ressourcen zwischen Angebot und Nachfrage – Entwicklung institutioneller Rahmenbedingungen für Konservierung und Nutzung. *Ber. Ldw.* 76: 74-86.
- FRIEDRICH, H. UND J. HEYM (1998): IT-Konzept zur Vernetzung der Datenbanken in BIG.
<http://www.dainet.de/genres/big/interna/daten/IT-Kon-r.htm>

- HETTERSCHIED, W.L.A. UND W.A. BRANDENBURG (1995): Culton versus taxon: conceptual issues in cultivated plant systematics. *Taxon* 44: 161-175.
- JIMÉNEZ KRAUSE, J. D., E. BOOS UND B. BECKER (1999): Das BIG-Konzept der zentralen Abfrage dezentraler Datenbanken. Beitrag zum Workshop des Arbeitskreises „Umweltdatenbanken“ der Fachgruppe 4.6.1 „Informatik im Umweltschutz“ der Gesellschaft für Informatik e.V., 10./11. Juni 1999 in Karlsruhe.
- OCHSMANN, J., (1998): Anmerkungen zur Kulturpflanzentaxonomie.
<http://www.dainet.de/genres/big/interna/daten/kultax.htm>
- SHAPIRO, C. UND H.R. VARIAN (1998): *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press. Boston, Massachusetts. 352 S.

Pflanzengenetische Ressourcen im Informationssystem Genetische Ressourcen (GENRES) – Online Demonstration

Plant Genetic Resources in the Information System on Genetic Resources (GENRES) – Online presentation

EBERHARD MÜNCH¹

Zusammenfassung

Das **Informationssystem Genetische Ressourcen (GENRES)** ist ein vom Informationszentrum für Genetische Ressourcen (IGR) der Zentralstelle für Agrardokumentation und -information (ZADI) betriebenes Verweissystem im Internet. Die Rubrik „Pflanzengenetische Ressourcen“ (PGR) ist nach den Hauptpunkten „PGR in Deutschland“, „PGR in anderen Staaten“ (geographisch), „PGR nach Fruchtarten“ und „PGR Online-Datenbanken“ unterteilt, welche weiter gegliedert sind. Im Beitrag wird näher auf die einzelnen Untergebiete eingegangen.

Abstract

The **Information System for Genetic Resources (GENRES)** is a reference system on the Internet created and maintained by the Information Centre for Genetic Resources (IGR) of the German Centre for Documentation and Information in Agriculture (ZADI). The category Plant Genetic Resources (PGR) contains data on PGR in Germany, PGR in other countries (geographic), PGR by crops and PGR online databases, which are then subdivided into further topics. The paper gives detailed information on the above mentioned subfields.

Das Informationssystem Genetische Ressourcen (GENRES)

Das Informationssystem Genetische Ressourcen (GENRES) ist ein vom Informationszentrum für Genetische Ressourcen (IGR) der Zentralstelle für Agrardokumentation und -information (ZADI) betriebenes Verweissystem im Internet. Analog zum Deutschen Agrarinformationsnetz (DAINet) als wichtigstem Produkt der ZADI ist es ein hypermediales System, welches Agrarinformationen im Internet über fach- und problemspezifische Verknüpfungen und Verweise zugriffsbezogen abbildet. Während der thematische Umfang des DAINet alle Bereiche der Ernährung, Land- und Forstwirtschaft erfaßt, ergibt sich das fachliche Spektrum von GENRES aus der spezifischen Aufgabenstellung des IGR.

¹ Zentralstelle für Agrardokumentation und -information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

Gegenwärtig wird im GENRES auf ca. 1.200 Internetquellen verwiesen. Monatlich erfolgen 30.000 bis 35.000 Zugriffe auf das System. Die Internetadresse der deutschen Version ist <http://www.genres.de>, die der englischen <http://www.genres.de/genres-e.htm> (Abb. 1).



Abb. 1: GENRES-Homepage

Fig. 1: GENRES-Homepage

In GENRES sind Informationen und Daten verfügbar über:

- pflanzengenetische Ressourcen (PGR), tiergenetische Ressourcen (TGR), forstgenetische Ressourcen (FGR), genetische Ressourcen bei Mikroorganismen (MGR) sowie zu politisch-rechtlichen Rahmenbedingungen;
- vom IGR zentral projektierte und aufgebaute Datenbanken mit zentralen Auswertungsdaten im Rahmen von GENRES;
- von den im GENRES mitwirkenden Partnern produzierte Datenbanken auf der Grundlage von Genbanken;
- vom IGR erarbeitete bzw. aufbereitete Informationen über Institutionen und Organisationen im nationalen Rahmen;
- vom IGR erarbeitete bzw. aufbereitete Informationen über nationale Maßnahmen zur *Ex-situ*- und *In-situ*-Erhaltung;

- vom IGR und Partnern im Rahmen europäischer oder internationaler Zusammenarbeit projektierte und aufgebaute Datenbanken;
- themenrelevante Volltexte;
- Ergebnisse von eigenen Recherchen nach nationalen und internationalen Informationsangeboten im Internet;
- Datenbanken mit Relevanz für GENRES, die vom DAINet angeboten werden (Literatur, Institutionen, FuE-Projekte, Stellenmarkt etc.).

Informationen zu pflanzengenetischen Ressourcen in GENRES

Für die Rubrik „Pflanzengenetische Ressourcen (PGR)“ ergibt sich die in Tabelle 1 dargestellte Hauptgliederung. Die einzelnen, unterschiedlich tief gegliederten Unterpunkte enthalten zum einen bei Recherchen im Internet ermittelte thematisch relevante Internetangebote, die in Form der sogenannten „Links“ (interaktive Verbindungen zu „Hypertexten“) anwählbar sind. Auf der anderen Seite werden Angebote nachgewiesen, die vom IGR in Zusammenarbeit mit anderen Partnern für das Internet erarbeitet oder aufbereitet wurden. Das betrifft Datenbanken zu PGR im nationalen, europäischen oder internationalen Rahmen, aber auch Informationen über Institutionen und Organisationen bzw. über Maßnahmen zur *Ex-situ*- und *In-situ*-Erhaltung.

Tab. 1: Hauptgliederung der Rubrik „Pflanzengenetische Ressourcen (PGR)“

Tab. 1: Structure of the category „Plant Genetic Resources“

Unterpunkte der Rubrik PGR	Inhalt
GENRES -Logo	Link zur GENRES - Homepage
Fachgebiete	Link zum GENRES - Hauptindex
Pflanzengenetische Ressourcen (PGR)	
PGR in Deutschland	siehe Tabelle 2
PGR in anderen Staaten (geographisch)	siehe Tabelle 3
PGR nach Fruchtarten	siehe Tabelle 4
PGR in Online-Datenbanken	siehe Tabelle 5
Pflanzenzüchtung im DAINet	Link zum DAINet - Unterpunkte: Allgemeines, allg. Pflanzenzüchtung, spez. Pflanzenzüchtung, PGR (GENRES!), Sorten/Saatgut
Stichwortsuche GENRES	Möglichkeit der Freitextsuche in allen GENRES-Datensätzen (Agrardatenkatalog)

Tab. 2: Struktur der Unterrubrik „PGR in Deutschland“

Tab. 2: Structure of the category „PGR in Germany“

Unterpunkte		Beispiele	Anzahl Links
Kultur- und Wildpflanzen		PGRDEU-Online (Datenbank zu pflanzengenetischen Ressourcen in Deutschland) http://www.dainet.de/genres/pgrdeu/	7
Adressen		Institutionen und Organisationen im Agrarbereich (Datenbank der ZADI) http://www.dainet.de:8080/INSTORGFOR M/SF	3
Ex-situ-Erhaltung	Deutsche Genbanken	Genbank IPK Gatersleben http://www.ipk-gatersleben.de/	6
	Botanische Gärten	SysTax – „Botanical Garden Information System“ http://www.biologie.uni-ulm.de/systax/infgard/	39
In-situ-Erhaltung und On-farm-Management		Erhaltung und Rekultivierung von Nutzpflanzen in Brandenburg (VERN) http://www.dainet.de/genres/infos/vern/index.htm	5
Nutzung	Forschung	Bundesanstalt für Züchtungsfor- schung an Kulturpflanzen http://www.bafz.de/	11
	Züchtung	Deutsches Maiskomitee e.V. http://www.Maiskomitee.de/	5
	Sorten/Saatgut	Bundessortenamt http://www.bundessortenamt.de	14
Aus- und Weiterbildung		Zentralstelle für Landwirtschaft und Ernährung (ZEL) http://www.dse.de/zell/zell.htm	1

