



Gemeinsame Nutzung von Daten am Beispiel von Biodiversitätsdaten

Gabriele Dröge, Neela Enke, Anton Güntsch & Walter Berendsohn

Abteilung Biodiversitätsinformatik & Labors
Botanischer Garten und Botanisches Museum Berlin-Dahlem
Freie Universität Berlin

Biodiversität



Naturhistorische Sammlungen



Naturhistorische Sammlungen

50-100.000 Naturhistorische Sammlungen weltweit
mit **1,5-2 Milliarden** Belegen

Primary Biodiversity Data is:

The digital text or multimedia data records that detail the instance of an organism – the '**what, where, when, how and by whom**' of the organism's occurrence and recording.

Primäre Biodiversitätsdaten

- Lebendes Exemplar
- Konserviertes Exemplar
- Beobachtung
- Multimedia-Dokument (Zeichnung, Foto, Tonaufnahme, Video)
- Literatur

→ 3-10 Milliarden Biodiversitätsdaten



Global Biodiversity Information Facility

www.gbif.org

Gründung: **2001**, Planung seit 1996

Sekretariat: Kopenhagen, **dezentrale Knotenstruktur**

Heute: Teilnahme von **55 Ländern** und **45 internationalen Organisationen**

Zugriff auf über **267 Millionen** Biodiversitätsdaten

GBIF - Aufbau

Dezentrale Struktur

- > nationale Knoten, innerhalb Deutschlands 8 Themen-Knoten (Prokaryonten, Botanik, Mykologie, Evertebraten I-III, Vertebraten, Paläontologie)
- > Daten werden lokal verwaltet
- > Datenprovider behält volle Kontrolle über seine Daten

Wrapper-Technologie (Provider Software)

- > Datenbanken können sehr heterogen im Aufbau sein
- > Umwandlung in standardisiertes Datenformat nötig
- > generische Abfragen möglich
- > Daten werden als Webservice zur Verfügung gestellt
- > nicht auf biologische Daten festgelegt



Wrapper Software, Datenschemata

Wrapper Software

1. Biological Collection Access Service (**BioCASE**)
2. Distributed Generic Information Retrieval (DiGIR)
3. TDWG Access Protocol for Information Retrieval (TAPIR)



XML-Daten-Schemata

- Access to Biological Collection Data (ABCD) (1)
- Darwin Core (DwC) (2,3)

BioCASE Provider Software

- Python -> plattformunabhängig
- Arbeitet mit jeder relationalen SQL-fähigen Datenbank zusammen (Access, Filemaker, MySQL, Oracle, Postgres, SQL Server, 4th Dimension)
- Arbeitet mit ABCD, generisches Design erlaubt jedoch Unterstützung beliebiger xml-Schemata
- Mit BioCASE/ABCD werden ca. 29 Millionen Daten an GBIF geliefert
- Erweiterungen von ABCD möglich (GeoCASE, DNA-Bank-Netzwerk)

BioCASE Provider Software / ABCD

Hierarchie/Wiederholbarkeit

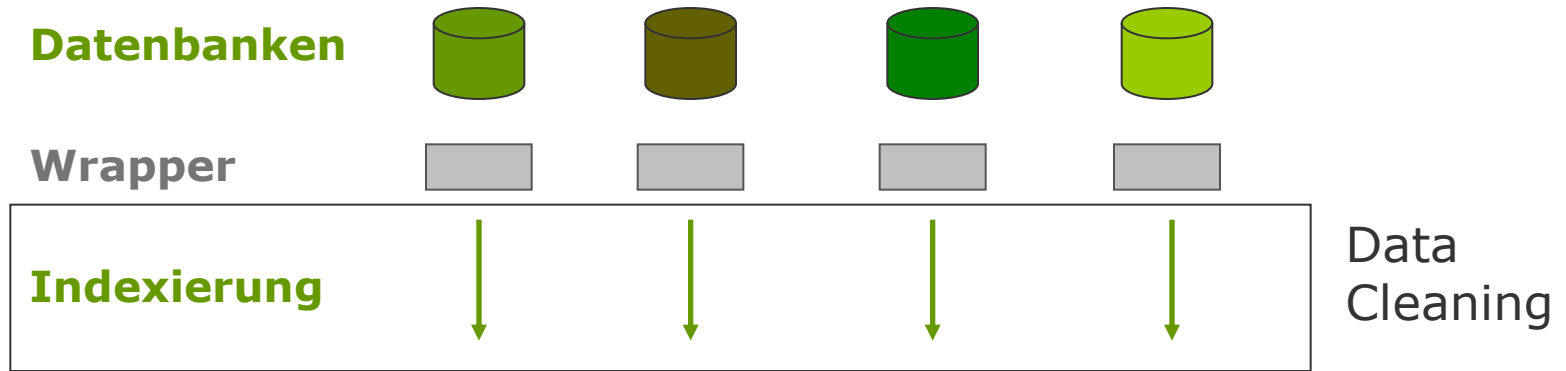
```

<Identifications>
  <Identification>
    <Result>
      <TaxonIdentified>
        <HigherTaxa>
          <HigherTaxon>
            <HigherTaxonName>Mesotaeniaceae</HigherTaxonName>
            <HigherTaxonRank>familia</HigherTaxonRank>
          </HigherTaxon>
          <HigherTaxon>
            <HigherTaxonName>Zygnematales</HigherTaxonName>
            <HigherTaxonRank>ordo</HigherTaxonRank>
          </HigherTaxon>
          <HigherTaxon>
            <HigherTaxonName>Conjugatophyceae</HigherTaxonName>
            <HigherTaxonRank>classis</HigherTaxonRank>
          </HigherTaxon>
        </HigherTaxa>
        <ScientificName>
          <FullScientificNameString>Ancylonema nordenskiöldii</FullScientificNameString>
        </ScientificName>
      </TaxonIdentified>
    </Result>
    <PreferredFlag>1</PreferredFlag>
    <Date>
      <DateText>1878-07-18</DateText>
    </Date>
  </Identification>
</Identifications>

