Introduction to Data Standards

Nelson Rios
Division of Informatics

YALE
PEABODY
MUSEUM

Standards

Documented agreement containing specifications, guidelines, or characteristics that ensure products, systems, or services work reliably and can operate together.

Ethernet, TCP/IP, HTTP etc.

Developed by recognized organizations (such as the through a consensus-based process, which involves input from industry experts, governments, and other stakeholders

IEEE, ISO, or W3C etc.

Data Standards

Semantic Structure: Data standards provide agreed-upon definitions and formats for data elements. This ensures that, regardless of where the data originates, it follows the same rules for structure and meaning.

Interoperability: By using consistent data definitions, different systems can exchange, compare, and analyze data seamlessly.

Data Quality: Data standards help maintain data quality by setting rules for what valid data looks like, including definitions, units of measurement, or code sets.

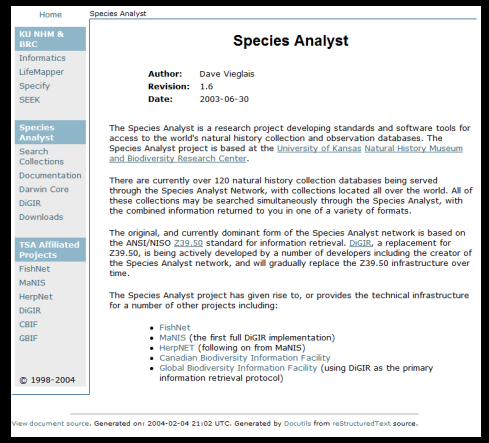
Evolution of Biodiversity Standards

1993

Fish Gopher

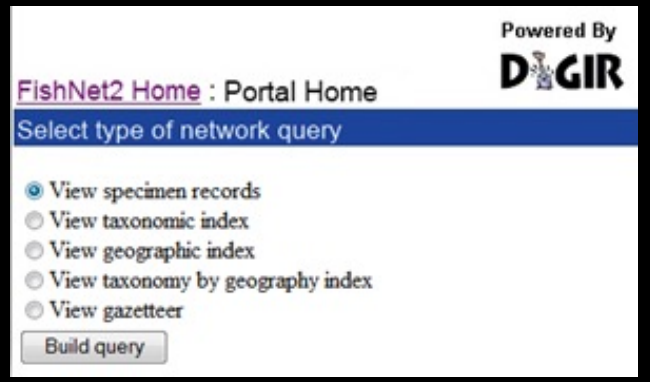
1998

ZBIG (Z39.50)
Darwin Core 1.0



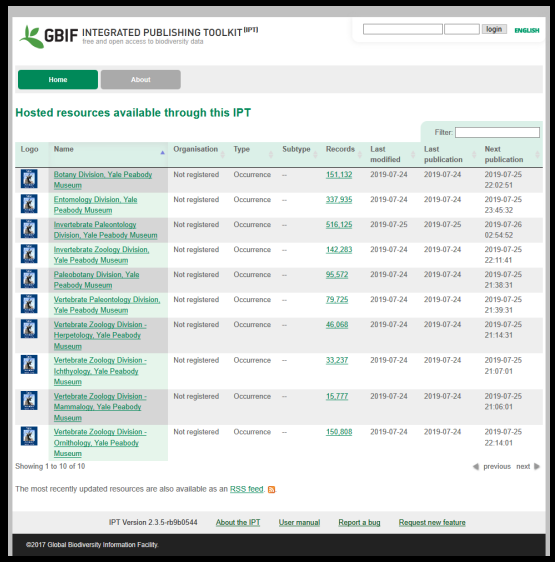
2001

DiGIR
Darwin Core 1.2



2009

IPT
Darwin Core Standard



Darwin Core

Darwin Core is a standard maintained by the Darwin Core Maintenance Interest Group. It includes a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to **facilitate the sharing of information about biological diversity** by providing identifiers, labels, and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

Occurrence

occurrenceID | catalogNumber | recordNumber | recordedBy | recordedByID
individualCount | organismQuantity | organismQuantityType | sex | lifeStage
reproductiveCondition | caste | behavior | vitality | establishmentMeans
degreeOfEstablishment | pathway | georeferenceVerificationStatus
occurrenceStatus | associatedMedia | associatedOccurrences
associatedReferences | associatedTaxa | otherCatalogNumbers
occurrenceRemarks

Occurrence	Class
Identifier	http://rs.tdwg.org/dwc/terms/Occurrence
Definition	An existence of a dwc:Organism at a particular place at a particular time.
Comments	
Examples	`a wolf pack on the shore of Kluane Lake in 1988` `a virus in a plant leaf in the New York Botanical Garden at 15:29 on 2014-10-23`