Tab. 3: Struktur der Unterrubrik „PGR in anderen Staaten (geographisch)“

Tab. 3: Structure of the category „PGR in other countries (geographical)“

Unterpunkte		Beispiele	Anzahl Links
Institutionen/Sammlungen	International	World Seed Info http://www.worldseed.org/	3
	(Über-)Regional	Europ. Programme ECP/GR http://www.cgiar.org/ecpgr/index.htm	13
		FAO Plant Genetic Resources Information http://web.icppgr.fao.org/	2
		CGIAR IPGRI http://www.cgiar.org/ipgri/index.htm	14
	Europa	z.Z. 14 Länder INRA Maize Genome Database http://moulon.moulon.inra.fr/imgd/	28
	Nordamerika	CA, MX, US GRIN-WWW-Service http://www.ars-grin.gov/	53
	Südamerika	PE, BR, CO International Potato Center (CIP) http://www.cgiar.org/cip/	3
	Afrika	KE, NG, ZM, CI SADC Plant Genetic Resources Centre (SPGRC) http://tor.ngb.se/sadc/spgrc.html	5
	Asien	JP, SY, IN Rice Genome Research Program (RGP) http://www.staff.or.jp/	3
	Australien/Ozeanien	NZ, AU, PH Australian Network for Plant Conservation http://www.anbg.gov.au/anpc/anpc97.html	5
Botanische Gärten	Europa	DK, FI, PL, UK, AT, IT The Royal Botanic Gardens http://www.rbgkew.org.uk/	11
	Nordamerika	CA, US Missouri Botanical Garden http://www.mobot.org/	41
	Australien/Ozeanien	AU, NZ National Australian Botanical Gardens http://155.187.10.12/anbg/anbg.html	4

Tab. 4: Struktur der Unterrubrik „PGR nach Fruchtarten“

Tab. 4: Structure of the category „PGR by species“

Unterpunkte		Beispiele	Anzahl Links
Getreide und Kleinkörnerfrüchte			
	Weizen	Triticum - Germplasm Resource Information Program for Wheat (GRIP I) http://www.psu.missouri.edu/grip/	11
	Reis	WARDA - West Africa Rice Development Association http://www.cgiar.org/warda	7
	Mais	INRA Maize Genome Database http://moulon.moulon.inra.fr/imgd/	7
	Gerste	GrainGenes - A Database for Small Grains and Sugarcane http://wheat.pw.usda.gov/graingenes.html	4
	Hafer	GrainGenes, the Triticeae Genome Gopher gopher://greengenes.cit.cornell.edu/html	4
	Hirse	MilletGenes - Pearl Millet http://probe.nalusda.gov:8000/plant/aboutmilletgenes.html	2
Leguminosen		Data Base Soybean http://mendel.agron.iastate.edu:8000/main.html	10
Zucker- u. Stärkepflanzen		The Dutch-German Potato Collection http://www.cpro.dlo.nl/cgn/potato/	9
Ölpflanzen		Eastern Cereal and Oilseed Research Centre http://res.agr.ca/ecorc/default.htm	3
Faserpflanzen		CottonDB Data Collection http://algodon.tamu.edu/cottondb.html	2
Futterpflanzen		Trifolium - The ECP/GR Trifolium Database http://www.ngb.se/Material/Trifolium.html	7
Gemüse		EU Lettuce Project http://www.cpro.dlo.nl/cgn/lettuce/	7
Obst		Lescrets Pomologie http://perso.club-internet.fr/lescrets/	7
Wein		Vitis - International Vitis Database http://www.dainet.de/genres/idb/vitis	2
Zierpflanzen		The Orchid House http://sciserv2.uwaterloo.ca/orchids.html	3
Andere Pflanzen			
	Arabidopsis	Arabidopsis thaliana Database http://genome-www.stanford.edu/Arabidopsis/	5
	Chlamydomonas	Chlamydomonas Genetics Center - DCMB Group http://www.botany.duke.edu/DCMB/chlamy.htm	2

Tab. 5: Struktur der Unterrubrik „PGR Online-Datenbanken“

Tab. 5: Structure of the category „PGR Online-databases“

Unterpunkte	Beispiele	Anzahl Links
Fruchtartenübergreifend	Informationssystem für Evaluierungsdaten (EVA) http://www.dainet.de/genres/eva/	9
Fruchtartenspezifisch	<i>Hordeum</i> - European Barley Database http://www.dainet.de/eccdb/barley/	20

Tab. 6: Struktur eines Datensatzes in der Datenbank „Agrardatenkatalog“ für die Verweissysteme DAINet und GENRES

Tab. 6: Structure of a record in the database „Agrardatenkatalog“ used for DAINet and GENRES

Feldbezeichnung	Feldinhalt (Beispieldatensatz)
Internet-Adresse	http://www.dainet.de/genres/pgrdeu/
Bezeichnung deutsch	PGRDEU-Online (Datenbank zu pflanzengenetischen Ressourcen in Deutschland)
Bezeichnung englisch	PGRDEU-Online (Plant Genetic Resources in Germany. Online Database)
Beschreibung deutsch	Genbanksammlungen zur deutschen Flora und zu Wildpflanzen mit aktueller oder potentieller Nutzung, Suchmöglichkeiten nach Pflanzennamen, Genbanken, Ursprungsland, online
Beschreibung englisch	Information about the <i>ex situ</i> collections of the most important German genebanks as well as information about the German flora and German wild plants with potential or actual utilization
Anbieter deutsch	Informationszentrum für Genetische Ressourcen (IGR);Zentralstelle für Agrardokumentation und -information (ZADI)
Anbieter englisch	Information Centre for Genetic Resources (IGR);German Center for Documentation and Information in Agriculture (ZADI)
Land ¹	DE
Thesaurus	GRP-Kultur- und Wildpflanzen; GRP-multi-crop; Genetische Ressourcen; Fakten; Volldatenbank; kostenfreier Onlinezugang; HTTP- Datenbank; Deutschland
Eingabe-Datum	Sep 10, 1997
Computer-Eingabedatum	Sep 10, 1997
Letzte Änderung	Jan 12, 1998
Gewicht ²	10,000

1) nach ISO 3166 Codes (Toplevel-Domains)

2) erlaubt Festlegung einer Rangfolge der Verweise bei der Recherche

Alle Verweise auf andere Internetadressen der beiden Verweissysteme DAINet und GENRES sind in einer zentralen Datenbank (Agrardatenkatalog) in der ZADI gespeichert und nach inhaltlichen und formalen Gesichtspunkten indexiert. In Tabelle 6 ist der prinzipielle Aufbau eines Datensatzes dargestellt. Die Speicherung der Datensätze für beide Informationssysteme in einer Datenbank erlauben ein redundanzfreies Arbeiten und zugleich die einfache wechselseitige Nutzung von für beide Systeme relevanten Datensätzen.

Die Präsentation der zu den einzelnen GENRES-Gliederungspunkten bereitgestellten relevanten Verweise erfolgt durch „pre-processed searches“ im Agrardatenkatalog. Diese bei Neuzugängen zur Datenbank ständig aktualisierte dynamische Version wird einmal pro Woche in eine statische Fassung überführt, die den Nutzern von GENRES im Internet zur Verfügung gestellt wird. Dadurch werden die Recherchezeiten und die Belastung des Datenbank-servers reduziert.

Ausgewählte internationale Informationssysteme zu pflanzengenetischen Ressourcen im Internet

Selected international internet-based information systems on plant genetic resources

SIEGFRIED HARRER, JULIAN DANIEL JIMÉNEZ KRAUSE UND FRANK BEGEMANN¹

Zusammenfassung

Informationssysteme spielen eine entscheidende Rolle bei der Erhaltung und nachhaltigen Nutzung von pflanzengenetischen Ressourcen (PGR). Aufgrund der aktuellen Entwicklung in der Informationstechnologie etablieren sich zunehmend mehr solcher Systeme als Online-Informationssysteme im Internet. Hierfür werden die grundlegenden Komponenten / Strukturen vorgestellt und anhand ausgewählter Informationssysteme exemplarisch erläutert.

Abstract

Information systems play a fundamental role in the conservation and sustainable use of plant genetic resources (PGR). As a result of the latest development in information technology, an increased number of such systems are created as online information systems in the Internet. The main components of such systems are presented and explained through some selected examples.

Einleitung

Die Erhaltung und nachhaltige Nutzung der Biodiversität wurde von der internationalen Staatengemeinschaft im Übereinkommen über die biologische Vielfalt (UNEP 1992) festgeschrieben. Die sich aus der Nutzung der Vielfalt ergebenden Vorteile sollen dabei fair und gerecht aufgeteilt werden. Für den Teil der biologischen Vielfalt, welcher relevant für die menschliche Ernährung ist, gibt der Globale Aktionsplan (GPA) der FAO (1996) Prioritäten zur Umsetzung dieser Ziele. Bei der Umsetzung dieser Maßnahmen spielen Dokumentations- und Informationssysteme eine entscheidende Rolle. Sowohl für die Erhaltung als auch für die nachhaltige Nutzung von pflanzengenetischen Ressourcen (PGR) sind umfassende Informationsmöglichkeiten zu PGR eine wesentliche Voraussetzung. Die Bereitstellung von solcher Information sollte dabei stets strukturiert erfolgen. Dabei sind ähnlich wie bei der biologischen Vielfalt selbst horizontale und vertikale Komponenten zu unterscheiden. Information zu PGR muß daher z.B. von der interspezifischen Stufe (z.B. Information auf Familien- / Gat-

¹ Zentralstelle für Agrardokumentation und -information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

tungsebene) über die intraspezifische Stufe (z.B. Informationen auf Akzessions- / Sortenebene) zunehmend feiner strukturiert angeboten werden. Hierfür stehen für Online-Informationssysteme (Internet) verschiedene „Werkzeuge“ zur Verfügung, welche im folgenden näher erläutert werden.

