

Natur im Museum

**Heft 4
2014**

Herausgeber:

Norbert Niedernostheide, Dr. Silke Stoll

Redaktion & Schriftleitung:

Prof. Dr. Annette Scheersoi

ISSN: 2191-6632



**Mitteilungen der Fachgruppe
Naturwissenschaftliche Museen
im Deutschen Museumsbund**

Sammlungsschätze online – Schritte zur Virtuellen Naturhistorischen Sammlung

Dagmar Triebel, Tanja Weibulat & Konstanze Bensch

Die Archive der Naturwissenschaftlichen Museen Deutschlands sind eine wahre Schatztruhe, gefüllt mit einer enormen Vielfalt einmaliger Objekte und historisch gewachsener Sammlungen. Nur ein Bruchteil dieser Objekte kann in den räumlich oft begrenzten Ausstellungsräumen gezeigt werden. Moderne Präsentationstechnik und Museumspädagogik verzichten sogar explizit auf eine Anhäufung vieler originaler Schaustücke, um stattdessen wenige oder gar nur Nachbildungen derselben in einen größeren thematischen Kontext zu stellen.

So finden sich in den Archiven auch kleinerer Museen in Deutschland oft wichtige Exponate bzw. größere Objektsammlungen von Tieren, Pflanzen und Pilzen, die vielleicht gut kuratiert, jedoch nur unvollständig inventarisiert, kaum erschlossen und für die Öffentlichkeit schlecht zugänglich sind. Neben dem historischen und musealen Wert dieser Archivbestände sind die Sammlungsstücke aber auch für die biologische Wissenschaft und Forschung von großem Interesse.

Einige naturwissenschaftliche Sammlungen begannen mit der digitalen Erschließung ihrer Bestände bereits in den 1980er und 1990er Jahren, indem

sie die Daten aus Inventarverzeichnissen oder Begleitdokumenten/Etiketten in die ersten relationalen Datenbanken, z. B. dBase, eingaben. Wertvolle Einzelobjekte, wie Typenexemplare, wurden fotografisch dokumentiert. Erst ab dem Jahr 2000 wurden in Deutschland größere Infrastrukturprojekte zur Bewältigung dieser Mammutaufgabe durch das Bundesministerium für Bildung und Forschung finanziert.

Objektdigitalisierung und Datenerschließung

Deutsche Museen unter öffentlicher Trägerschaft beherbergen schätzungsweise 140 Millionen naturkundliche Objekte (DNFS 2013) an etwa hundert Standorten. Die Objektdigitalisierung und Datenerschließung stehen heute nach wie vor am Anfang. Man schätzt, dass bisher nur 2–5 % der Bestände digital erschlossen werden konnten. Bedingt ist dies – neben fehlender Finanzierung – vor allem durch den enormen Aufwand und die Komplexität der Aufgaben.

Die einzelnen Einrichtungen beherbergen oft eine große Vielfalt von Objekten (Abb. 1). Diese werden aufgrund ihrer unterschiedlichen Eigenschaften, Objektgröße und Konservierungszustand in verschiedener Weise aufbewahrt. So existieren getrocknete Sammlungsobjekte und in

