Module 3: Global databases, applications and the future of identification

This module connects organism identification to **global biodiversity infrastructure**, real-world applications, and emerging directions in biodiversity science. Learners explore how identification data flows into international databases, supports conservation and policy decisions, enables large-scale monitoring, and powers future tools such as phylogenetic visualization and open science platforms.

1.1 Why global biodiversity infrastructure matters

Global biodiversity infrastructures like GBIF, BOLD, and IUCN Red List enable standardized species names and classifications, integrating morphological, molecular, and ecological data from diverse sources for seamless comparison. They facilitate long-term monitoring of distribution shifts and extinction risks, powering evidence-based conservation policies and automated alerts for invasives. Without these shared frameworks, siloed identifications hinder scalability and global response to biodiversity loss.

1.2 Global taxonomic backbones

Catalogue of Life (CoL)

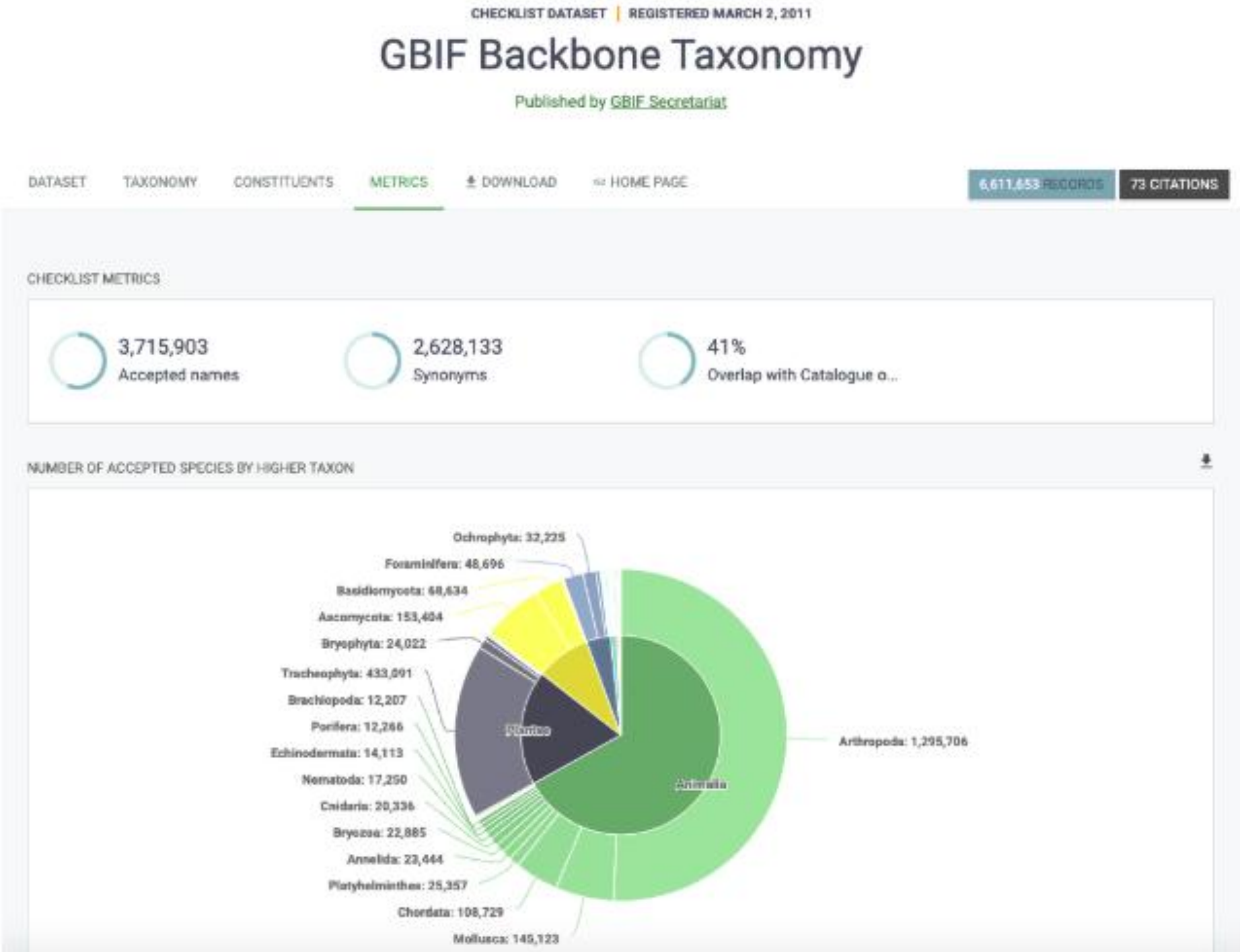
The Catalogue of Life (CoL) aims to compile a complete, authoritative list of all known species through global taxonomic collaboration. Key features include verified scientific names from expert databases, hierarchical classification spanning Domain to species, and links to specialist sources like WoRMS and ITIS for detailed taxonomy.