Komponenten von PGR-Informationssystemen im Internet

Informations- und Kommunikationsplattformen

Bei Informations- und Kommunikationsplattformen handelt es sich im einfachsten Fall um normale HTML-Dokumente, welche PGR-relevante Information vermitteln. Die Realisierung eines solchen Angebotes ist demgemäß sehr einfach und stellt keine großen Anforderungen an Hard- und Software. Nachfolgend sind beispielhaft einige Angebote dieser Informationskomponenten aufgeführt:

- Verweise auf relevante Systeme
- Veröffentlichungen, Literatur
- Aktuelle Forschungsprojekte, Projektkooperation
- Newsletter, Nachrichten
- Diskussionsforen
- Veranstaltungskalender
- Jobbörsen, Praktika und Stipendien

Metadaten-Informationssysteme

Bei Metadaten-Informationssystemen handelt es sich um komplexere Informationssysteme, welche generelle Information über Informationsquellen und deren Struktur vermitteln. In ihrer allgemeinsten Form sind dies z.B. reine Verweissysteme zu detaillierterer Information („Gelbe-Seiten-Prinzip“). Aufgrund der steigenden Informationsflut im Internet gewinnen diese Informationssysteme nicht nur im PGR-Bereich zunehmend an Bedeutung. Die Beschreibung der Informationsquellen im Internet erfolgt anhand semantischer, struktureller und medialer Merkmale. Diese Vorgehensweise wird von FRIEDRICH & POHLMANN (1997) ausführlich für den Agrardatenkatalog des Deutschen Agrarinformationsnetzes (DAINet) dargestellt. Eine Realisierung ist zwar auch hier mit reinen HTML-Dokumenten möglich. Für umfangreiche Verweissysteme bietet sich aber für die Verwaltung der relevanten Informationen (z.B. Internetadresse (URL), Beschreibung und Struktur der Information) die Verwendung von Online-Metadatenbanken an. Dies stellt dann allerdings schon höhere Anforderungen an die Etablierung und den Unterhalt eines solchen Systems, ist aber aufgrund der hohen Dynamik, besonders was das ständig sich ändernde und zunehmende Informationsangebot im Internet betrifft, die bessere Lösung. Dabei erlaubt eine Online-Metadatenbank à la Agrardatenkatalog z.B. dynamisches Angebot, Indexierung, automatische Überprüfung der URLs, interaktive Aktualisierung (URL Eingabe/Änderung) etc.. Ein Beispiel ist das Informationssystem für Genetische Ressourcen (GENRES, siehe Beitrag MÜNCH).

Online-Faktendatenbanken

Während Metadatenysteme quasi „horizontal“ einen Überblick über vorhandene Information ermöglichen, können strukturierte Faktendaten als die „vertikale“ Komponente von Dokumentationssystemen gesehen werden. Im Falle von PGR geben Faktendaten z.B. Auskunft über Muster/Akzessionen in PGR-Sammlungen. Sie liefern detaillierte Informationen zum Material selbst und zu den Zugangswegen zu ihm. Wegen Umfang und Art der Faktendaten zu PGR (Daten zur Identifizierung, Herkunft, Eigenschaften, etc.) werden diese idealerweise in strukturierten Dateien, vorzugsweise in Datenbanken dokumentiert. Der optimale Zugang zu PGR-Faktendaten wird also über Online-Faktendatenbanken gewährleistet, d.h. über strukturierte Dateien, die interaktiv und nutzerfreundlich übers WWW abfragbar sind. Für die Realisierung von Online-Datenbanken werden hohe Anforderungen an Soft- und Hardware gestellt. So ist z.B. ein Datenbankserver mit WWW-Schnittstelle und einer permanenten Leitung zum Internet notwendig. Für eine effiziente Online-Darstellung von Online-Datenbanken ist auch die Qualität der benutzten Internetverbindung von großer Bedeutung. Nicht zu unterschätzen ist der zusätzliche Aufwand an personellen Ressourcen (Datenbankspezialisten, Netzwerkmanager), der bei der Realisierung eines solchen Unternehmen anfällt.

Eine technisch leichter realisierbare Form, strukturierte Dateien online anbieten zu können, ist es, Datenbankdateien (z.B. in einem der üblichen Datenbankformate dBASE, FoxPro, Excel, Access etc.) zum Herunterladen (*download*) zur Verfügung zu stellen, so daß sie auf der lokalen Festplatte gespeichert und vom Interessenten lokal (*offline*) mit der entsprechenden Datenbanksoftware benutzt (gelesen, abgefragt) wird.

Europäische Ebene - European Cooperative Programme for Crop Genetic Resources Networks (ECP/GR)

Das ECP/GR ist ein europäisches Kooperationsprogramm u.a. mit den Zielen, die Erhaltung von PGR sowie deren nachhaltige Nutzung in Europa zu unterstützen. Hierzu arbeiten innerhalb dieses Programms 10 fruchtarten- oder themenspezifische (bezogen auf PGR) Netzwerke.

European Information Platform on Crop Genetic Resources

Die European Information Platform on Crop Genetic Resources (<http://www.cgiar.org/ecpgr/platform/index.htm>) ist die Informationsplattform des ECP/GR. Ausgehend von einem ECP/GR-Dokumentationsworkshop 1996 in Budapest (LIPMAN ET AL. 1997) wurde in der Folge dieses Meta-Informationswerkzeug geschaffen, welches u.a. verweist auf (Abb. 1):

- Regionale / Globale Fruchtartenspezifische PGR-Datenbanken
- Relevante PGR-Institutionen / Einrichtungen
- Serviceangebote (Newsletter, Informationen, Verweise etc.).

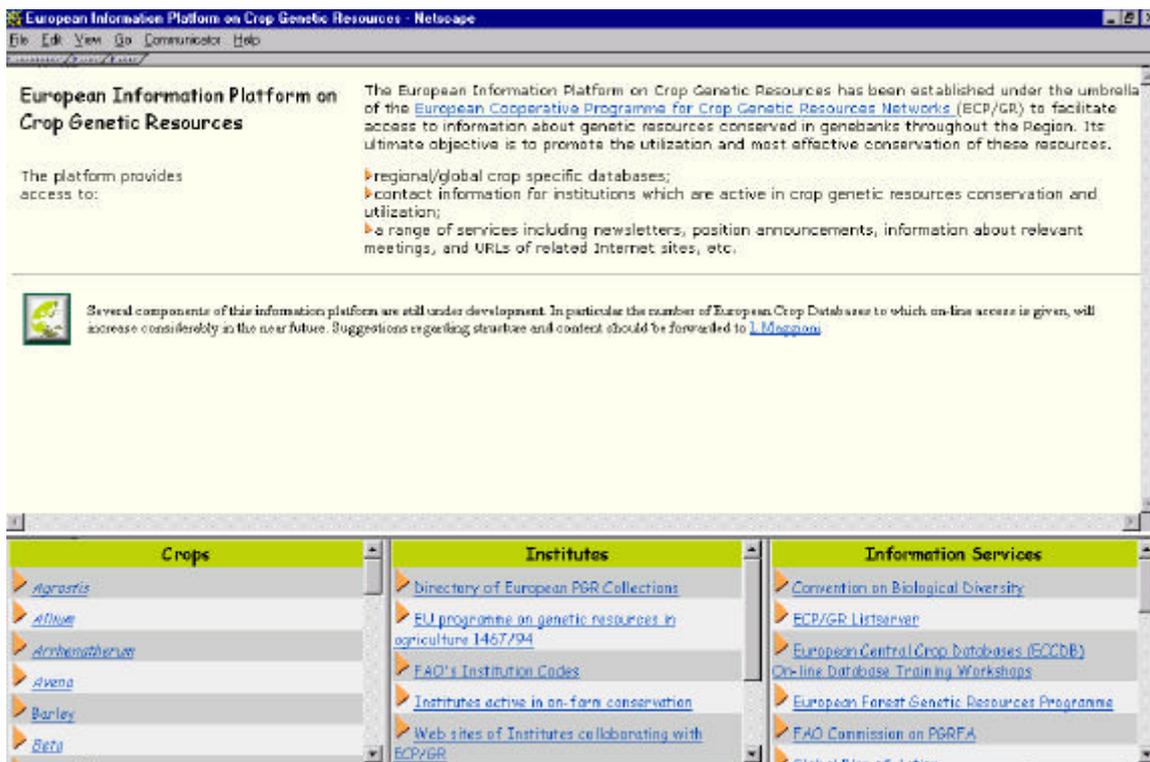


Abb. 1: Informationsplattform des ECP/GR

Fig. 1: Information platform of ECP/GR

European Central Crop Databases (ECCDB)

Eine der Hauptaufgaben der *Crop Networks* oder fruchtartenspezifischen Netzwerke des ECP/GR ist die Kompilierung von fruchtartenspezifischen Faktendatenbanken (European Central Crop Databases (ECCDB)) mit dem Ziel, das Informationsangebot über europäische PGR zu erhöhen. (VAN HINTUM 1997). Ein ECCDB-Manager ist für Datenbankdesign, Datensammlung und –kompilierung sowie Zusammenstellung und Veröffentlichung der ECCDB verantwortlich. Die Daten zu der spezifischen Fruchtart werden ihm zugelifert von jenen europäischen Genbanken, die Muster zu dieser Fruchtart erhalten. Dies bedeutet, daß der ECCDB-Manager Daten aus verschiedenen Quellen zu homogenisieren und zusammenzufügen hat, welches unter Umständen zu einem erheblichen Konvertierungsaufwand führt. Um diesem zu vermindern, wurde in dem Dokumentationsworkshop in Budapest 1996 empfohlen, ein standardisiertes Datenaustauschformat für den Datentransfer zu benutzen, die Multicrop Passport Descriptor List der FAO/IPGRI (LIPMAN ET AL 1997).