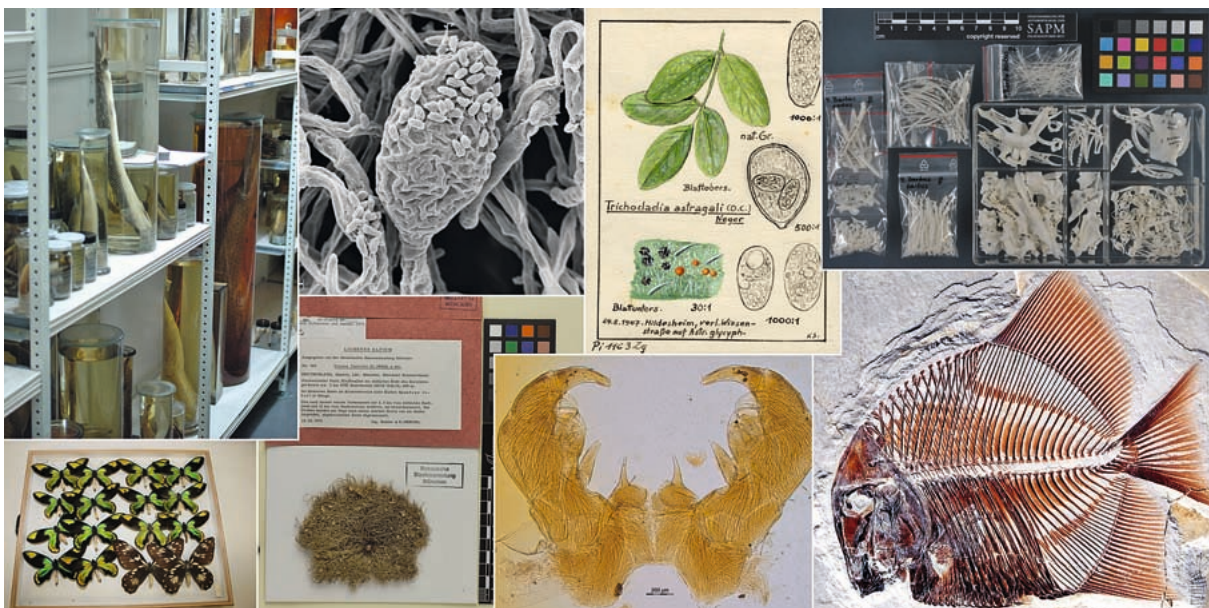


Abb. 1: Diversität der Bestände naturhistorischer Sammlungen.

Hochauflösende Kameraeinrichtung mit softwaregesteuert bewegbarem Tisch optimiert für dreidimensionale Makroobjekte (Kleinpilze) und hohe Schärfentiefe

Contents [hide]

- 1 High-resolution camera system with software-controlled movable table optimized for three-dimensional macro objects (small fungi) and high depth of field
- 1.1 Camera hardware
- 1.2 Lighting
- 1.3 Digitization workstation - hardware and software
- 1.4 Digitizing server infrastructure
- 1.5 Digitizing workflow
- 1.6 Contact person(s)

High-resolution camera system with software-controlled movable table optimized for three-dimensional macro objects (small fungi) and high depth of field

Camera hardware

- DSLR: Nikon D2X (12.5 Megapixel)
- Macro lens: Micro Nikkor 60 mm 1:2.8 D

Lighting

- 1 x CH-126 LED Video Lighting (5400 K, 650 lx, diffuser filter)
- 2 x LED-5001 Video Light (5500 K, 700 lx, diffuser filter)

Digitization workstation - hardware and software

Computer hardware:

- Intel Core i7-870 CPU (8M Cache, 2.93 GHz)
- 8 GB RAM
- 2 x Western Digital WD5002ABYS-02B1B0 (500 GB, SATA2) RAID 0 HDD
- NVIDIA Quadro 2000 graphics card (1 GB DDR5 frame buffer)

ABOVE: Artificial image constructed from 18 single photo applying focus stacking and super-resolution processing.

LEFT: Original image shot with Nikon D2X DSLR and Micro Nikkor 60 mm 1:2.8 D macro lens.

Abb. 2: Beschreibung einer Digitalisierungsstation an der Botanischen Staatssammlung München (<http://wiki.gbif.de>).

Vielen konservierte Alkoholpräparate, in Schubladen gelagerte Fossilien und Mineralien, große Schauexponate oder auf Objektträgern eingebettete Mikropräparate, z. B. von Kieselalgen, Tiergewebe und Sporen, sowie Organpräparate. Ferner gibt es Fell-, Geweih- und Vogelbalgsammlungen sowie Präparate von Fischschuppen und mikroskopisch kleinen Ohrsteinen, Insektenkästen und Herbarbögen mit gepressten Pflanzen. Begleitende Dokumente werden separat oder objektnah aufbewahrt.

Die Digitalisierung der geschilderten großen und kleinen Sammlungsschätze erfordert in verschiedener Hinsicht optimierte Arbeitsstationen. Beschreibungen einiger Stationen zur digitalen Fotografie, wie sie an den Forschungssammlungen in Berlin, Bonn, Braunschweig und München (BGBM, DSMZ, MfN, SNSB und ZFMK) betrieben werden, können auf der Wiki-Plattform von GBIF Deutschland unter <http://wiki.gbif.de> eingesehen werden (Abb. 2). Dort finden sich auch Schilderungen zum Arbeitsablauf, zur verwendeten Geräteausstattung sowie den eingesetzten Datenbanken zur Aufnahme von Metadaten, zur benötigten mittleren Arbeitszeit pro Objekt sowie zu den Ansprechpartnern vor Ort.

