



Improving Named Entity Recognition in BHL with Machine Learning

Katie Mika and Alicia Esquivel

12/6/17 | SWIB17

kmika@fas.harvard.edu

esquivelndsr@gmail.com

*Share your thoughts on social media using **#BHLib***


 **BHL**
Biodiversity Heritage Library



53+ MILLION
PAGES

131,000+ | 217,000+
TITLES | VOLUMES

180+ MILLION
INSTANCES OF TAXONOMIC NAMES



715+
IN-COPYRIGHT TITLES LICENSED FOR BHL

AGREEMENTS
WITH | **300+**
LICENSORS

NDSR and BHL



The Missouri Botanical Garden resident will explore user interface modifications to the BHL portal to enable image searching, browsing, and display.

Mentors: Douglas Holland & Trish Rose-Sandler



The Chicago Botanic Garden resident will address content and gap analysis, reviewing the domain of biodiversity literature underpinning the field of biodiversity, estimating the amount of that literature in the public domain, and exploring methodologies to visualize digital collections.

Mentor: Leora Siegel

NATURAL HISTORY MUSEUM
LOS ANGELES COUNTY

The Natural History Museum Los Angeles County resident will consult with BHL partners such as DPLA and Europeana to determine how BHL data works in these large-scale digital libraries and categorize high value tools and services.

Mentor: Richard Hulser



The Harvard University MCZ resident will develop methodologies and propose tools for verification and integration of crowd-sourced data corrections, building on previous work undertaken by the Ernst Mayr Library and the Missouri Botanical Garden (MBG).

Mentors: Constance Rinaldo & Joseph deVeer



The Smithsonian Libraries resident will work with the BHL Secretariat to determine additional digital library needs and services that will provide increased value to a broader set of BHL users.

Mentors: Martin Kalfatovic & Carolyn Sheffield



Global Names Architecture

[HOME](#)[NEWS](#)[APPS](#)[PAPERS](#)[DOCUMENTATION](#)[SOURCE](#)

Global Names Architecture

Scientific Names Services

The Global Names Architecture (GNA) is a system of web-services which helps people to register, find, index, check and organize biological scientific names and interconnect on-line information about species.

Resolve

Check if the names you have are real, spelled right, and currently used.

[Resolver](#) →

Find

Discover scientific names in texts, PDFs, and even images you took in a Natural History Museum.

[Name Finder](#) →

Register

If you are a zoologist -- you can officially register species electronically.

[Zoobank](#) →

Global Names Architecture

ZooBank

GN Usage Bank (GNUB)

BioGUID

GN Parser

GN Index (GNI)

GN Resolver (GNR)

GNRD

GN NameSpotter

GN BiblioSpotter

GNITE



Global Names Architecture

ZooBank

GN Usage Bank (GNUB)

BioGUID

GN Parser

GN Index (GNI)

GN Resolver (GNR)