Event

eventID | parentEventID | eventType | fieldNumber | eventDate | eventTime
startDayOfYear | endDayOfYear | year | month | day | verbatimEventDate
habitat | samplingProtocol | sampleSizeValue | sampleSizeUnit
samplingEffort | fieldNotes | eventRemarks

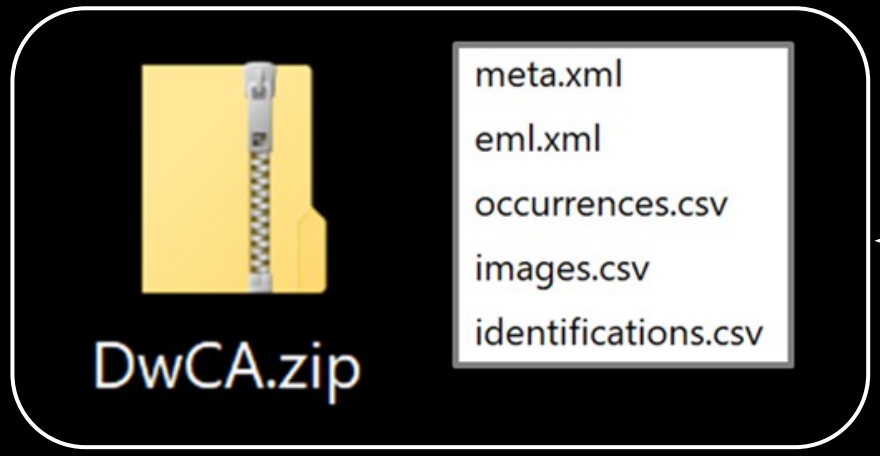
Event	Class
Identifier	http://rs.tdwg.org/dwc/terms/Event
Definition	An action that occurs at some location during some time.
Comments	

Examples	`a specimen collecting event`
	`a camera trap image capture`
	`a marine trawl`

vitality	
Identifier	http://rs.tdwg.org/dwc/terms/vitality
Definition	An indication of whether a dwc:Organism was alive or dead at the time of collection or observation.
Comments	Recommended best practice is to use a controlled vocabulary. Intended to be used with records having a dwc:basisOfRecord of `PreservedSpecimen`, `MaterialEntity`, `MaterialSample`, or `HumanObservation`. This term has an equivalent in the dwciri: namespace that allows only an IRI as a value, whereas this term allows for any string literal value.

Examples	`alive`
	`dead`
	`mixedLot`
	`uncertain`
	`notAssessed`

Eastern Seaboard
Mobilizing Millions of Marine Mollusks



Home

Hosted resources available through this IPT

10 resource(s) currently available

Filter:

Logo	Name	Organization	Type	Subtype	Records	Last modified	Last publication	Next publication
	Botany Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	186,529	2025-01-07 13:00:07	2025-01-07 13:00:59	2025-01-08 13:00:00
	Entomology Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	413,821	2025-01-07 13:05:09	2025-01-07 13:06:50	2025-01-08 13:05:00
	Invertebrate Paleontology Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	557,942	2025-01-07 13:10:00	2025-01-07 13:12:21	2025-01-08 13:10:00
	Invertebrate Zoology Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	149,090	2025-01-07 13:15:01	2025-01-07 13:15:42	2025-01-08 13:15:00
	Paleobotany Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	98,035	2025-01-07 13:20:02	2025-01-07 13:20:33	2025-01-08 13:20:00
	Vertebrate Paleontology Division, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	79,770	2025-01-07 13:25:03	2025-01-07 13:25:24	2025-01-08 13:25:00
	Vertebrate Zoology Division - Herpetology, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	75,760	2025-01-07 13:30:04	2025-01-07 13:30:31	2025-01-08 13:30:00
	Vertebrate Zoology Division - Ichthyology, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	77,592	2025-01-07 13:35:01	2025-01-07 13:35:22	2025-01-08 13:35:00
	Vertebrate Zoology Division - Mammalogy, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	18,866	2025-01-07 13:40:02	2025-01-07 13:40:13	2025-01-08 13:40:00
	Vertebrate Zoology Division - Ornithology, Yale Peabody Museum	Yale University Peabody Museum	Occurrence	–	163,560	2025-01-07 13:45:03	2025-01-07 13:45:44	2025-01-08 13:45:00

Showing 1 to 10 of 10

previous 1 next

The most recently updated resources are also available as an RSS feed.