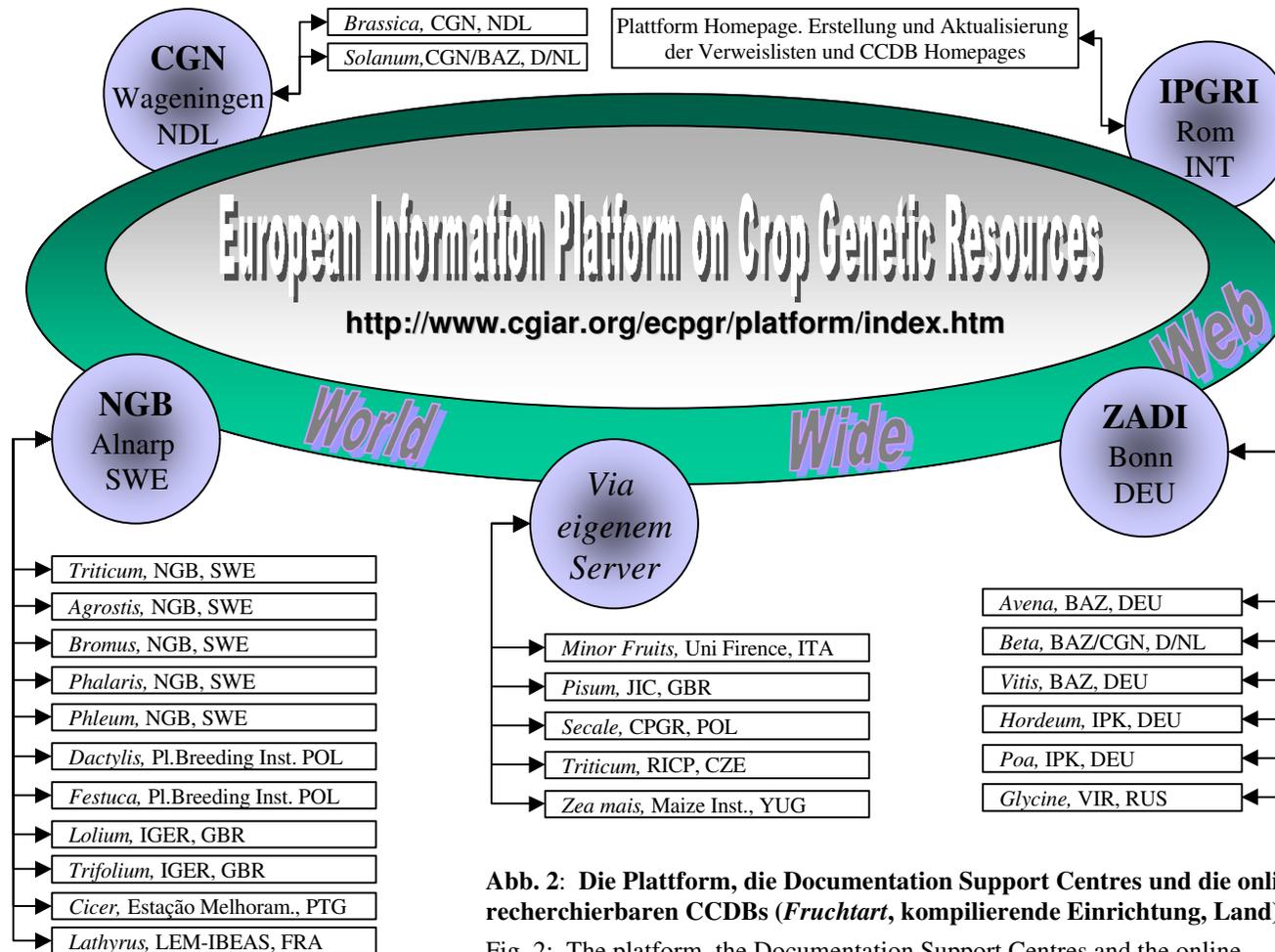


Abb. 2: Die Plattform, die Documentation Support Centres und die online recherchierbaren CCDBs (Fruchtart, kompilierende Einrichtung, Land).

Fig. 2: The platform, the Documentation Support Centres and the online searcheable CCDBs (*crop*, compiling institute, country)

ECCDB online, Documentation Support Centres (DSCs) und Internet Advisory Group

Das WWW bietet einen optimalen Zugang zu den ECCDB. In dem Dokumentationsworkshop in Budapest 1996 wurde als eine Empfehlung ausgesprochen, das Angebot der ECCDB über WWW zu fördern. Optimalerweise sollte jede ECCDB als Datenbankdatei zum herunterladen und/oder als Online-Datenbank bereitgestellt werden. Jedoch, nur wenige PGR Einrichtungen verfügen über das technische und personelle Potential zur praktischen Umsetzung und Implementierung von Online-Datenbanken. Daher wurde in dem Workshop weiter empfohlen, solche PGR-Einrichtungen mit den notwendigen technischen und personellen Voraussetzungen, sowie der Bereitschaft, andere Einrichtungen bei der Online-Verfügbarmachung ihrer Datenbestände zu unterstützen, zu lokalisieren. Sie bildeten sogenannte Documentation Support Centres (DSC) (Dokumentationsunterstützende Einrichtungen, siehe Abb. 2). Darüber hinaus wurde ein informationstechnologisches Beratergremium (Internet Advisory Group), bestehend aus Kollegen aus DSC mit dem entsprechenden Know-How, gegründet. DSC und Internet Advisory Group unterstützen seither die ECCDB-Datenbankmanager bei Fragen zu Internet und Online-Datenbanken und beraten bei der Online-Verfügbarmachung einzelner ECCDB über die DSC mitwirken bzw. fungieren als Online-Host für die ECCDBs (s. Abb.2). Mittlerweile sind zudem zwei Dokumentationstagungen vom ECP/GR und der Internet Advisory Group organisiert worden, um die ECCDB-Manager in Sachen Internet und Online-Datenbanken zu trainieren.

Das Konzept und die Vorgehensweise des ECP/GR zur Umsetzung und Realisierung eines effektiven und umfassenden Dokumentations- und Informationssystems pflanzengenetischer Ressourcen sind zweifelsohne ein hervorragendes Beispiel für Pragmatismus und optimale Nutzung vorhandener technischer und personeller Kapazitäten, das sich als Grundmodell auf analoge regionale und internationale Informationssysteme übertragen läßt.

Internationale Ebene – FAO und CGIAR

Auf internationaler Ebene spielen neben anderen regionalen, nationalen (z.B. dem GRIN der USA) und institutionellen Informationssystemen insbesondere die Systeme der FAO und der CGIAR eine führende Rolle. Als Informationsplattform mit horizontalem Charakter bietet sich das von der FAO entwickelte World Information and Early Warning System (WIEWS, <http://apps2.fao.org:8080/wiews.new/>) an. WIEWS kann dabei in der Kooperation mit dem World Agricultural Information Centre (WAICENT, www.fao.org/waicent/search/default.htm) der FAO von einem entsprechenden Know-how-Vorteil profitieren. Die horizontale Funktion wäre hier in Anlehnung an die European Platform oder das deutsche GENRES (<http://www.genres.de>) ein Verweissystem auf bestehende regionale und nationale Informationsplattformen. Hinzu kommen normative Informationsdienstleistungen von WIEWS, wie z.B. das Angebot der Minimum Passport Descriptor List zur Vereinfachung des Datenaustauschs.