GNRD

GN NameSpotter

GN BiblioSpotter

GNITE



GNA and BHL

- GNRD
 - TaxonFinder
 - Neti Neti

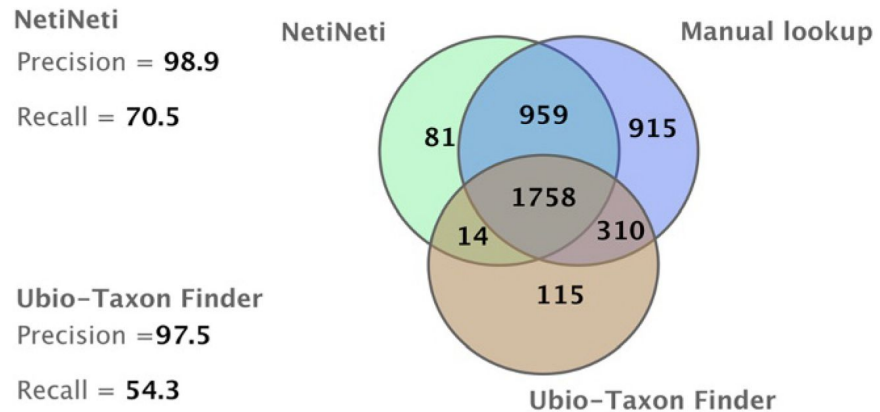


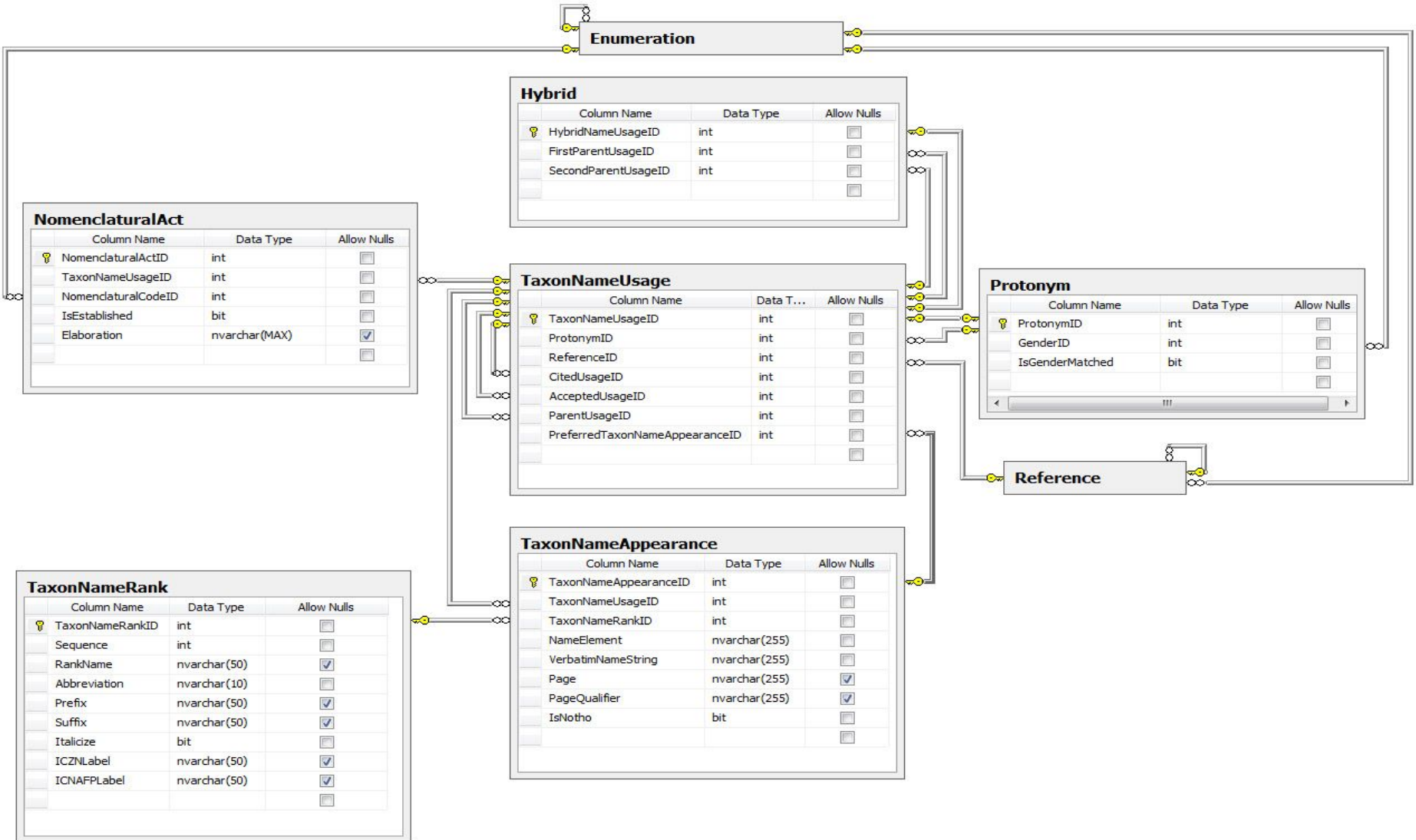
Figure 2 Comparison between NetiNeti and TaxonFinder on *American Seashells* Book in BHL.

<http://www.biomedcentral.com/1471-2105/13/211>

GNUB and TNUs

- Agents
- References
- Taxon Name Usage (TNU) instances
 - unique and persistent identifiers
 - link to reference (at page level)
 - link to Protonym TNU
 - indication of taxonomic rank
 - exact spelling of name
 - link to TNU of parent taxon

GNUB and TNUs



Current Semantic Links



BHL

About Help Feedback

BHL Store

Biodiversity Heritage Library



Inspiring discovery through free access to biodiversity knowledge.

The Biodiversity Heritage Library improves research methodology by collaboratively making biodiversity literature openly available to the world as part of a global biodiversity community.

BHL also serves as the foundational literature component of the Encyclopedia of Life (EOL).