Die fotografische Aufnahme von Objekten ist meist gekoppelt mit der Metadatenerfassung. Letztere kann aber auch unabhängig durchgeführt werden, zunächst als reine Erhebung von Daten aus Dokumenten, Etiketten und Inventarbüchern. Seit Beginn der Entwicklung relationaler Datenbanken kommen verschiedene Datenbanksysteme und Software-Anwendungen zum Einsatz. Die ersten IT-Werkzeuge waren zumeist sowohl hinsichtlich ihrer technischen Spezifikationen als auch ihrer Bedienbarkeit mehr oder weniger begrenzt. Sie wurden nach den Anforderungen der Mitarbeiter der jeweiligen Institution konzipiert und waren dementsprechend wenig flexibel.

Standardisierte Arbeitsflüsse und Daten-Pipelines, welche die Schritte der Objektdigitalisierung und Datenerschließung miteinander verbinden und beide Informationsschichten zusammen über ein gemeinsames Standard-Austauschformat den Netzwerken und Internetportalen bereitstellen, wurden bisher selten eingerichtet. Dies kann am besten von Fachdaten-Servicezentren mit umfangreicher Datenbank-, Server- und Netzwerkadministration geleistet werden. Derzeit fördert die DFG Infrastrukturprojekte, welche integrierte Datenflüsse für zoologische, mykologische,

botanische und paläontologische Sammlungen an deutschen Museen entwickeln. Beispiele hierfür sind das Verbundprojekt IDES (<http://ides.snsb.info>) für zoologische und paläontologische Sammlungen sowie das Verbundprojekt BiNHum „BiodiversitätsNetzwerk des Humboldt-Rings“ für Sammlungen der Humboldt-Ring-Institutionen (<http://wiki.binhum.net/>).

An mehreren naturwissenschaftlichen Sammlungen werden zurzeit Vorstudien zur Etablierung fließbandunterstützter Arbeitsflüsse für Sammlungsobjekte in Standardformaten oder Standard-Schubläden durchgeführt. Voraussetzungen sind dabei eine logistisch aufwändige Organisation der Materialverlagerung sowie präzise organisierte Arbeitsabläufe. Eine Rentabilität dieser Anlagen ist nur bei hoher Objektzahl denkbar. Nach der Objektdigitalisierung wartet dann noch die immense Aufgabe der Bearbeitung der Digitalisate wie Herbarbögen oder Insektenkästen und die Erschließung der Information zu den Digitalisaten, d. h., die Erfassung der Metadaten in Datenbanksystemen sowie ein in sich nachhaltiges Datenmanagement.

Datenkuration, Datenhosting und Nachhaltigkeit

Um die Qualität der erhobenen Objekt-assoziierten Fachdaten zu verbessern, werden Objektdigitalisierung und Datenerschließung durch entsprechende Maßnahmen und Arbeitsschritte begleitet. Dazu gehören Qualitätskontrolle, Korrekturverfahren und eine Verknüpfung der Datenbestände mit Vokabularen, Taxonomien und Thesauri. Wesentlich ist dabei eine klare Trennung zwischen der Erhebung von Originaldaten und deren Interpretation sowie einer datenbankinternen Dokumentation der Historie der vorgenommenen Änderungen. Hier spielen die Konzepte einer modernen Datenhaltung in relationalen SQL-Datenbanksystemen und die verwendeten Werkzeuge eine große Rolle.

Ein Beispiel hierfür ist die Sammlungsdatenbank DiversityCollection, eine der derzeit dreizehn Datenbankkomponenten des Datenmanagementsystems Diversity Workbench (DWB; <http://www.diversityworkbench.net>, Abb. 3). Sie werden unter Koordination des IT Centers der Staatlichen Naturwissenschaftlichen Sammlungen Bayerns (SNSB) entwickelt und sind frei verfügbar. Bei sämtlichen

Datenbanken handelt es sich um SQL-Server-Datenbanken mit Rich Clients für den Datenzugriff. Sie können so installiert werden, dass sie mit den anderen DWB-Datenbanken verbunden sind. Die DWB erlaubt die Verwaltung von Multimedia-Daten wie Scans von Etiketten, hochauflösenden digitalen Fotos, Audio- und Video-Files zusammen mit objekt-bezogenen Fachdaten in ein- und demselben Datenbank-Framework. Virtuelle Arbeitsumgebungen und Daten-Pipelines können über Client-Server-Netzwerke betrieben werden. Management-Werkzeuge zum Import und Export existieren.