```



GBIF - Aufbau



GBIF - Aufbau



... free and open access to biodiversity data
GLOBAL BIODIVERSITY INFORMATION FACILITY

search species/country/dataset

Search

HOME SPECIES COUNTRIES DATASETS OCCURRENCES SETTINGS ABOUT



Occurrence search

Please add filters and click Search to perform a search for occurrences records. Specify more filters to narrow your search to get more accurate results for the species, area or time period that is of interest.

Add search filter

Scientific name

Taxonomy

- Scientific name
- Common name
- Classification
- Type status

Geospatial

- Country
- Region
- Bounding box
- Latitude
- Longitude
- Altitude
- Depth
- Coordinate status
- Coordinate issues
- Protected Area

Datasets

- Data publisher
- Dataset name
- Host country

any records that have a matching name given for the identification regardless of how the organism is classified.

Your current search

Your current search is empty.

Example searches:

- Observations of *Strix aluco* in Poland
- Observations of *Puma concolor* in North America



GBIF - Aufbau

Actions

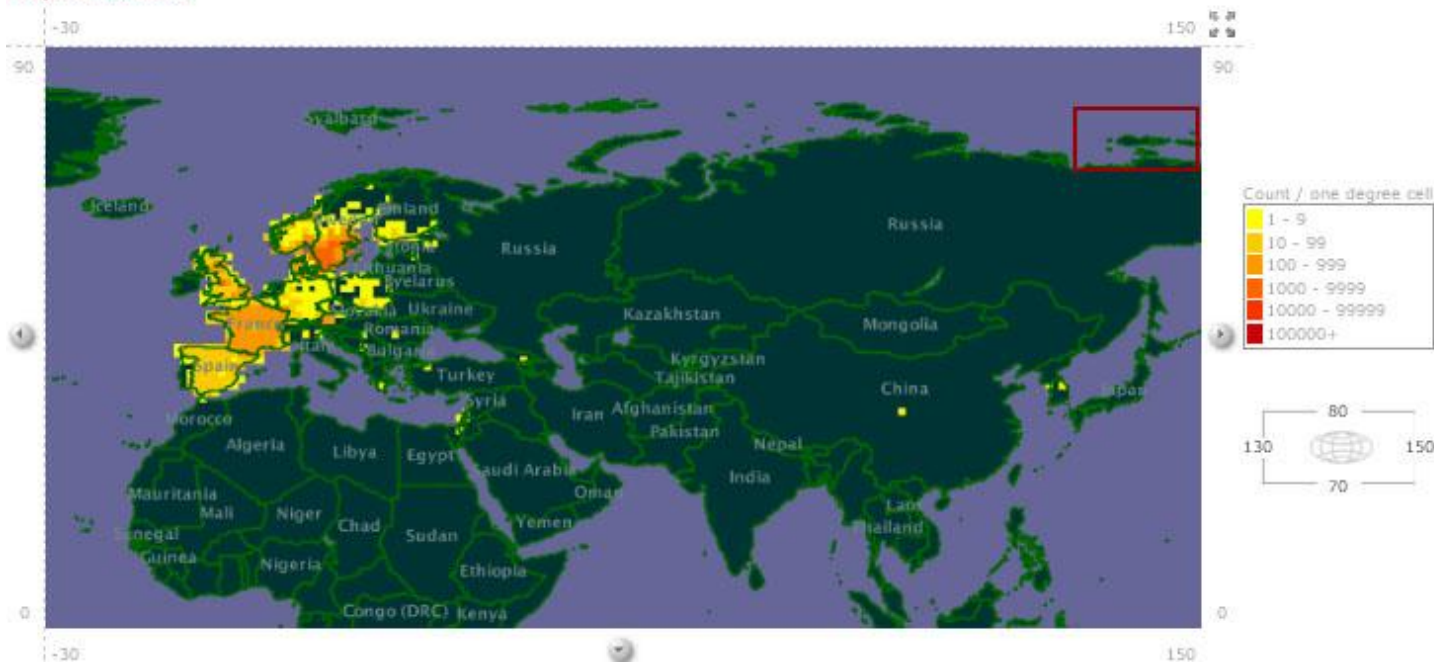
- View:** [Matching records as table](#) [Matching records on map](#)
- Specify:** [Data publishers to be included in search](#) [Datasets to be included in search](#) [Countries to be included in search](#)
- Download:** [Spreadsheet of results](#) [Darwin core \(maximum 100,000\)](#) [Google Earth \(maximum 50,000\)](#) [Species in results](#)
- Create:** [Niche Model](#)