It serves as a reference for accepted names, reducing synonymy and inconsistencies by providing stable identifiers integrated with GBIF and BOLD. This backbone taxonomy enables standardized identification workflows across biodiversity platforms.(https://www.catalogueoflife.org/)

GBIF Species Backbone

The GBIF Species Backbone integrates taxonomic data from 150+ sources like Catalogue of Life into a unified canonical classification covering 2.2M+ species. Key functions include standardizing names across datasets via synonymy resolution, enabling interoperability through stable taxonKeys, and supporting global mapping of occurrences with higher ranks.

Its impact allows consistent comparison of billions of worldwide occurrence records, powering analyses of range shifts and invasion patterns while minimizing nomenclatural mismatches. (https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c)



CHECKLIST DATASET | REGISTERED MARCH 2, 2011

GBIF Backbone Taxonomy

Published by GBIF Secretariat

CHECKLIST METRICS

3,715,903 Accepted names

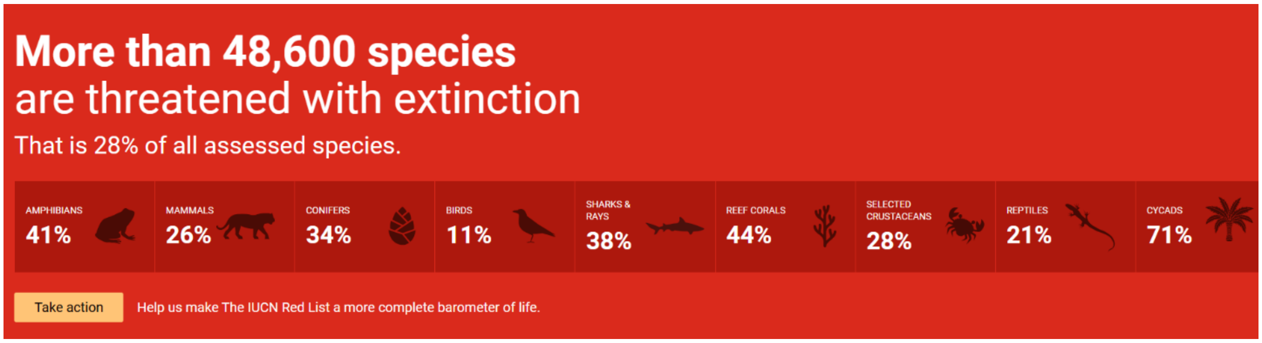
2,628,133 Synonyms

41% Overlap with Catalogue o...

NUMBER OF ACCEPTED SPECIES BY HIGHER TAXON

IUCN Red List

The IUCN Red List assesses global conservation status using categories from Least Concern (lowest risk) through Vulnerable, Endangered, Critically Endangered (extreme risk), to Extinct. Accurate identification is crucial because misidentification distorts risk assessments, leading to misguided priorities and invalid legal protections under conventions like CITES. (https://www.iucnredlist.org/)



More than 48,600 species
are threatened with extinction

That is 28% of all assessed species.

AMPHIBIANS	MAMMALS	CONIFERS	BIRDS	SHARKS & RAYS	REEF CORALS	SELECTED CRUSTACEANS	REPTILES	CYCADS
41%	26%	34%	11%	38%	44%	28%	21%	71%

Help us make The IUCN Red List a more complete barometer of life.