Die DWB ist an mehreren naturhistorischen Sammlungen und größeren Museen als unabhängige Installation für die langfristige Datenkuration im Einsatz. Kleinere Museen und Herbarien lassen ihre Datenbestände bisher ganz oder auch nur teilweise durch das Datenzentrum der SNSB hosten, im Allgemeinen in Datenbanken, welche auf virtuellen Servern eingerichtet sind. Dies erlaubt eine langfristige Kuration der jeweils eigenen Datenbestände und die Nutzung zusätzlicher Dienste des SNSB IT Centers wie die der Langzeit-Archivierung. Das Konzept des Datenhostings in virtuellen Arbeitsumgebungen als Service für naturwissenschaftliche Sammlungen wird seit 2010 an den SNSB im Rahmen des BMBF-Verbundvorhabens GBIF-D: ‚Kompetenzzentren innovativer Datenmobilisierung‘ (<http://www.gbif.de>) umgesetzt und derzeit im Rahmen des DFG-Verbundprojektes BiNHum erweitert bzw. den Bedürfnissen neuer Partner angepasst.

Datenintegration in Internetauftritten von Museen, Wissensportalen, nationalen und internationalen Netzwerken sowie Technologietransfer

Ein wichtiger Schritt zur Erschließung der Sammlungsschätze in Deutschland ist die Integration der mobilisierten Datenbestände in Wissensportale und internationale Forschungsnetzwerke, aber auch in institutseigene Internetauftritte der jeweiligen Museen und Sammlungen. Um diese Aufgaben auf nationalem Niveau leisten zu können und entsprechende Daten-Pipelines zu etablieren, ist ein dauerhafter Erfahrungsaustausch und Technologietransfer zwischen den verschiedenen Einrichtungen erforderlich. Wesentlich sind dabei Workshops und Trainingskurse, im Rahmen derer Sammlungsmitarbeiterinnen, d. h., Wissenschaft-