Eine besondere Rolle kommt der FAO im Bereich der Überwachung politisch-rechtlicher Vereinbarungen zu. Von besonderer Bedeutung sind in diesem Zusammenhang die zwischen der FAO und den internationalen Agrarforschungszentren der CGIAR und dem Coconut Network (COGENT) geschlossenen Vereinbarungen, welche die Sammlungen dieser Partner unter die Treuhandenschaft der FAO stellen und auch den Zugang zu diesen Sammlungen in besonderer Weise regeln. Das WIEWS der FAO würde sich als Informationswerkzeug zur Überwachung dieser Vereinbarungen anbieten und in diesem Fall auch horizontale, d.h. akzessionsbezogene Informationen beinhalten müssen. Sollten derartige Informationen von den jeweiligen Vertragspartnern, hier insbesondere von den Zentren der CGIAR selbst dokumentiert werden, bestünde die Rolle von WIEWS in der partnerschaftlichen Referenz zu diesen Informationen.

Das vom systemweiten Programm der CGIAR zu genetischen Ressourcen entwickelte Informationssystem SINGER (<http://noc1.cgiar.org/>) dokumentiert eben diese Angaben (u.a. auch die einzelnen Material Transfer Agreements (MTA)). Daneben bietet es als eine institutionelle Datenbank Zugang zu verschiedenen Fruchtarten. SINGER ist daher eine Multi Crop Datenbank, die sich eher mit anderen institutionellen oder nationalen Datenbanken vergleichen läßt. Sie beinhaltet akzessionsbezogene Daten. Der Schwerpunkt liegt auf den Passportdaten, während die Charakterisierungs- und Evaluierungsdaten aus eigenen Forschungsarbeiten der Zentren in der zweiten Projektphase nun vervollständigt werden sollen.

Schlußfolgerungen

Es sollte ein international abgestimmtes System von nationalen, regionalen und internationalen Informations- und Kommunikationsplattformen im oben geschilderten horizontalen Sinne etabliert werden. Daneben sollten Schritte zum Aufbau von vertikalen Datenbanken mit akzessionsbezogenen Daten, nämlich von abgestimmten Central und Multi Crop Datenbanken unternommen werden mit dem Ziel, im Idealfall internationale Central Crop Datenbanken (ICCDB) zu etablieren. Beispiele sind hier u.a. die International Vitis Database (<http://www.dainet.de/genres/idb/vitis/>) oder die International Beta Database (<http://www.dainet.de/genres/beta/>), die von BAZ-Instituten gemanaget werden. Wo dies nicht möglich ist, sollten regionale, d.h. in Europa, europäische ECCDB und, wo auch dies nicht möglich ist, nationale bzw. institutionelle Multi Crop Datenbanken eingerichtet werden. Die Datenbanken bzw. Informationssysteme sollten online recherchierbar sein. Beim Zugang zu den Informationen werden künftig noch zu entwickelnde Wege beschritten werden müssen, die insbesondere den Schutz der politisch-rechtlichen und institutionellen Rahmenbedingungen sicherstellen.

Literatur

- FRIEDRICH, H. UND J.M. POHLMANN (1997): Metadatenbanken als Steuer- und Leitsysteme für das Deutsche Agrarinformationsnetz. *Berichte der GIL*, 10, S. 31-36.
- BECKER, B. (1999): Bundesinformationssystem Genetische Ressourcen – Konzept und erste Ergebnisse. In: F. BEGEMANN, S. HARRER, J. D. JIMÉNEZ KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. *Schriften zu Genetischen Ressourcen*, Band 12, ZADI, Bonn.
- FAO (1996c): Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture. FAO, Rome, 63 pp.
- HARRER, S. (1999): EVA - Aufbau eines Informationssystems für Evaluierungsdaten pflanzengenetischer Ressourcen in der Bundesrepublik Deutschland. In: F. BEGEMANN, S. HARRER, J. D. JIMÉNEZ KRAUSE (eds.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. *Schriften zu Genetischen Ressourcen*, Band 12, ZADI, Bonn.
- VAN HINTUM, TH. (1997): Central Crop Databases – an overview. In: LIPMAN, E., M.W.M. JONGEN, TH.J.L. VAN HINTUM, T. GASS AND L. MAGGIONI (eds.) (1997): Central Crop Databases: Tools for Plant Genetic Resources Management 17-19, International Plant Genetic Resources Institute, Rome, Italy/CGN Wageningen, The Netherlands
- LIPMAN, E., M.W.M. JONGEN, TH.J.L. VAN HINTUM, T. GASS AND L. MAGGIONI (eds.) (1997): Central Crop Databases: Tools for Plant Genetic Resources Management. 75-78. International Plant Genetic Resources Institute, Rome, Italy/CGN, Wageningen, The Netherlands
- MÜNCH, E. (1999): Pflanzengenetische Ressourcen im Informationssystem Genetische Ressourcen (GENRES) – Online Demonstration. In: F. BEGEMANN, S. HARRER UND J. D. JIMÉNEZ KRAUSE (Hrsg.): Dokumentation und Informationssysteme im Bereich pflanzengenetischer Ressourcen in Deutschland. *Schriften zu Genetischen Ressourcen*, Bd. 12, ZADI, Bonn.
- UNITED NATIONS ENVIRONMENT PROGRAMM (1992): Convention on Biological Diversity, UNEP, Nairobi.

Das Internet-Informationsangebot des N. I. Vavilov Institute of Plant Industry (VIR)

Internet Information System of the N. I. Vavilov Institute of Plant Industry (VIR)

SIEGFRIED HARRER¹ UND ANDREW OMELCHENKO²

Abstract

Since its establishment in 1894, the N. I. Vavilov Institute of Plant Industry (VIR), St Petersburg, Russia has accumulated impressive quantities of data associated with germplasm collections held there. Until recently, these data were recorded on cards and paper reports. As a result of a joint project between VIR and the German Centre for Documentation and Information in Agriculture (ZADI) a VIR information system on the Internet was established. The most important part of this Information System is the online database. Currently, the Institute holds more than 330.000 accessions, representing 2.539 species within 304 genera of 155 botanical families. To date, 250.000 accessions have been verified and the passport data of approximately 85% of this collection is now available on the Internet (<http://www.genres.de/vir>).

Einführung

Seit seiner Gründung im Jahre 1894 konnte das N.I. Vavilov Institute of Plant Industry (VIR), St. Petersburg, eine der weltweit bedeutendsten und umfangreichsten Sammlungen zu pflanzen genetischen Ressourcen aufbauen. Grundlagen waren u.a. die Sammelexpeditionen des weltweit anerkannten Wissenschaftlers N.I.Vavilov, welcher das Institut von 1921 – 1940 leitete. In dieser Zeit nahmen die Wissenschaftler des VIR an 180 Sammelexpeditionen in 65 Ländern teil und bis 1940 hatte die Sammlung schon einen Umfang von 200.000 Mustern erreicht.

Heute umfaßt die Sammlung des VIR ca. 330.000 Muster von 2539 Pflanzenarten aus 155 Familien. Die darin enthaltene Diversität wurde und wird durch das VIR mit seinen Außenstationen über einen weiten ökogeographischen Bereich hinweg wissenschaftlich und auch im Hinblick auf züchterischer Nutzung untersucht. Die umfangreichen Ergebnisse dieser Arbeiten stellen eine wesentliche Grundlage für die Nutzung der Sammlung dar. Der größte Teil dieser wertvollen Informationen stand bisher nicht EDV-gerecht zur Verfügung, wodurch der Zugang und die Nutzung, besonders für Außenstehende, erheblich eingeschränkt war.

¹ Zentralstelle für Agrardokumentation und –information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

² N. I. Vavilov Institute of Plant Industry (VIR)
St. Petersburg, Rußland

In dem vom Bundesministerium für Ernährung, Landwirtschaft und Forsten (BML) geförderten Projekt „Beratungshilfe für das VIR im Bereich der Dokumentation und Information pflanzengenetischer Ressourcen“ zwischen dem VIR und der Zentralstelle für Agrardokumentation und -information (ZADI) wurde in den letzten Jahren vornehmlich durch technische Beratung der Aufbau eines VIR-Informationssystems unterstützt. Bereits Ende der ersten Projektphase (1994-1996) waren die wichtigsten Passportdaten von 80% der Sammlungen online im Internet recherchierbar. Im Zuge der zweiten Projektphase (1997-1999) wird das Internetangebot inhaltlich wesentlich ausgebaut. Das jetzige VIR-Informationssystem bietet allgemeine Informationen zu Struktur und Aufgaben des VIR sowie verschiedene Online-Datenbanken, die ständig ausgebaut werden.

