What you probably didn’t know about Biodiversity Information Serving Our Nation (BISON) and the Integrated Taxonomic Information System (ITIS)

Dr. Stinger Guala
USGS Core Science Systems, Science Analytics and Synthesis
9 Oct. 2019
US National Clearinghouse for biological occurrence data
• 464+ million records & growing
  • 16M are non-native occurrences
  • 100M available nowhere else
• US Node Application of GBIF
• Nearly all species in US
  • Including 10K+ non-natives
• Taxonomic standardization
• Every state and county
• Who, what, when, where for every record (at a minimum)
• Saves repeated re-investment in data integration.

“I have been spending a lot of weekends using BISON and the data have been extremely helpful in my analyses. I am very grateful to have this resource - thank you for all you’ve done to make this what it is.”
September 14, 2016. Laura T. Bortolin, Harvard Medical School, Dept. of Genetics, Boston, MA

“...I am very impressed with BISON and am happy the USGS is making biodiversity data available as it is desperately needed. Keep up the good work!”

“...we're really excited at its performance in producing and mapping extremely large search-result sets (I generated one with 1.8 million hits in a matter of seconds).”
April 22, 2013. Scott L. Cross, Ph.D.. NOAA National Oceanographic Data Center/National Coastal Data Development Center, Charleston, SC.

“...thanks a million for this application. It is really adding to the number of records I have been able to acquire for numerous amphibian and reptile species.”
April 29, 2013. Bill Sutton, Ph.D., Postdoctoral Research Associate, Clemson University, School of Agricultural, Forest and Environmental Sciences.
What BISON does

- It provides the only fully integrated and comprehensive clearinghouse for US species occurrence data with extensive Web Services to power other websites, scientific analyses and applications.

- It delivers Federal data sets, increasing usage, visibility and consistency.
  - Birds and Plants are of special emphasis in Federal data.

- Identification and mobilization of targeted data sets (Federal and Non-Federal) to produce comprehensive coverages.
  - Emphasis groups: invasives, birds & insects.

- US relevant taxonomic and geospatial binning (e.g. ITIS, Counties, States).
  Automatic checklists for anywhere in the US.

- Community engagement, training and support.

More than a million different people have collected the data points now in BISON
Custom Maps & Checklists

For Any Area Including User Polygons
Documentation with a Click & Download
Refine Search and Environmental Layers
It’s in there…

1000’s Global Providers

Natureserve

Smithsonian

VertNet

eBird

iNaturalist

Federal Data Sets (e.g. FIA,NPS,BLM)

Non-Gov Data for BISON analyses (e.g. VegBank)

USDA PLANTS

GBIF

BISON

More than a million different people have collected the data points now in BISON
Obtaining and Cleaning Species Occurrence Data for Models

1. Acquire occurrence data from publicly available databases, including: GBIF (www.gbif.org), eBird (www.ebird.org), BISON (bison.usgs.gov), iNaturalist (www.inaturalist.org), Berkeley Ecoinformatics (ecoengine.berkeley.edu), VertNet (vertnet.org), iDigBio (www.idigbio.org), OBIS (obis.org), as well as other appropriate taxon specific databases. These include vouchered museum records, as well as records from verified citizen science databases, such as iNaturalist and eBird.

2. Data will be acquired using the valid taxonomic name, as well as any species synonyms according to the Integrated Taxonomic Information System (ITIS) database https://www.itis.gov/. In some cases, additional data for a species may be available, but listed under another valid and related species. This can occur when one species is split into two or more species, based on newly published taxonomic guidance. In many cases, data for the previous species will not have been transferred to the newly recognized species. Therefore, it
Power in Data from Multiple Sources
Species coverage in checklists of National Parks
  • Checklists from data in NPSpecies vs. BISON

Undergraduate Capstone Project by Theo Burton, USGS Patuxent Wildlife Research Center
Patuxent Wildlife Research Center Bird Banding Lab and Other Bird Data
Invasion of Rare Native Pecos Pupfish by Invasive Sheepshead Minnow

Where they come into contact, they hybridize, eradicating the pupfish.

An example of the power of having integrated T&E and Invasives data.

At least forty percent of T&E species are pressured by invasives.
Spread of *Pueraria montana* var. *lobata* (Kudzu) in the US, by displaying BISON data cumulatively by year.
98% of records in BISON are covered by ITIS
Integrated Taxonomic Information System

• Authoritative standard for names of biological organisms in Federal agencies
• Federal Partnership: 9 agencies
• Complete hierarchy of life and extensive synonymy allows data integration across biology
• Taxonomic Serial Numbers serve as the only truly reliable standard for Globally Unique Identifiers on species names
• Foundation for most other broad taxonomic databases
• International partnership produces the Species2000 & ITIS Catalogue of Life every year.
ITIS Taxonomic Serial Numbers are Social Security Numbers for Species Names
ITIS is the Rosetta Stone

ITIS web services and imbedded tables are heavily used in thousands of applications

15 Articles

Marginaria polypodioides

814 Articles

With ITIS synonomy

Pleopeltis polypodioides
Polypodium polypodioides

USGS
The Importance of Species Name Synonyms in Literature Searches

Gerald F. Guala*
- 39 fields for every Taxonomic Serial Number
- Multiple pre-computed hierarchy fields for high speed rendering on websites and advanced calculations
- JSONP – so you don’t need a server to run code from it in your HTML
- Facets – and pivot facets
- Highly configurable fuzzy matching
- Motif and association support for data mining
- Multiple output formats
- No limits - way faster…

https://services.itis.gov/?q=
https://www.itis.gov/solr_documentation.html
https://www.itis.gov/solr_examples.html
Regular Expressions.

- matches only monomials
  - matches only binomials
  - matches only trinomials

Use the ITISscientificName in BISON – it will give you the records under synonyms automatically.

Use the Hierarchy Strings in ITIS and BISON to search on higher level taxa and generate hierarchies.

Get Fuzzy
Try a Polygon and a facet or two:

https://bison.usgs.gov/solr/occurrences/select/?q=ITISscientificName:[A-Za-z]*[%20]%7B1,1%7D[A-Za-z]*/&fq=geo:%22IsWithin(POLYGON%20((-85.635500451549888%2037.611442854627967,%20-85.636395091190934%2037.612390553578734,%20-85.636666836217046%2037.61264057084918,%20-85.6467796266079%2037.617916146293283,%20-85.647250449284911%2037.61264057084918,%20-85.652201883494854%2037.618336495012045,%20-85.740281673148274%2037.530058404430747,%20-85.738804267719388%2037.529313048347831,%20-85.729063972830772%2037.529499590396881,%20-85.639235144481063%2037.605515029281378,%20-85.635500451549888%2037.611442854627967))%20distErrPct=0.1%22&facet=true&facet.pivot=ITISscientificName,resourceID,ITIScommonName&facet.limit=-1&facet.mincount=1&rows=0&wt=json

Generates a species (only species) list with common names and provider for a small National Park property.

And try a CSV feed if the WMS is challenging:

https://bison.usgs.gov/solr/occurrences/select/?q=ITISscientificName:"Poa annua"&fl=decimalLatitude,decimalLongitude&wt=csv
Use the geo filter for mobile apps...

https://bison.usgs.gov/solr/occurrences/select/?q=*&fq={!geofilt pt=" + latitude + "," + longitude + " sfield=geo\%20d="+d+"}&facet=true&facet.field=ITISscientificName
TIME
for
QUESTIONS
( Thank you )