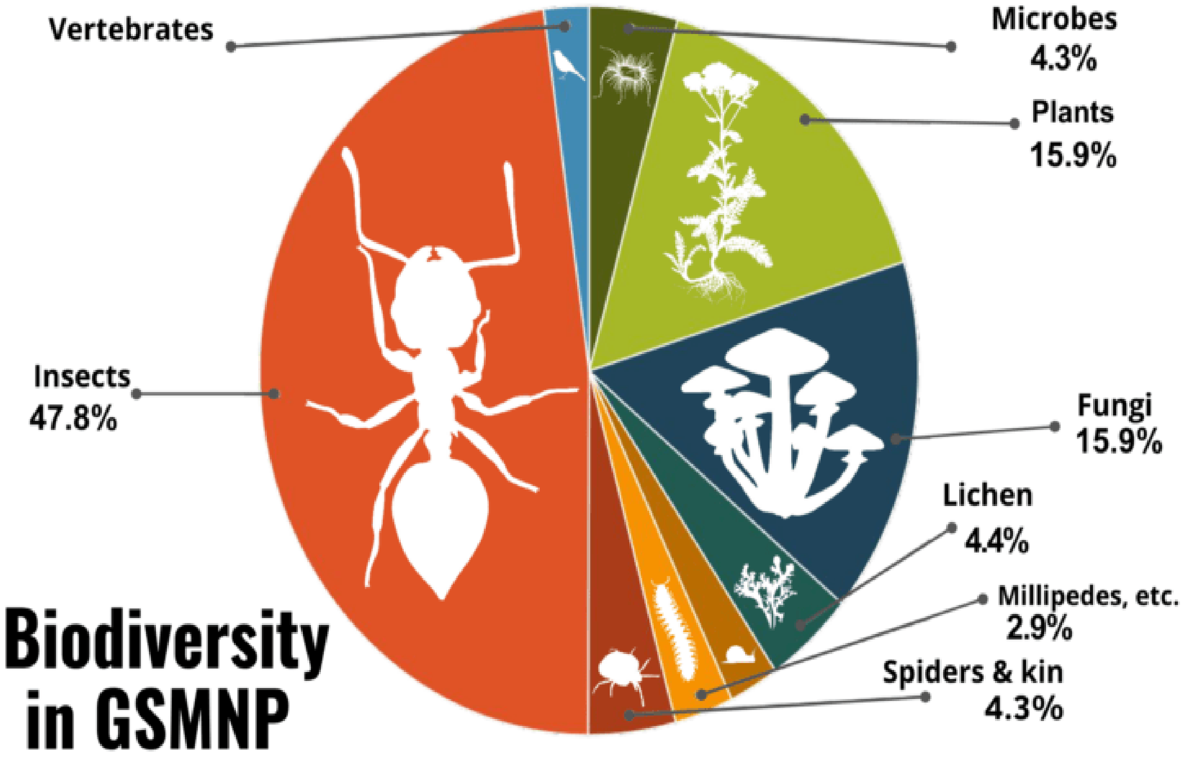
1.3 All-Taxa Biodiversity Inventories (ATBIs)

ATBIs (All Taxa Biodiversity Inventories) are comprehensive efforts to document all living species within defined geographic areas, typically biodiversity hotspots like national parks.

Great Smoky Mountains Case Study

This pioneering ATBI targets ~80,000 species across microorganisms, plants, fungi, invertebrates, and vertebrates, doubling known park diversity to 20,000+ species and describing 1,000+ new ones through parataxonomists and experts.

Most diversity resides in cryptic arthropods, fungi, and microbes; integrative taxonomy (morphology + DNA) proves essential; success demands sustained collaboration among institutions, specialists, and citizen scientists. (https://dlia.org/about/atbi/)



1.4 Species spotlight

Microorganisms are identified primarily through molecular tools like 16S rRNA (bacteria), ITS (fungi), and MALDI-TOF MS protein profiling due to minimal morphology, powering ecosystem functions, health diagnostics, and biotechnology.

Plants rely on morphological keys (leaves, flowers) bolstered by rbcL/matK barcoding, linking identification to habitat conservation via herbarium networks.

Fungi exhibits high cryptic diversity resolved by ITS barcoding, as morphology fails; they regulate decomposition and symbioses essential to ecosystems.

Invertebrates (~97% animal diversity) integrate morphology, COI barcoding, and AI imaging for rapid processing in biodiversity surveys.

Vertebrates combine morphology, behavior, geography, and genetics due to high policy stakes like CITES listings.

Algae/Diatoms, key aquatic producers, use microscopy (frustule shape) and molecular markers for water quality and paleoclimate reconstruction.

1.5 Applications of species identification

Conservation Biology leverages precise IDs for species risk assessment (IUCN Red List), habitat protection planning, and population trend monitoring via mark-recapture or camera traps.

