

Lothar Seidewitz

*Institut für Pflanzenbau und Pflanzenzüchtung der FAL,
Braunschweig-Völkenrode*

DATENDOKUMENTATION ZU PFLANZENGENETISCHEN POTENTIALEN
IM INSTITUT FÜR PFLANZENBAU UND PFLANZENZÜCHTUNG DER FAL

Einleitung

Der Aufbau einer sogenannten "Genbank" im Institut für Pflanzenbau und Pflanzenzüchtung der FAL entspricht dem Mitte der 60er Jahre geäußerten Wunsch der Pflanzenzüchter nach einem Dokumentations- und Informationssystem in Verbindung mit der langfristigen Erhaltung pflanzen genetisch wertvollen Materials. Mit dieser Einrichtung sollte die Bereitstellung beschriebenen Materials aus den entsprechenden Sammlungen pflanzen genetischer Potentiale für verschiedene Verwendungszwecke ermöglicht werden. Ferner sollten Informationen und Resultate aus Evaluierungsarbeiten zu einem gezielten Nachweis von Ausgangsmaterial beitragen. Die mit dieser Zielsetzung verbundenen Aktivitäten des Instituts für Pflanzenbau und Pflanzenzüchtung der FAL führten zum Aufbau einer gegenwärtig etwa 30.000 Muster umfassenden Sammlung pflanzen genetischer Potentiale sowie zur Entwicklung eines Dokumentations- und Informationssystems.

Aufbau des Dokumentations- und Informationssystems

Wie bereits hervorgehoben wurde, haben die Genbankaktivitäten des Instituts für Pflanzenbau und Pflanzenzüchtung zum Ziel, pflanzen genetische Potentiale einerseits langfristig zu erhalten und sie andererseits für Evaluierungs- und Züchtungsvorhaben bereitzustellen. Aus diesem Aufgabenkomplex resultieren zahlreiche Informationen, die in Form von Daten die Grundlage eines Dokumentations- und Informationssystems bilden. Dieses System wurde unter dem Gesichtspunkt der Unterstützung sowohl institutsinterner als auch -externer Arbeiten konzipiert. Im Mittelpunkt des Systems stehen eine computerunterstützte Datenbank und ein entsprechendes Datenbanksystem, das für die Speicherung von Daten sowie für die selektive Wiedergewinnung von Daten und Informationen dient.

Für die Unterstützung der sowohl institutsintern als auch -extern durchgeführten Bearbeitung pflanzen genetischer Potentiale durch das Dokumentationsystem des Instituts stehen Daten aus verschiedenen Quellen zur Verfügung. Diese Quellen können analog in 2 Gruppen als

institutsinterne und
institutsexterne Daten-Quellen

aufgeteilt werden. Genbankaktivitäten sind in besonderem Maße durch eine weitreichende Zusammenarbeit nationaler und internationaler Institutionen gekennzeichnet. Daraus ergibt sich ein Austausch von Material und Informationen (Daten), dem auf nationaler und internationaler Ebene in dem Maße immer größere Bedeutung zukommt, wie die unter natürlichen Bedingungen verfügbare Formenmannigfaltigkeit als Ausgangsbasis für Züchtungsarbeiten abnimmt. Insbesondere mit dem Datenaustausch und der Einbeziehung fremder Daten in die eigenen Arbeiten wird ein umfassender und gezielter Zugriff auf international verfügbares Genmaterial ermöglicht.

Eine hohe Wahrscheinlichkeit des Nachweises relevanter Informationen (Daten) und des den jeweiligen Erfordernissen entsprechenden Materials macht eine Zentralisierung bzw. Konzentration von Daten zwingend erforderlich. Da die Darstellung von Erkenntnissen und Sachverhalten zu

pflanzen genetischen Ressourcen üblicherweise unter individuellen Gesichtspunkten erfolgt, treten bei der Zusammenführung von Daten die in der Dokumentation bestehenden Unterschiede auf logischer wie auf technischer Ebene zutage. Dieser Umstand erklärt die Notwendigkeit eines koordinierenden Wörterbuches und entsprechender Regeln für die austauschgerechte Darstellung von Sachverhalten. Dieser Notwendigkeit entspricht die

Entwicklung eines "Thesaurus für die internationale Standardisierung der Dokumentation von Genbanken".

Mit diesem Thesaurus verbinden sich zwei Ziele:

- 1) die Schaffung einer semantisch-terminologischen Grundlage für die verarbeitungsgerechte Darstellung von Sachverhalten einerseits sowie für die selektive Informationsgewinnung andererseits;
- 2) die Erleichterung kommunikativer Beziehungen zwischen Institutionen verschiedener Länder, die bei der Sammlung, Erhaltung, Evaluierung und Nutzung pflanzen genetischer Potentiale zusammenarbeiten.

Damit dient der Thesaurus sowohl als Grundlage für die eigene Arbeit als auch als koordinierendes Wörterbuch für den Austausch von Daten und Informationen. Er unterstützt die im Dokumentationsprozeß wichtige Kontrollfunktion des Sachbearbeiters und trägt mit der Definition entsprechender Regeln zur Überwindung der verschiedenen Aspekte der Inkompatibilität zwischen bestehenden und sich entwickelnden Dokumentationssystemen bei.