Your current search

Classification includes Species: *Strix aluco*

[Change your current search](#)

Map of results



GBIF - Aufbau

RBG Flora exsiccata Bavarica 1801 - GBIF Portal - Mozilla Firefox

http://data.gbif.org/occurrences/273130/

Occurrence Details: RBG Flora exsiccata Bavarica 1801
Aconitum napellus

Actions

- Find:** [Records of *Aconitum napellus* from Flora exsiccata Bavarica](#)
- Retrieve:** [Retrieve original record from data provider](#)
- View:** [View information for *Aconitum napellus*](#) [Classification of *Aconitum napellus* according to Flora exsiccata Bavarica](#)
- Send:** [Feedback to Botanic Garden and Botanical Museum Berlin-Dahlem](#)

Dataset

Data Provider: [Botanic Garden and Botanical Museum Berlin-Dahlem](#)

Dataset: [Flora exsiccata Bavarica](#)

Institution code: RBG


Collection code: Flora exsiccata Bavarica

Catalogue No: 1801

Basis of record: Specimen (inferred)

Collector name: Kurt Harz

Date collected: Day:26 Month:8 Year:1917 (interpreted as 26.08.1917)

Image: 

Fertig

GBIF – Nutzung der Daten

Thematischer Fokus

Atlas of Living Australia (Flora/Fauna Australiens)

Iabin (Bestäuberarten Amerikas)

BioCASE (Europäische Daten)