Ecology and Monitoring computes biodiversity indices (Shannon, Simpson), tracks ecosystem health with indicator species, and deploys eDNA metabarcoding for non-invasive community surveys.

Bioinformatics fuses taxonomic, genomic (barcodes), and occurrence data (GBIF) for large-scale modeling of distributions and phylogenies.

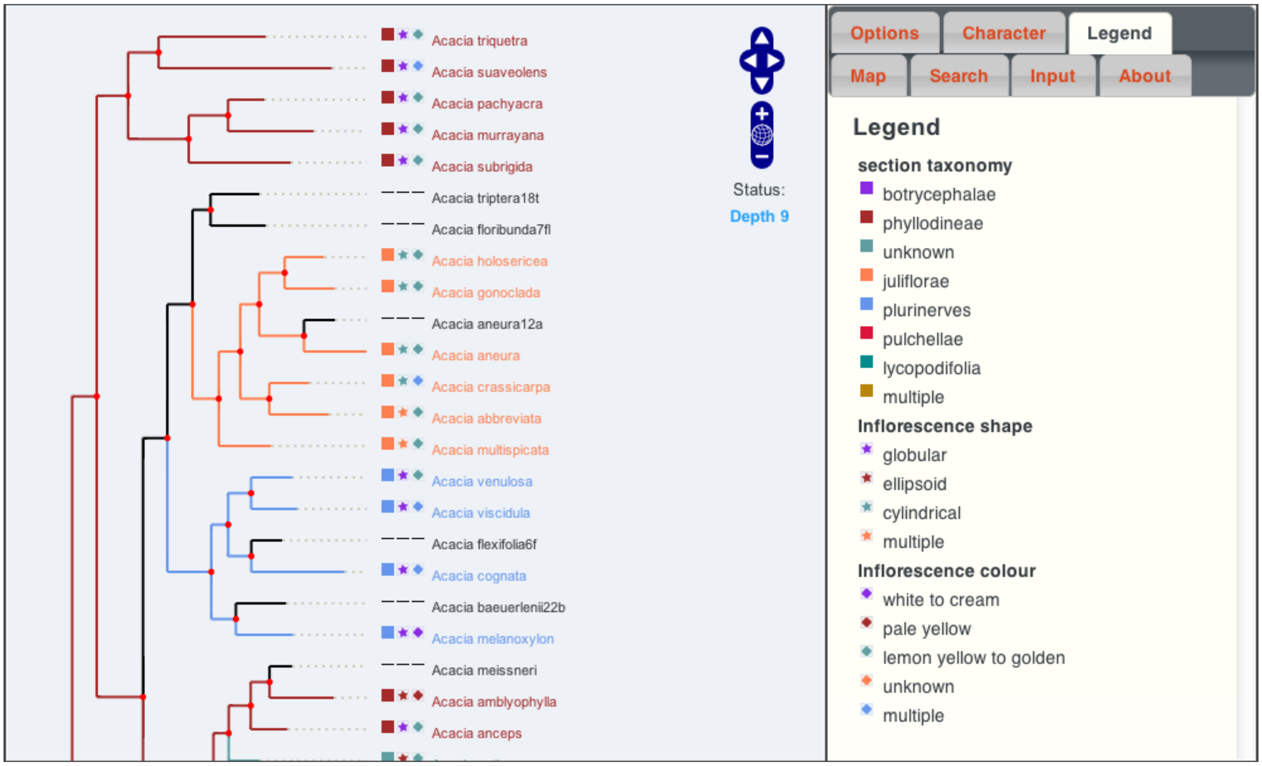
Citizen Science harnesses platforms like iNaturalist for image-verified records, broadening coverage with AI-assisted IDs.

Policy and Regulation inform environmental impact assessments, invasive management, and treaties like CBD or CITES through verified data.

1.6 Phylogenetic visualization

Phylogenetic trees provide an evolutionary framework for understanding biodiversity.

PhyloJIVE (Phylogeny JavaScript Information Visualiser and Explorer)



PhyloJIVE is an open-source tool that:

- Visualizes evolutionary relationships interactively
- Links species names to occurrence data, traits, and maps
- Integrates data from Catalogue of Life, GBIF, and regional biodiversity services

Phylogenetic visualization helps users move from species lists to evolutionary understanding. (https://www.ala.org.au/blogs-news/phylojive-integrating-biodiversity-data-with-phylogenies/)