Die inzwischen aus über 1.000 Deskriptoren bestehende Dokumentations-sprache macht eine Neuauflage des Thesaurus erforderlich, mit der inzwischen begonnen wurde. Diese Neuauflage folgt der bisherigen Konzeption, die eine Aufteilung des Wortgutes auf insgesamt 6 Thesaurusteile vorsieht:

- I Getreide, einschließlich Mais, Reis und Hirse
- II Futterpflanzen, einschließlich Rasengräser
- III Hackfrüchte
- IV Gemüse, einschließlich Öl- und Faserpflanzen
- V A Krankheiten, Krankheitserreger und Schädlinge
- V B landläufige und wissenschaftliche (taxonomische) Pflanzennamen
- VI nicht-fruchtartenspezifische Deskriptoren.

Es ist ein Sonderteil als technisches Wörterbuch für die Definition von Begriffen aus der Dokumentation und Datenverarbeitung vorgesehen.

Die gesamte Thesaurusarbeit stützt sich neuerdings auf den Einsatz der elektronischen Datenverarbeitungstechnik. Damit ist es möglich, revidierte Neuauflagen binnen kurzer Zeit zu erstellen. Ebenso lassen sich mehrsprachige Thesaurusausgaben leichter bearbeiten. Die Thesaurus-Ausgabe in mehreren Sprachen verfolgt das Ziel, die Datenbank des Dokumentationssystems zukünftig einem simultanen Zugriff mittels Suchbegriffen verschiedener natürlicher Sprachen zugänglich zu machen. Die seitens der Europäischen Gemeinschaften geförderte Entwicklung von EURONET wird zukünftig die erforderliche technische Voraussetzung für die Realisierung dieser Zielsetzung bieten.

Informationsaufnahme, Datenorganisation und Informationsgewinnung

Für das im Institut für Pflanzenbau und Pflanzenzüchtung der FAL aufgebaute Dokumentations- und Informationssystem gilt dasselbe Prinzip wie für jedes andere auf faktische Daten gestützte System, indem allen Prozessen der Speicherung, Verarbeitung und Wiedergewinnung eine strukturierte Form von Informationen zugrundeliegt.

Die Datenbankkonzeption stützt sich auf ein sogenanntes "interaktives Datenbanksystem" (oder Datenbankprozessor) und auf eine entsprechend

leistungsfähige Computer-Anlage. Das Datenbanksystem wird als die Gesamtheit aller Programme verstanden, und zwar für folgende Aufgaben:

- Organisation der Datenspeicherung
- Datenwiedergewinnung und
- Datenklassifizierung
- interne Reorganisation der Datenbank nach Datenänderung oder Datenaufnahme
- Datenschutz
- Dialogbetrieb.

Unter Datenbank wird in diesem Zusammenhang die Gesamtheit aller für die Genbankaktivitäten wichtigen Daten verstanden. In diesem Beispiel wird die Datenbankarbeit mit dem Prozessor "FIDAS" der Gesellschaft für Mathematik und Datenverarbeitung (GMD), Bonn, abgewickelt.

Datenbankarbeit als Handhabung von Informationen zu pflanzengenetisch wertvollem Material ist aus der Sicht der gewonnenen eigenen Erfahrung mit dem zu bearbeitenden Objekt gleichbedeutend mit einer umfangreichen Dateiverwaltung. Gemäß ihrer Herkunft aus internen und externen Quellen werden die Daten 2 verschiedenen Kategorien von Dateien zugeordnet:

- Dateien mit Direktdaten
- Dateien mit Hinweisdaten.

Die "Direktdaten" beziehen sich auf das in Braunschweig unmittelbar verfügbare Material, während "Hinweisdaten" auf das in anderen Sammlungen pflanzengenetischer Potentiale oder Genbanken verfügbare Material verweisen.

Während der Thesaurus des Dokumentationssystems die Darstellung von Sachverhalten und den Datenaustausch regelt, wird der Zugriff auf die Datenbank durch eine besondere, in Vorbereitung befindliche Publikation geregelt. Diese gibt Auskunft über die tatsächlich abrufbaren Daten, ihre Lokalisierung in den einzelnen Dateien des Systems sowie die Art ihrer Adressierbarkeit über die entsprechenden Deskriptoren.

Das hier skizzierte Beispiel einer Datensammlung und dementsprechender Computer-Anwendung dokumentiert den besonderen Beitrag, den das Institut für Pflanzenbau und Pflanzenzüchtung der FAL sowohl zur langfristigen Sicherung pflanzengenetischer Potentiale als auch zur Entwicklung einer kommunikativen Grundlage für die weltweite, netzwerkartige Zusammenarbeit der auf diesem Gebiet tätigen Institutionen leistet.

DISKUSSIONSBEITRAG

BRUNDKE: Besteht zwischen dem Thesaurus FAL und dem von anderen Stellen eine Übereinstimmung?

v. SEIDWITZ: Nein, aber, so HAAS (Luxemburg), dies ist nur eine EDV-technische Frage der Zuordnung von einem Thesaurus: Was ist Deskriptor und was ist synonym?