Search Records

Help Reset

search all fields

Must have media Must have map point

Filters | Mapping | Sorting | Download

Add a field Clear

Scientific Name dwc:scientificName

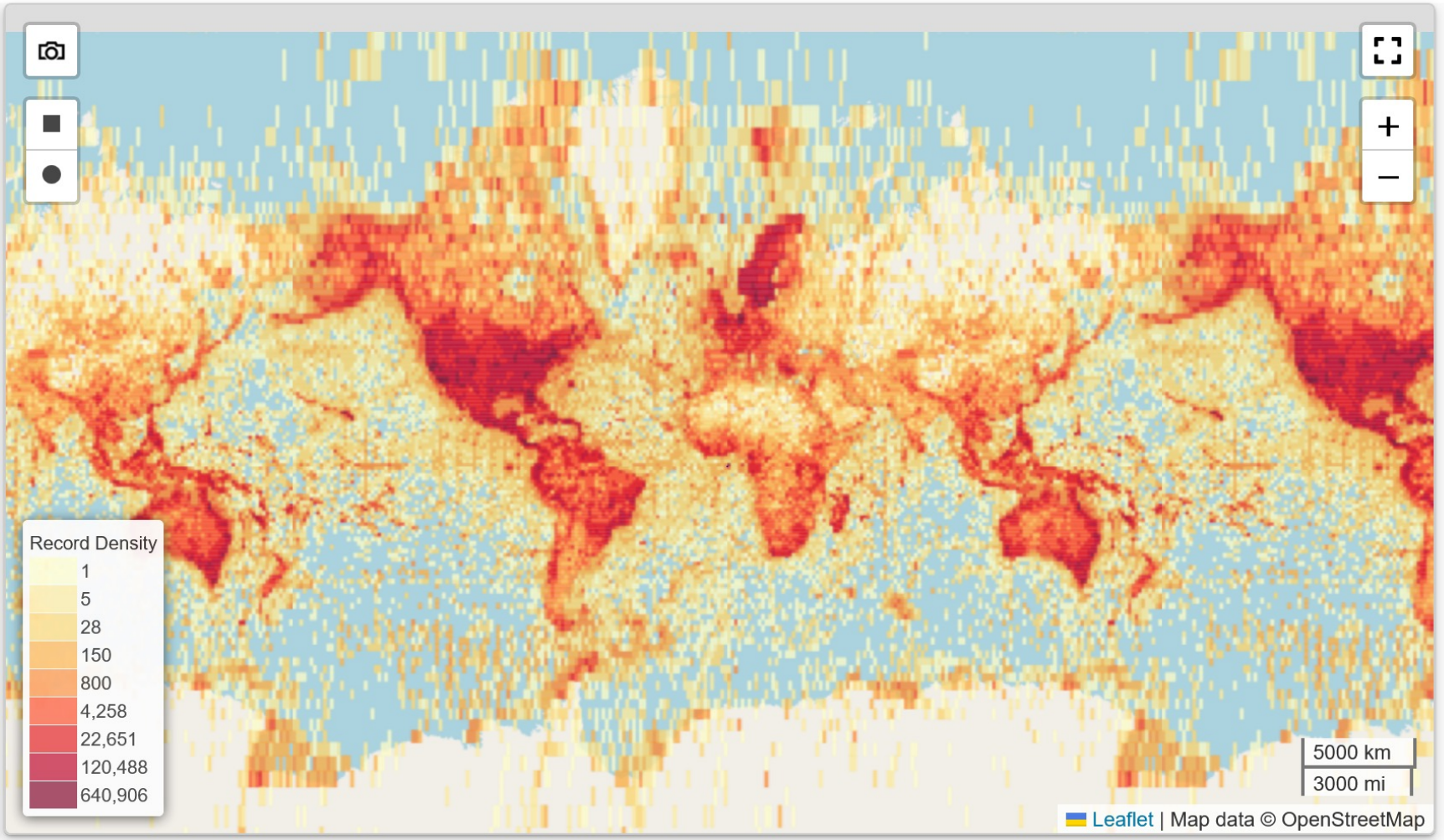
Present Missing Fuzzy

Date Collected Start: yyyy-mm-dd End: yyyy-mm-dd

Present Missing

Country dwc:country

Present Missing



Thematic Collections Networks

- InvertNet
- Tri-trophic (TTD)
- Lichens & Bryophytes (LBCC)
- New England Vascular Plants (NEVP)
- PaleoNICHES
- MacroFungi Collection Consortium (MaCC)
- Symbiota Collections of Arthropods (SCAN)
- Fossil Insect Collaborative (FIC)
- Vouchered Animal Communication Signals (VACS)
- Great Lakes Invasives
- InvertEBase
- The Key to the Cabinets
- Microfungi Collections Consortium (MiCC)
- Fossil Marine Invertebrates EPICC
- Cretaceous World
- LepNet
- The Mid-Atlantic Megalopolis MAM
- oVert
- SoRo
- Endless Forms
- The Pteridological Collections Consortium
- Capturing California's Flowers
- Pacific Island Land Snails
- Vascular Flora of the South-Central United States
- Parasite-Host Associations
- DigIn
- ESB
- Bryophytes and Lichens
- Tropical African Plants (TAP)

Thank You