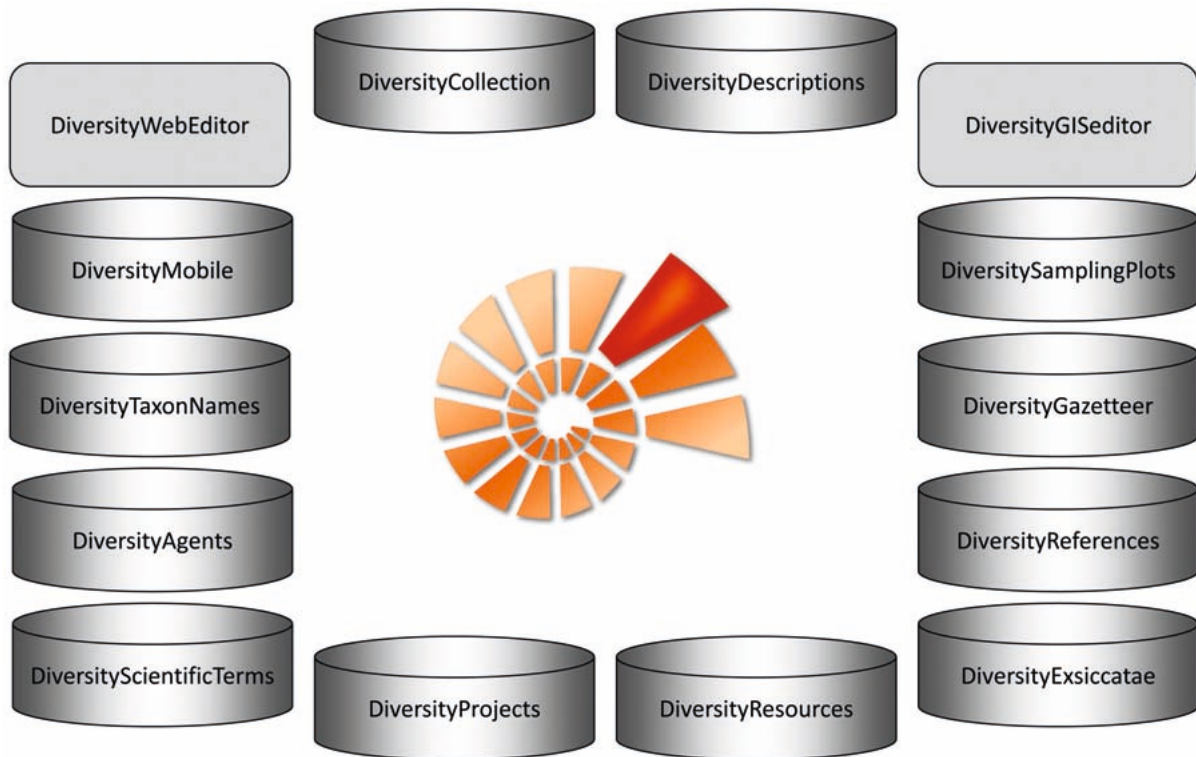


Abb. 3: Die 13 Komponenten der Diversity Workbench. Jede Datenbank ist eigenständig und für das Management eines Fachdaten-Typs zuständig, z. B. DiversityDescriptions für strukturierte Beschreibungen von Objekten und Organismen.

lern, Präparatoren und Datenkuratoren, Kompetenz im Umgang mit den verschiedenen IT-technischen Werkzeugen vermittelt wird. Die regelmäßigen eintägigen Workshops in München haben dies zum Ziel (<http://www.snsb.info/Workshops.html>).

Arbeitsumgebungen und Daten-Pipelines, zugeschnitten auf die Bedürfnisse kleinerer Museen, können zunächst an Datenzentren größerer Einrichtungen, wie z. B. am SNSB IT Center (siehe Abb. 4) aufgebaut und gehostet werden. Nach Aufbau der entsprechenden Infrastruktur an den kleineren Institutionen können die Datensammlungen und IT-Infrastruktur dann dorthin transferiert werden. Eine Alternative dazu ist das dauerhafte Hosting der Datenbestände an größeren Datenzentren, um Synergien zu nutzen und kein IT-Fachpersonal an kleineren Einrichtungen binden zu müssen. Derartige Konzepte werden zurzeit im DFG-geförderten Verbundprojekt „German Federation for the Curation of Biological Data“ (GFBio) vertieft diskutiert und im BiNHum-Projekt bereits pilothaft umgesetzt.

Gemeinsames Ziel aller naturwissenschaftlicher Sammlungen Deutschlands sollte es sein, ihre

Datenbestände aus datenkuratorisch gut betreuten Datenbanken möglichst aktuell und dynamisch bereitzustellen. Standardschemata wie ABCD und Wrapper-Software, die diese Schemata nutzen, z. B. BioCAsE (<http://www.biocase.org/>), können dabei helfen, die Daten gemeinsam online zur weiteren Nutzung anzubieten. Beispiele dafür sind das Datenportal der Global Biodiversity Information Facility (GBIF) und – in Zukunft – auf nationaler Ebene Portale, wie sie im Rahmen von BiNHum und GFBio entstehen sollen. Die Integration von Datenbeständen einzelner Museen in deren institutseigene Internetauftritte kann parallel dazu erfolgen.

Ausblick

Angeht die wachsende Komplexität der Aufgaben, wie sie im Bereich Biodiversitätsinformatik, Sammlungsdigitalisierung und -erschließung auf die naturwissenschaftlichen Museen zukommen, werden flexible Systeme zum Management von Sammlungsdaten, Terminologien und primären Forschungsdaten, wie z. B. Messreihen und Beschreibungen von Organismen und deren Merkmalen, benötigt. Diese Systeme werden sich sowohl technisch wie auch konzeptionell lau-