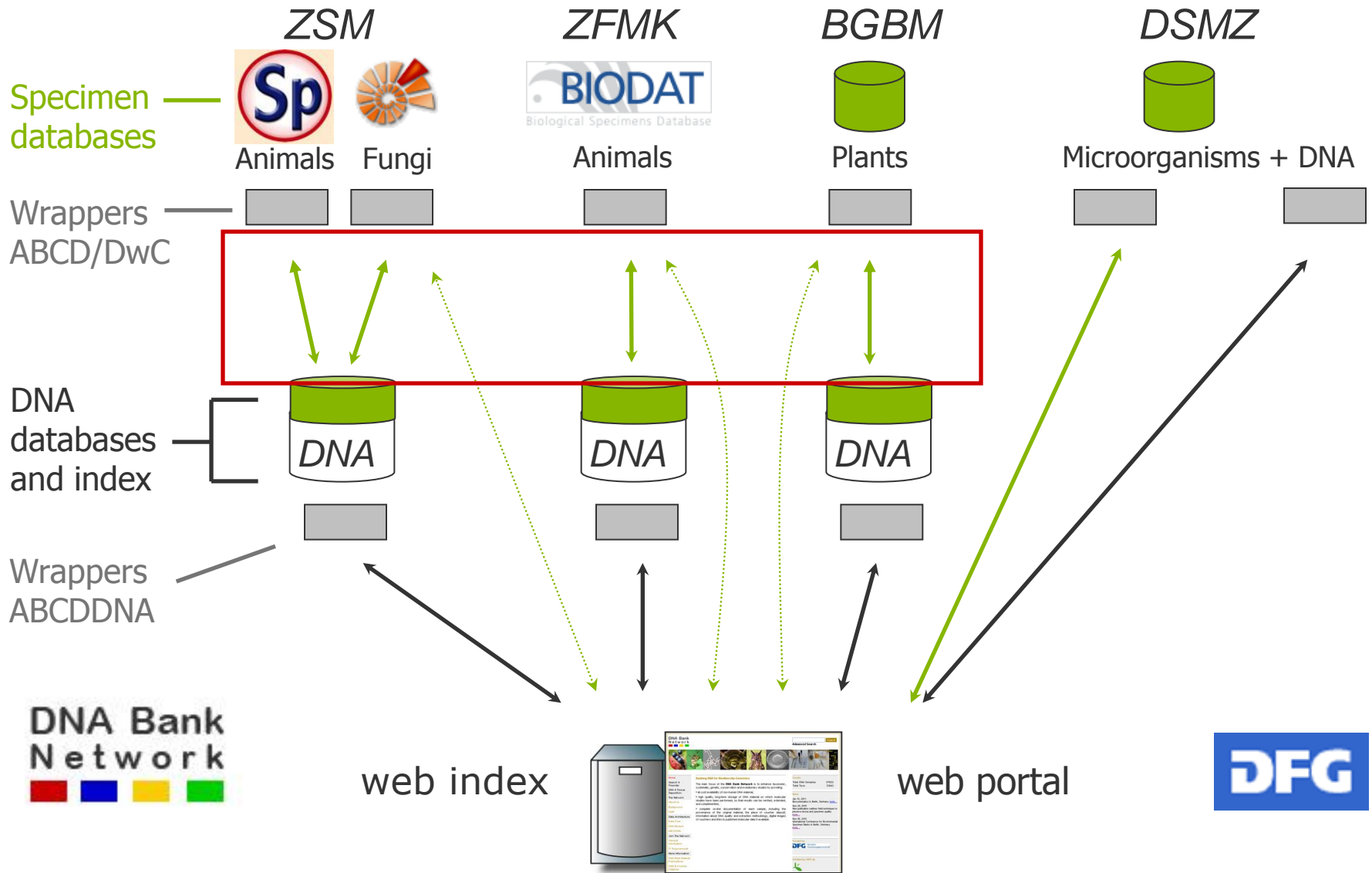
GeoCASE (Geowissenschaftl. Daten, ABCDEFG)

VertNet (Wirbeltiere)

DNA-Bank-Netzwerk (DNA-+ Beleg-Daten, ABCDDNA)



DNA-Bank-Netzwerk www.dnabank-network.org



DNA-Bank-Netzwerk

DNA Details:

Preferred Taxon Name: **Cichorium intybus L.**

Unit ID/DNA Bank No: DB 6
 Institution Code: BGBM
 Collection Code: DNA Bank
 Record Basis: DNA Sample
 Kind of Unit: gDNA

Relation to voucher:

DNA and specimen from the same population
 Tissue: Leaf Preservation: unknown

Extraction:

Extraction Date: 13.08.2002 Extraction Method: DNeasy Plant Mini Spin Kit Qiagen
 Extraction Staff: Banesmer, Jana

Quality:

Concentration: 17,14ng/µl
 Ratio of Absorbance: 1,38 OD_{260nm}/OD_{280nm} Ratio of Absorbance: 0,77 OD_{230nm}/OD_{260nm}
 Quality Check Date: 15.04.2008

Specimen details (voucher)


Unit ID: B 10 0066577
 Institution Code: BGBM
 Collection Code: Herbarium Berolinense
 Record Basis: PreservedSpecimen
 Kind of Unit: herbarium sheet

Identification(s):

Name: **Cichorium intybus L.**
 Taxonomy: COMPOSITAE (familia) ;

Multimedia:

1. [Show image \(high quality\)](#)
 Format: fpx

2. [Show image](#)

 Format: jpg

Gathering:

Locality	Collector(s)	Other info
Country: Spain (ES) Continent: Europe Locality: Bailén Prov. Jaen.	B. Valdés, S. Talarera, G. Alziar, D. Jeanmonod, N. Galland, U. Matthás, V. Stevanovic, P. Minissale, S. Fici, B. Foggi, M. Watson, P. Hinz & J. M. Romero	Biotope: In roadside ditches, basic soils Project: 1st Iter Mediterranean of OPTIMA; South-Eastern Spain, June-July 1988

Other data:

Collector(s) field number: Rec. It2993/68

DNA-Daten
 Provider Nr.1

Beleg-Daten
 Provider Nr.2

Vielen Dank für Ihre Aufmerksamkeit!



g.droege@bgbm.org