Bugs and BISON:
130+ Million Mapped Species Records Now Includes ITIS-Enabled Search

BISON: A national unified resource for discovery, linkage and reuse of species occurrence data.

http://bison.usgs.ornl.gov

Annie Simpson
U.S. Geological Survey; Core Science Analytics, Synthesis, & Libraries (CSASL) Program
Core Science Systems (CSS) Mission Area

16-June-2014

U.S. Department of the Interior
U.S. Geological Survey

The BISON system *(a review)*

- **Developed by the U.S. Geological Survey**
  - **Core Science Analytics, Synthesis, & Libraries (CSASL) Program**

- **A species occurrence data aggregator providing 137+M species occurrence records for the U.S. and Territories including:**
  - Almost every US species (animals, plants, fungi)
  - Various data types: observation-based, natural history collections (specimen-based), and literature-based
  - Federal and non-federal data

BISON Goals

Increase Data Access, Exposure, Discoverability, and Quality

Data Mobilization through integration and application of standards, open data technologies, machine readable access

Data Layers

Combine specimen and observation data with other data layers to detect relationships and patterns.

What's new in BISON?

- **Search is now more than "exact match" of a name, so taxon groups can be searched, including synonyms**
- **Uses powerful name resolution ability of the Integrated Taxonomic Information System**

USGS
science for a changing world

Search example – *Halyomorpha halys*

Only 13 matches currently

Found **13** matches for **Halyomorpha halys** using ITIS taxonomy
13 georeferenced

Record Details generated with ITIS taxonomy.

Resource: KNWR Entomology Collection
Basis of Record: specimen
Scientific Name: *Halyomorpha halys*
Common Names: brown marmorated stink bug
State: Alaska
County: Kenai Peninsula Borough
Provider: Arctos
Collector: unknown
Latitude: 60.31663
Longitude: -151.26102
Precision: N/A
Year: 2012
Occurrence Date: 2012-11-25

For more information, or to make comments on this occurrence, please contact the provider.

≋ USGS
science for a changing world

Data Fields

- **Darwin Core standard fields desirable**
- **Minimum fields necessary for new datasets to be successfully added into BISON**

	Field
Who -	provider - resource
What -	scientific_name
Where -	decimal_latitude / decimal_longitude (or a place name e.g. state, county…)
When -	occurrence_date

Additional data fields are welcome/encouraged

USGS
science for a changing world

Search example – Halyinae, Pentatominae

Halyini = 701 records
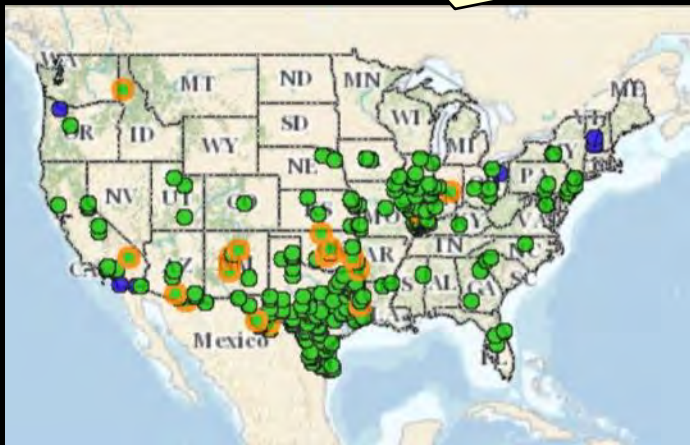
Pentatominae = 12,342 records

ITIS search for BMSB hosts in BISON

Prunus persica
common cherry (2,524)

Prunus spp.
stone fruit (34,642)

Ailanthus altissima
tree of heaven (2,443)

Ailanthus spp.
(2,847)

What else is new in BISON?

- **40+ data fields now supported, including several verbatim from provider**
- **Support for voucher images, video, audio files (later this summer)**

USGS
science for a changing world

BISON fields (43)

BISON FIELD NAME	BISON FIELD NAME
valid_accepted_scientific_name	calculated_county_name
valid_accepted_tsn	calculated_fips
itis_common_name	calculated_state_name
basis_of_record	centroid
occurrence_date	provided_county_name
year	provided_fips
provider	provided_state_name
provider_url	thumb_url
resource	associated_media
resource_url	associated_references
occurrence_url	general_comments
catalog_number	id
collector	provider_id
collector_number	resource_id
clean_provided_scientific_name	provided_common_name
itis_tsn	provided_kingdom
provided_scientific_name	~~calculated_tsn~~
provided_tsn	geodetic_datum
latitude	coordinate_precision
longitude	coordinate_uncertainty
verbatim_elevation	verbatim_locality
verbatim_depth	iso_country_code