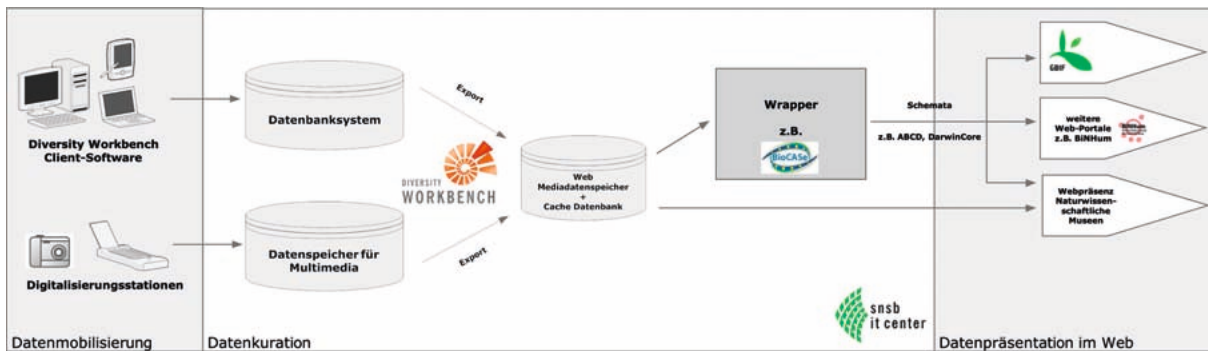


Abb. 4: Beispiel für einen Datenfluss von der Digitalisierung bis zur virtuellen Sammlung im Web

fend weiterentwickeln. Das beteiligte technische und wissenschaftliche Personal wird sich ständig weiterbilden, wobei sich vielleicht das Berufsbild eines Fachdaten-Kurators im Bereich naturwissenschaftlicher Sammlungen herausbilden wird.

Die generierten Informationen sollen als „Virtuelle Naturhistorische Sammlungen“ der Gesellschaft und Öffentlichkeit zur Verfügung stehen. Sie sollen aber auch Teil der internationalen Wissensnetzwerke und der Megascience-Plattformen sein (TRIEBEL et al. 2012). So können die Daten zu Sammlungsschätzen aus den naturhistorischen Museen internationalen Großprojekten wie dem Barcode of Life-Projekt (iBOL) oder dem INSDC Konsortium zur Verfügung gestellt werden und die Sammlungen und Museen aufgrund ihrer fachlichen Expertise wichtige Partnereinrichtungen bei der Mobilisierung von Datenbeständen für die Forschung sein.

Projekte zur Sammlungsdigitalisierung und dem Aufbau entsprechender IT-Infrastruktur an den Staatlichen Naturwissenschaftlichen Sammlungen Bayerns wurden in den letzten Jahren durch das Bayerische Staatsministerium für Wissenschaft, Forschung und Kunst, das Bundesministerium für Bildung und Forschung (BMBF), Projekt 01

LI 1001 B 'GBIF-D', die Deutschen Forschungsgemeinschaft (DFG) INST 747/1-1 u. a. (IBF), HA 2598/16-1 (BiNHum), TR 290/5-1 u. a. (IDES) und TR 290/7-1 (GFBio) sowie die Andrew Mellon-Foundation, Projekt „Global Plants Initiative“ gefördert. Für die Bereitstellung von Bildmaterial danken die Autoren der Bayerischen Staatssammlung für Anthropologie und Paläoanatomie, der Botanischen Staatssammlung München, dem Juramuseum Eichstätt und der Zoologischen Staatssammlung München.

Literatur

- Pressemitteilung der DNFS (Deutsche Naturwissenschaftliche Forschungssammlungen e.V.) und des DMB (Deutscher Museumsbund), 19.2.2013: Deutschlands naturkundliche Sammlungen – Die Erhaltung der Vielfalt ist gesellschaftlich bedeutsam. Online via: http://www.dnfs.de/sites/default/files/pm_tagung_naturkundliche_sammlungen_berlin15.02.2013.pdf
- TRIEBEL, D., HAGEDORN, G. & RAMBOLD, G. (2012): An appraisal of megascience platforms for biodiversity information. In: *Mycology* 5, S. 45–63. doi:10.3897/mycokeys.5.4302; pdf.

Dr. Dagmar Triebel, Tanja Weibulat, Dr. Konstanze Bensch
SNSB IT Center und Botanische Staatssammlung München
Menzinger Straße 67
80638 München
triebel@bsm.mwn.de