Online Angebot des VIR

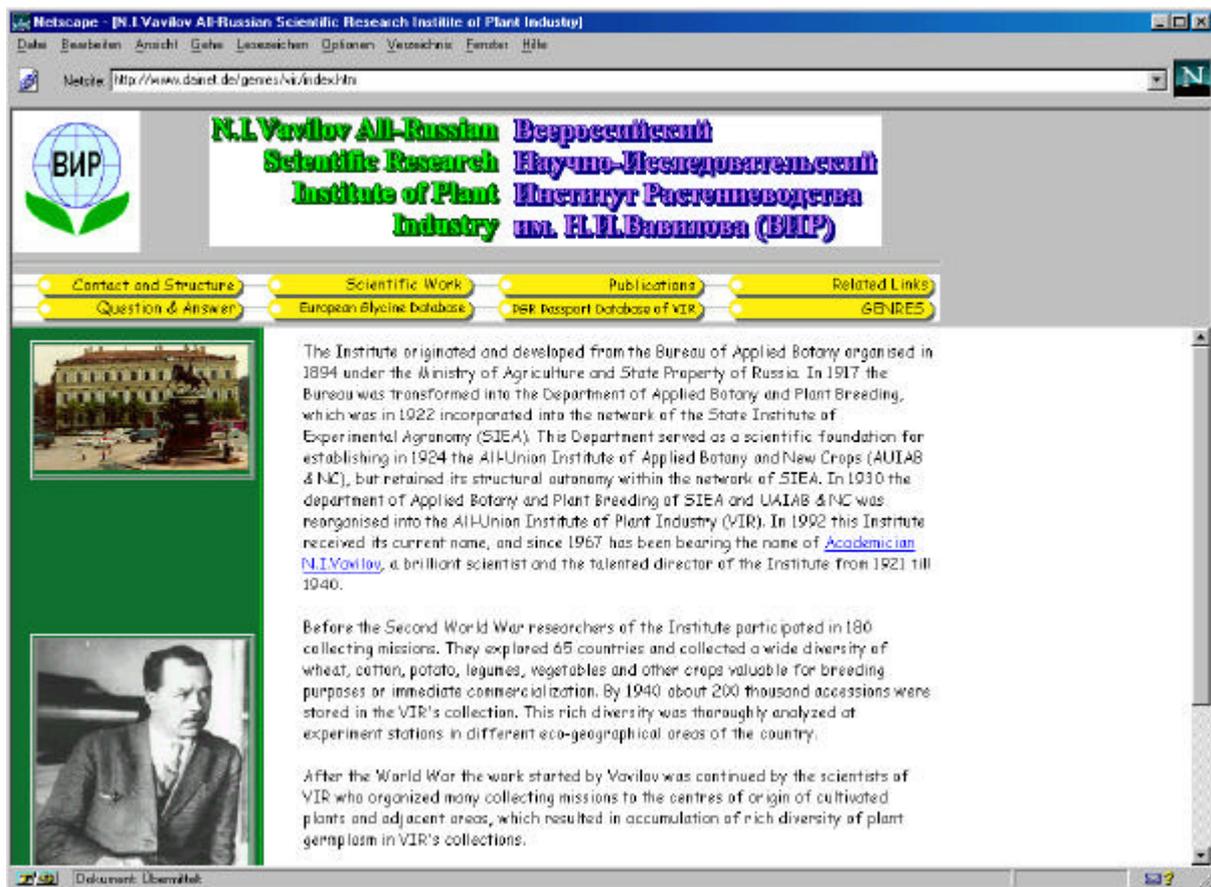


Abb.1: Homepage des VIR-Informationssystems
Fig. 1: Homepage of the information system of VIR

Allgemeine Informationen. Die Homepage (<http://www.dainet.de/genres/vir/>) des VIR-Informationssystems bietet einen zentralen Einstieg zum Informationsangebot des VIR. Wie aus Abbildung 1 ersichtlich wird, besteht die Möglichkeit, sich über die geschichtliche Entwicklung und jetzige Struktur des VIR zu informieren. Für die einzelnen Abteilungen und Außenstationen stehen dabei u.a. jeweils in einer Kurzbeschreibung Angaben zu Art und Umfang der Aufgaben und Sammlungen sowie Namen der verantwortlichen Wissenschaftler. Weitere Angaben wie Kontaktadressen der leitenden Personen runden das allgemeine Informationsangebot ab.

Informationen zu wissenschaftlichen Arbeiten. Unter dem Menüpunkt „*Scientific Work*“ ist geplant, die wissenschaftliche Arbeit der einzelnen Abteilungen detailliert vorzustellen und damit den internationalen wissenschaftlichen Austausch zu unterstützen. Bisher stehen diese Angaben allerdings erst für wenige Abteilungen in englischer Sprache zur Verfügung. Das beste Angebot hierzu bietet augenblicklich die Abteilung Roggen und Hafer; u.a. wird ein botanischer Bestimmungsschlüssel sowie Metainformation zu Evaluierungsarbeiten beim Hafer angeboten. Der Menüpunkt „*Publications*“ informiert über aktuelle Veröffentlichungen aus dem VIR und bietet die Möglichkeit, Kopien zu ordern. Am weiteren Ausbau dieses Informationsangebotes wird mit Nachdruck gearbeitet.

Documents	Crop	Accession Number	Accession Name	Status of Sample	Form of Life
381	Barley	3524	ALBERT		SP
382	Barley	3529	DE PRINTEMPS		SP
383	Barley	3530	D'HIVER		WI
384	Barley	3531	ESCOURGEON D'HIVER		WI
385	Barley	3532	H-378		SP
386	Barley	3537	ENGLISH ARCHER		SP
387	Barley	3542	K-21		WI
388	Barley	3543	H-341		SP
389	Barley	3544	H-410		SP
390	Barley	3552	H-57		SP
391	Barley	3554	H-16		SP
392	Barley	3555	H-19		SP

Abb. 2: Suche nach Gerstenmustern (Barley) über die Artenliste ergibt Passportdaten zu 17.021 Sammelmustern

Fig. 2: Search for barley via the crop list results in passport data for 17.021 accessions

Online-Datenbanken. Die online-Datenbanken stellen den wichtigsten Teil des Informationssystems dar. Von den ca. 330.000 Mustern sind derzeit ca. 250.000 überprüft und in den offiziellen Katalog übernommen worden. Zu 85% hiervon bietet die **Passportdatenbank** einen komfortablen Informationszugang über das Internet. Die Struktur dieser Datenbank ist weitestgehend kompatibel zur *Multicrop Passport Descriptor List* von IPGRI (LIPMAN ET AL 1997) und folgt somit internationalen Standards. Der Zugang zur Passportdatenbank ist benutzerfreundlich gestaltet:

- Über eine Fruchtartenliste (s. Abb. 2) ist ein schneller Zugang zu den Passportdaten der einzelnen Fruchtarten in der VIR-Kollektion möglich. Dies gestattet einen schnellen Informationszugang und einen ersten Überblick über Umfang und Zusammensetzung der Kollektion.
- Mittels eines Suchformulars können kombinierte Suchen über alle 13 Felder der Passportdatenbank durchgeführt und somit auch komplexere Fragestellungen beantwortet werden (s. Abb. 3).

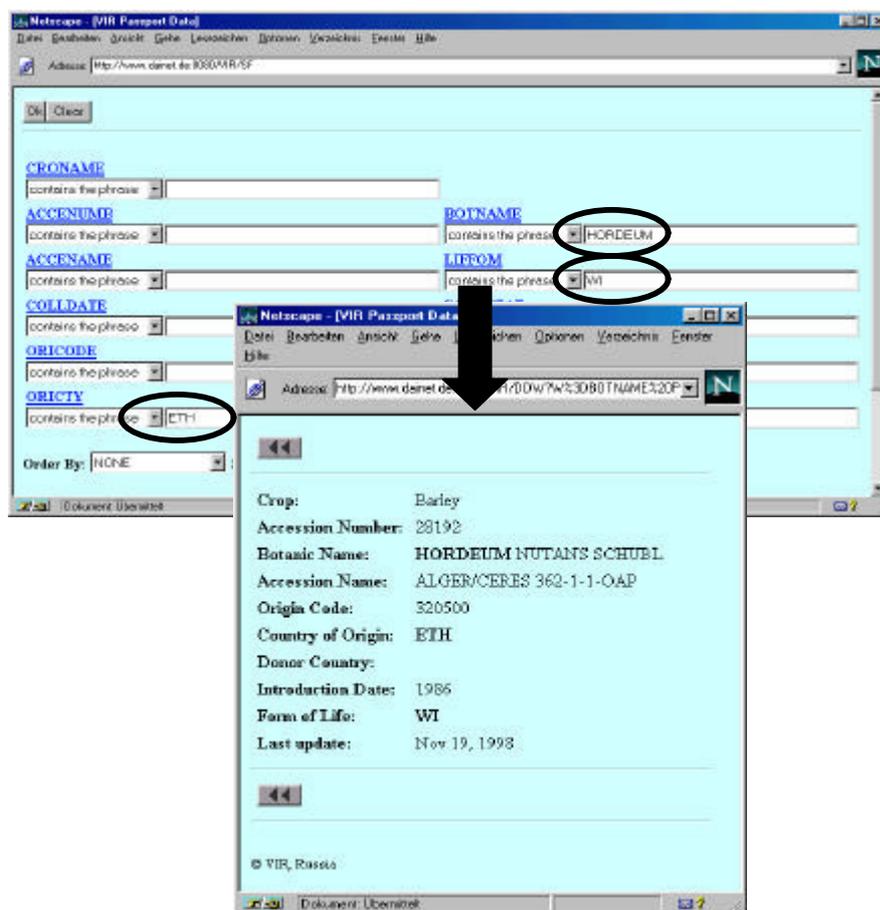


Abb. 3: Kombinierte Suche nach *Hordeum* & Herkunftsland Äthiopien (ETH) & Lebensform Winter (WI)

Fig. 3: Combined search for „*Hordeum*“ & country of origin „Ethiopia“ (ETH) & life form „winter“ (WI)

Als weiteres Datenbankangebot steht seit kurzem die **Datenbank des VIR-Herbariums** online zur Verfügung. Dieses weltberühmte Herbarium beherbergt über 250.000 Herbarbelege von Kulturpflanzen sowie deren verwandter Wildarten. Online abrufbar sind derzeit die Arten (ca. 1.000), zu welchen beim VIR Herbarbelege zum wissenschaftlichen Austausch bereitliegen.