USGS
science for a changing world

Darwin Core Fields (148)

occurrenceID	associatedSequences	locationAccordingTo	identificationID
type	associatedTaxa	locationRemarks	identifiedBy
modified	eventID	verbatimCoordinates	dateIdentified
language	samplingProtocol	verbatimLatitude	identificationReferences
rights	samplingEffort	verbatimLongitude	identificationRemarks
rightsHolder	eventDate	verbatimCoordinateSystem	identificationQualifier
accessRights	eventTime	verbatimSRS	typeStatus
bibliographicCitation	startDayOfYear	decimalLatitude	taxonID
institutionID	endDayOfYear	decimalLongitude	scientificNameID
collectionID	year	geodeticDatum	acceptedNameUsageID
datasetID	month	coordinateUncertaintyInMeters	parentNameUsageID
institutionCode	day	coordinatePrecision	originalNameUsageID
collectionCode	verbatimEventDate	pointRadiusSpatialFit	nameAccordingToID
datasetName	habitat	footprintWKT	namePublishedInID
ownerInstitutionCode	fieldNumber	footprintSRS	taxonConceptID
basisOfRecord	fieldNotes	footprintSpatialFit	scientificName
informationWithheld	eventRemarks	georeferencedBy	acceptedNameUsage
dataGeneralizations	locationID	georeferenceProtocol	parentNameUsage
dynamicProperties	higherGeographyID	georeferenceSources	originalNameUsage
catalogNumber	higherGeography	georeferenceVerificationStatus	nameAccordingTo
occurrenceDetails	continent	georeferenceRemarks	namePublishedIn
occurrenceRemarks	waterBody	geologicalContextID	higherClassification
recordNumber	islandGroup	earliestEonOrLowestEonothem	kingdom
recordedBy	island	latestEonOrHighestEonothem	phylum
individualID	country	earliestEraOrLowestErathem	class
individualCount	countryCode	latestEraOrHighestErathem	order
sex	stateProvince	earliestPeriodOrLowestSystem	family
lifeStage	county	latestPeriodOrHighestSystem	genus
reproductiveCondition	municipality	earliestEpochOrLowestSeries	subgenus
behavior	locality	latestEpochOrHighestSeries	specificEpithet
establishmentMeans	verbatimLocality	earliestAgeOrLowestStage	infraspecificEpithet
occurrenceStatus	verbatimElevation	latestAgeOrHighestStage	taxonRank
preparations	minimumElevationInMeters	lowestBiostratigraphicZone	verbatimTaxonRank
disposition	maximumElevationInMeters	highestBiostratigraphicZone	scientificNameAuthorship
otherCatalogNumbers	verbatimDepth	lithostratigraphicTerms	vernacularName
previousIdentifications	minimumDepthInMeters	group	nomenclaturalCode
associatedMedia	maximumDepthInMeters	formation	taxonomicStatus
associatedReferences	minimumDistanceAboveSurfaceInMeters	member	nomenclaturalStatus
associatedOccurrences	maximumDistanceAboveSurfaceInMeters	bed	taxonRemarks

BISON Associated Media

USGS GeoPlatform

- **A National and Federal focal point where governmental, non-governmental, private, and public data and applications can be visualized together to inform and address national and regional issues**

GeoPlatform.gov Conceptual Model
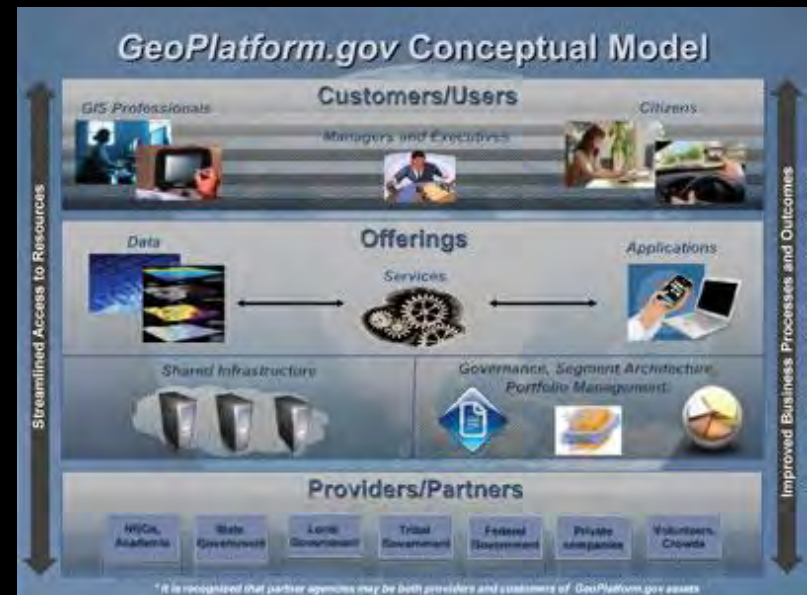
BISON Data Workflow Products

- **Final Package – original dataset + BISON (enhanced) dataset + metadata record (linked and archived) + ReadMe file (record of BISON data modifications)**

- **Standardized data – Darwin Core format, Scientific Name mapping to ITIS, FIPS Code location references**

- **Data updates – ongoing for *living* datasets**

- **Data Quality – improvement reports/recommendations sent to Data Provider**

- **Web and Web Services access – inc. mapping and visualization, and integration with other data layers**

- **Multi-format data download – .csv, .kml, .zip**

- **Machine access – via API**

USGS
science for a changing world

BISON & BMSB-NEIPM

How can BISON help members of the BMSB-NEIPM working group?