Serviceangebot. Das Informationsangebot wird durch einen Question&Answer-Service und eine noch weiter auszubauende Linksammlung abgerundet. Ersteres ermöglicht auf unkomplizierte und schnelle Art die Inanspruchnahme des Fachwissens der Wissenschaftler des VIR. Die Linksammlung ist im Aufbau begriffen und hat ihren Schwerpunkt im russischen / osteuropäischen Raum und soll auch nicht-englischsprachige Internetangebote aufführen.

Ausblick. Aufgrund mangelnder technischer Voraussetzungen beim VIR ist das Internetangebot derzeit auf dem Webserver der ZADI untergebracht. Allerdings ist im Zuge des weiteren Ausbaus des VIR-Intranets sowie der Weiterentwicklung des dortigen Informations- und Dokumentationssystems ein eigener VIR-Webserver vorgesehen. Zur Unterstützung dieser Aktivitäten ist eine dritte Projektphase geplant. In dieser Phase sollen verstärkt Daten zu Werteeigenschaften (Evaluierungsdaten) in das VIR-Informationssystem übernommen werden und damit die Nutzbarkeit vor allem in Hinblick auf eine züchterische Nutzung der in den Kollektionen enthaltenen Diversität erheblich gesteigert werden.

Literatur

LIPMAN, E., M.W.M. JONGEN, TH.J.L. VAN HINTUM, T. GASS AND L. MAGGIONI (eds.) (1997): Central Crop Databases: Tools for Plant Genetic Resources Management. 75-78. International Plant Genetic Resources Institute, Rome, Italy/CGN, Wageningen, The Netherlands.

Entwicklung des deutschen Clearing-House Mechanismus (CHM)

Development of the German Clearing-House Mechanism (CHM)

HORST FREIBERG¹ UND RUDOLF SPECHT²

Abstract

The implementation of the Clearing-House Mechanism (CHM) of the Convention on Biological Diversity (CBD) was decided at the 2nd Conference of the Parties (COP2) in 1995. Art. 18(3) of the Convention on Biological Diversity (CBD) provides the basis for the development of the CHM aimed to promote and facilitate technical and scientific cooperation among the Contracting Parties of the Convention and active partners. Furthermore, through the CHM global access to and exchange of information on biodiversity and its sustainable use will be facilitated. The Information Centre for Genetic Resources (IGR) of the German Centre for Documentation and Information in Agriculture (ZADI) has been contracted in November 1995 by the Federal Ministry of the Environment (BMU) through the Federal Agency for Nature Conservation (BfN) to give support to the development of the German CHM under the Convention on Biological Diversity. The German CHM has contributed actively to the development of the international CHM. The shape and identification of the specific information niche of its national CHM has lead to a concept of three pillars of the CHM: the Information-, Convention- and Cooperation pillar. Through these pillars (1) access to domestic information is made available (2) a monitoring service is provided for the implementation of the CBD and related COP-decisions and (3) besides, emphasis is lead in the development of the German CHM on innovative interactive tools which should stimulate and facilitate scientific and technical cooperation via the CHM.

Aufgaben und Entwicklung des CHM

Für die „Förderung der wissenschaftlichen und technologischen Zusammenarbeit“ sowie die Vermittlung und den Zugriff auf Informationen und Daten, die mit der Umsetzung der Ziele des Übereinkommens über die biologische Vielfalt (ÜBV) (Convention on Biological Diversity - CBD) und der Beschlüsse seiner Vertragsstaatenkonferenzen zusammenhängen, haben sich die Vertragsstaaten den Clearing-House Mechanismus (CHM) in Artikel 18(3) des Übereinkommens geschaffen. Der CHM ist damit die zentrale Informations- und Kommunikationsdrehscheibe des Übereinkommens. In der 2.Vertragsstaatenkonferenz (VSK) im No-vember

¹ Zentralstelle für Agrardokumentation und –information (ZADI)
Informationszentrum für Genetische Ressourcen (IGR)

² Bundesamt für Naturschutz (BfN)

1995 in Jakarta, wurden die informationstechnologischen Eckwerte zum Aufbau des internationalen und nationalen CHM festgelegt. Besonders im Vordergrund steht dabei die Forderung, daß der CHM dezentral, schrittweise und in unmittelbarer Anlehnung an den Bedürfnissen seiner Nutzer und den anfallenden Erfahrungen entwickelt werden soll.

Alle Maßnahmen und Entwicklungen im internationalen CHM sind Bestandteil einer 3-jährigen Pilotphase (1996-1998). Das Sekretariat der Konvention mit seinem Sitz in Montréal „moderiert“ diese Entwicklung. Mit Auslaufen der Pilotphase und ihrer anschließenden Evaluierung in 1999, werden Art und Form der erwarteten Dienstleistungen und Funktionalitäten des CHM sowohl aus nationaler als auch internationaler Sicht klarer herausgearbeitet sein.

Evolution: vom „Clearing-House“ zur „Informationsdrehscheibe“

In der mehrsprachigen Übersetzung des CBD-Textes wird in Artikel 18(3) vom „Clearing-House Mechanismus“(E), „Centre d'Échange“(F) und vom „Vermittlungsmechanismus“(D) gesprochen. Die originäre Bedeutung des Wortes Clearing-House ist der englischen Bedeutung der Clearing-Stelle, einer früher im Bankwesen mit der Verrechnung von Gut- und Lastschriften beauftragten Einrichtung, entlehnt. Einige Länder sehen gerne den CHM in der Rolle eines Maklers, der Angebot und Nachfrage in Verbindung zu bringen versucht. Die Mehrheit der Länder lehnen jedoch diese Makler-Rolle („brokerage-function“) ab und befürworten, daß der CHM transparent und dezentral organisiert wird, auf Meta-Datenebene funktioniert und nur auf öffentlich verfügbare Informationen verweist. Auf der 2. Vertragsstaatenkonferenz der CBD im November 1995 wurde beschlossen, den „Clearing-House“ daher nicht zum Brokerinstrument, sondern zur „Informationsdrehscheibe“ für das Übereinkommen zu entwickeln. Zu Beginn der Verhandlungsrunden sah man die Aufgabe des Clearing-House in der Förderung des Technologie-Transfers mit einem eigens eingerichteten Finanzierungs-Fonds. Erst Anfang 1992 - kurz vor dem Abschluß der Verhandlungen - einigte man sich auf die Förderung der technischen und wissenschaftlichen Zusammenarbeit als die eigentliche Aufgabe des CHM und fand letztlich im Artikel 18 („Technische und wissenschaftliche Zusammenarbeit“) den geeigneten Platz für die Beschreibung seiner Aufgaben.

Deutschland und der CHM

Von Seiten der Bundesregierung, vertreten durch das Bundesumweltministerium (BMU) wird dieser Entwicklungsprozeß des CHM seit der 2.VSK im Rahmen eines Forschungs- und Entwicklungsvorhabens (F&E), für das das IGR der ZADI Auftragnehmer ist, mit fachlicher Betreuung durch das Bundesamt für Naturschutz (BfN) aktiv unterstützt. Über dieses F&E-Vorhaben wurden bisher mehrere international anerkannte Beiträge geleistet, u.a.:

- die Entwicklung eines CHM-Logos, das heute weltweit den CHM visualisiert
- die Übernahme der deutschen CHM-Internet-Leitseite durch das Sekretariat des Übereinkommens der biologischen Vielfalt am Tag seiner offiziellen Eröffnung am 6.5.1996 in Montréal
- die Pressekonferenz am 12.6.1996 mit dem Exekutiv-Sekretär der CBD in Bonn, u.a. zur offiziellen Eröffnung des deutschen CHM
- der Aufbau einer nationalen CHM-Arbeitsgruppe
- der internationale CHM-Workshop Ende Juni 1997 in Bonn, der die Informationsgrundlage für die anschließenden regionalen CHM-workshops in Afrika, Asien, Lateinamerika und den osteuropäischen GUS-Staaten erarbeitete
- der „CHM-Balloon“ ein doppelseitig bedrucktes Informationsblatt, das sowohl in gedruckter als auch in elektronischer Form verteilt wird
- ein gedrucktes CHM-Informationfaltblatt zur dritten und vierten VSK in Buenos Aires 1996 bzw. Bratislava 1998
- eine Tragetasche mit dem CHM-Logo und der Klimaschutz-Initiative des BMU
- verschiedene Poster zum „Deutschen Clearing-House Mechanismus“
- Idee, Konzept und Koordination der „Natur-Detektive“ als Kooperationsprojekt mit Schulen-ans-Netz im Bereich „Umweltbildung/Öffentlichkeitsarbeit“ (Artikel 13, CBD)
- Mitarbeit in der Arbeitsgruppe „Biodiversität“ im G8 Pilotprojekt „Environment and Natural Resources Management“ (ENRM) von BMU und BfN
- Unterstützung des Beratergremiums des Sekretariats des Übereinkommens der biologischen Vielfalt für den CHM
- Dokumentendatenbank mit den englischen Originaltexten der Entscheidungen der 4 VSKs und den deutschen Übersetzungen sowie relevanten Informationsquellen der nationalen Umsetzung dieser Entscheidungen
- Unterstützung zweier Vorhaben der Technischen Zusammenarbeit (TZ) mit Kolumbien und Kamerun beim Aufbau nationaler CHM-Strukturen und der modellhaften Einrichtung einer Internet-gestützten Plattform zur Förderung der wissenschaftlichen und technologischen Zusammenarbeit