1. **"Submit-a-dataset"**
2. **Smarter search including synonyms and children; search for host species and species groups**
3. **Post BISON API search results on your Web pages as canned search results**

   http://bisonapi.usgs.ornl.gov/index.html

4. **Use BISON data or visualizations in your publications**

1. Submit a dataset to BISON

[this is you] → QA/QC → (US) **BISON** — All Data → (Global) **GBIF**

US Data → **Consumer Access via Web Services & Web** http://bison.usgs.ornl.gov

Machine Access BISON API

Consumer Access http://DataProvider.com

All Data → Consumer Access http://www.gbif.org

USGS
science for a changing world
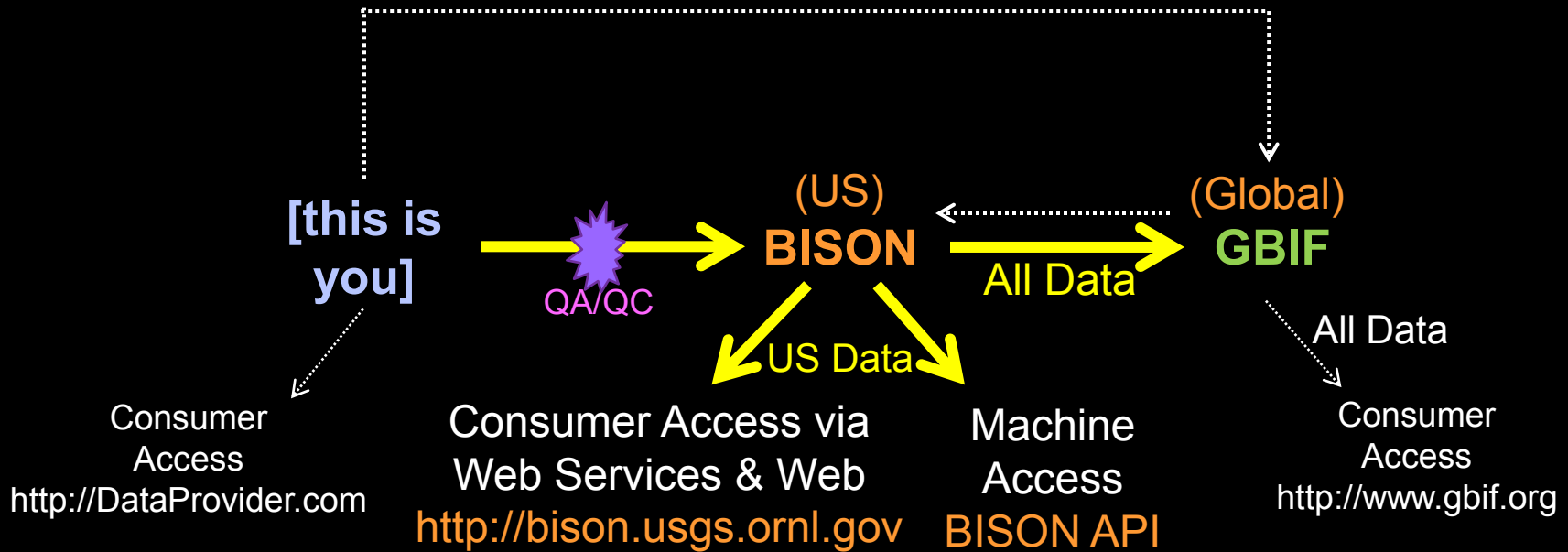
Future Attractions

- **Link to images of specimens and other media**

- **User interface redesign**
  (first phase coming this summer)

- **Addition of non-native status tags to species**

- **Automated species lists by protected area (e.g. refuge or park)**

![USGS logo]
USGS
science for a changing world

Data Retrieval and Download

Biodiversity Information Serving Our Nation (BISON):

http://bison.usgs.ornl.gov

For questions, additional information:
bison@usgs.gov

BISON Team (…@usgs.gov)

- **Gerald (Stinger) Guala [Director] - gguala**
- **Derek Masaki - dmasaki**

- **Elizabeth Martin – elizabeth_martin**
- **Elizabeth Sellers – esellers**
- **Annie Simpson - asimpson**

USGS
science for a changing world