Anforderungen an den CHM

Auf der 4. Vertragsstaatenkonferenz (VSK) in Bratislava, im Mai 1998 wurden weitere Herausforderungen an den CHM formuliert, die besonders die Bereiche betreffen:

- Technologietransfer und Capacity Building
- „decision support function“ für die Vertragsstaaten u.a. im Bereich der Trendbeobachtung zur Veränderung der biologischen Vielfalt
- Einbindung des Privat Sektors
- der Entwicklung des CHM zum Informationsinstrument für alle biodiversitätsrelevanten Konventionen (u.a. CMS, CITES, Ramsar)
- der Entwicklung von „interaktiven“ Nutzungsmöglichkeiten und
- Nutzung des CHM als Instrument der Bewußtseinsbildung bei der Bevölkerung für die Ziele der Konvention.

Dies erfordert von allen Vertragsstaaten Anstrengungen, ganz gezielt konkrete Beispiele aus ihrer nationalen Arbeit zur Umsetzung der Beschlüsse der VSK sowohl zu den einzelnen Artikeln, den wichtigen Themen und den übergeordneten Zielen der Konvention vorzulegen. Diese werden dann über den CHM allen Staaten verfügbar gemacht.

Die Informationssäulen der deutschen Homepage

Die deutsche Homepage des CHM (<http://www.biodiv-chm.de>) soll dem individuellen Nutzer möglichst einen für ihn „intuitiven“ Zugang zu den vorhandenen Informationsquellen der biologischen Vielfalt in Deutschland, aber auch weltweit, gestatten (s. Abb. 1). Die deutsche Homepage ist zweisprachig aufgebaut. Die Originalstartseite ist deutschsprachig mit einem Hinweis auf die englischsprachige Kopie.

Die zentralen Informationsfelder der Startseite sind in drei „Säulen“ aufgeteilt, die für die künftige Entwicklung des CHM als richtungsweisend betrachtet werden:

Unter dem Bereich „**Informationsdienste**“ stehen allgemein interessante Informationsquellen wie Datenbanken, Presseinformationsquellen usw..

Unter dem Titel „**Konvention**“ werden schrittweise, neben den Originaltexten der Entscheidungen der Vertragsstaatenkonferenzen der Konvention, auch die Übersetzungen dieser Originaltexte in deutsch, gemeinsam mit Verweisen bzw. Dokumenten der nationalen Umsetzung dieser Beschlüsse ab 1999 zugänglich gemacht werden. Hier findet der Besucher auch den Einstieg über „Artikel und Themen“ zu allen wesentlichen nationalen Informationsquellen

mit direktem Bezug entweder zu den zentralen Artikeln des Übereinkommens - wie u.a. Artikel 8 + 9 (*In-situ*- und *Ex-situ* Erhaltung) oder zu Artikel 15 (Genetische Ressourcen). Oder der Nutzer hat die Möglichkeit, über eines der großen Themen des Übereinkommens, wie u.a. Genetische Ressourcen, Biotechnologie, Agrobiodiversität zu den entsprechenden Informationsquellen zu gelangen. Hierfür wurde für jedes Thema eine entsprechende „Themenseite“ eingerichtet, die nicht als fortlaufend anwachsende Linksammlung aufgebaut ist sondern eine gezielte Auswahl von national ergiebigen Informationsquellen darstellt. Diese Liste wird ganz bewußt in dieser kompakten Form geführt und soll möglichst nicht über eine Seite anwachsen. Diese Themenseiten sollen künftig zusammen mit dem deutschen CHM-Balloon, dem elektronischen Informationsdienst des deutschen CHM, ein gebündeltes aktuelles Informationsmedium dem Nutzer anbieten. Angestrebt ist hierbei, daß die Themenseiten von Mitgliedern der deutschen CHM-Arbeitsgruppe oder aber auch anderen Institutionen moderiert werden und aus dieser Querschnittssicht aktuelle Kurztextbeiträge für den CHM-Balloon fortlaufend angeboten werden können. Damit soll der Balloon zum aktuellen Informationsdienst des deutschen CHM weiter entwickelt werden.

Unter dem Titel „**Kooperation & Dialog**“ werden schließlich insbesondere Aktivitäten entwickelt, die einen konkreten und praktischen Beitrag zur direkten Umsetzung des Artikel 18.3 „Förderung der wissenschaftlichen und technologischen Zusammenarbeit“ darstellen. Hier werden 1999 zwei bilaterale Kooperationsvorhaben mit Kolumbien und Kamerun, finanziert im Rahmen des Sektorvorhabens „Umsetzung der Biodiversitätskonvention“ der Gesellschaft für Technische Zusammenarbeit (GTZ) eine zentrale und richtungsweisende Rolle einnehmen.



Abb. 1: Homepage des CHM

Fig. 1: Homepage of CHM

Die deutsche CHM-Arbeitsgruppe

Für eine breitgetragene Umsetzung des deutschen CHM hat das Bundesumweltministerium bereits Ende 1995 eine nationale Arbeitsgruppe eingerichtet. Die Mitglieder der Arbeitsgruppe setzen sich aus verschiedenen gesellschaftlich eingebundenen Institutionen und Gruppen zusammen: dem BMU (als federführendem Ressort) gemeinsam mit dem Bundesamt für Naturschutz (BfN) und dem Umweltbundesamt (UBA), der Deutschen Gesellschaft für Technische Zusammenarbeit (GTZ), der Zentralstelle für Agrardokumentation und –information (ZADI) und dem Informationszentrum für Genetische Ressourcen (IGR), der Arbeitsgruppe „Biologische Vielfalt“ des Forums Umwelt & Entwicklung (NRO), dem Botanischen Institut der Uni Bonn (für die Universitäten) sowie dem Internationalen Transferzentrum für Umwelttechnik (ITUT) als Vertreter des privat-industriellen Sektors. Alle Vertreter verfügen ihrerseits wiederum aus ihrem Arbeitsfeld über ein breites Informationsangebot, das sie im Rahmen des deutschen CHM weiter ausbauen; die Arbeitsgruppe berät außerdem die nationale Kontaktstelle des CHM in der Weiterentwicklung des nationalen Informationssystems.

Vision: Clearing House Mechanismus 2000

Der CHM kann nicht auf dem Entwicklungsniveau eines „bloßen“ Verweissystems verharren. Das Verweissystem ist notwendig und auch grundlegend, denn erst mit ihm wird eine „evolutive Ausgangssubstanz“ geschaffen, mit dem der Wissensraum der biologischen Vielfalt in seiner vernetzten Komplexität sektor- und fachübergreifend erschlossen werden kann. Hierzu zählen:

- die Entwicklung der Nationalen Berichte (Artikel 26) zu zentralen Informationsquellen
- die Visualisierung und geo-referenzierte Verschneidung von Meta-Informationen gleich welcher Herkunft zu Trendabschätzung, Monitoring und Indikatorenbestimmung der Veränderung der biologischen Vielfalt
- das „fraktale Zoomen“ über aggregierte, flache Informationsstrukturen globalen Maßstabes zu vielschichtigen Detailstrukturen auf nationaler und sogar lokaler Ebene
- die Integration und Vernetzung mit globalen und regionalen Informationsstrukturen über die Nutzung weniger „Schlüsselwörter“ (Attribute list) die ihrerseits über existierende Thesauri neue Informationsräume öffnen (G7 ENRM, OECD Megascience Initiative)
- die Entwicklung eines „CHM Tool-Kit“ zur Unterstützung der Partner des CHM bei Aufbau, Integration und Austausch eigener Informationsbestände in einer miteinander abgestimmten Grundinformationsstruktur des CHM
- die Entwicklung neuer, nutzerfreundlicher Such- und Abfragemechanismen, die den Informationsraum der Biodiversität schrittweise selbständig erschließen helfen.

So kann der Clearing-House 2000 zu einer global gelenkten Informationsdrehscheibe in gemeinsamer Verantwortung aller Vertragsstaaten werden.