Search across books and journals, scientific names, authors and subjects



[ADVANCED SEARCH](#)



BHL

Biodiversity Heritage Library

Current Semantic Links

Advanced Search

Books/Journals

Articles/Chapters

Authors

Subjects

Scientific Names

Title:

Author Last Name:

Volume:

Edition:

Year (YYYY):

Subject:

Content Analysis at BHL

- download name file from BHL export page
- create a dictionary in Python and split into smaller lists
- submit CSV files to GN list resolver
- Analyze classification paths with Python script
- compare kingdom counts to species estimates

Content Analysis at BHL



Scientific Names List Resolver

Help

Step 1:
Upload a CSV File With
Scientific Names



Step 2:
Map Headers



Step 3:
Pick Reference Data

Upload a CSV File With Scientific Names

This app compares your list of scientific names with other datasets. For a successful matching of names make sure that your CSV file meets the following requirements:

CSV Headers Format: Corresponds to one of the [examples](#)

Encoding: UTF-8

CSV field separator: comma, semicolon, tab (, ; \t)

UPLOAD CSV

MIT License Version 0.0.5-dev

Content Analysis at BHL

Step 1:
Upload a CSV File With
Scientific Names



Step 2:
Map Headers



Step 3:
Pick Reference Data

Map Headers

Green headers show terms that were recognized automatically. You can manually change the mapping of terms by manual editing.

There are two mutually exclusive approaches. In one scientific names are given in one field and mapped to a **scientificName** term, or they can be split into **genus**, **species**, **scientificNameAuthorship** etc.

Each term can happen only once and is removed from available terms list if it is already used. White-colored headers will be ignored during comparison.

CONTINUE

scientificname → scientificName

match with

scientificName



Pick Reference Data

Choose a data source to match your names against

CONTINUE

- Catalogue of Life
- Wikispecies
- ITIS
- NCBI
- Index Fungorum
- IRMNG (Interim Register of Marine and Nonmarine Genera)
- WoRMS (World Register of Marine Species)
- GBIF Backbone Taxonomy
- EOL
- Tropicos - Missouri Botanical Garden
- IPNI (The International Plant Names Index)
- uBio NameBank
- PaleoBioDb (The Paleobiology Database)
- The Reptile Database
- The Mammal Species of The World
- BirdLife International
- FishBase Cache
- OTT (Open Tree of Life Reference Taxonomy)
- iNaturalist

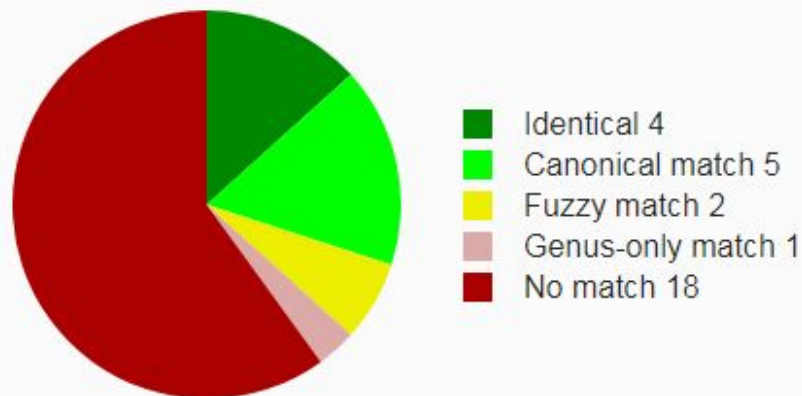
Content Analysis at BHL

Matching names from your file against "Catalogue of Life" data

When the process is complete you will be able to download results of the name matching.

Ingestion Status: Done (Processed 30 names in 0h, 0m, 0s)

Resolution Status: Done (Processed 30 names in 0h, 0m, 2s)



Download name-matching results:

[CSV file](#)

[XLSX building... \(0%\)](#)

Content Analysis at BHL

0

classification

Animalia(kingdom), Arthropoda(phylum), Insecta(class), Coleoptera(order), Curculionoidea(superfamily), Curculionidae(family), Mecysmoderes(genus)

Animalia(kingdom), Arthropoda(phylum), Insecta(class), Coleoptera(order), Curculionoidea(superfamily), Curculionidae(family), Mecysolobus(genus)

Animalia(kingdom), Arthropoda(phylum), Insecta(class), Coleoptera(order), Curculionoidea(superfamily), Curculionidae(family), Mecysolobus(genus), Mecysolobus lixoides(species)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeops(genus)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeus(genus)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeus(genus)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeus(genus), Medaeus elegans(species)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeus(genus), Medaeus grandis(species)

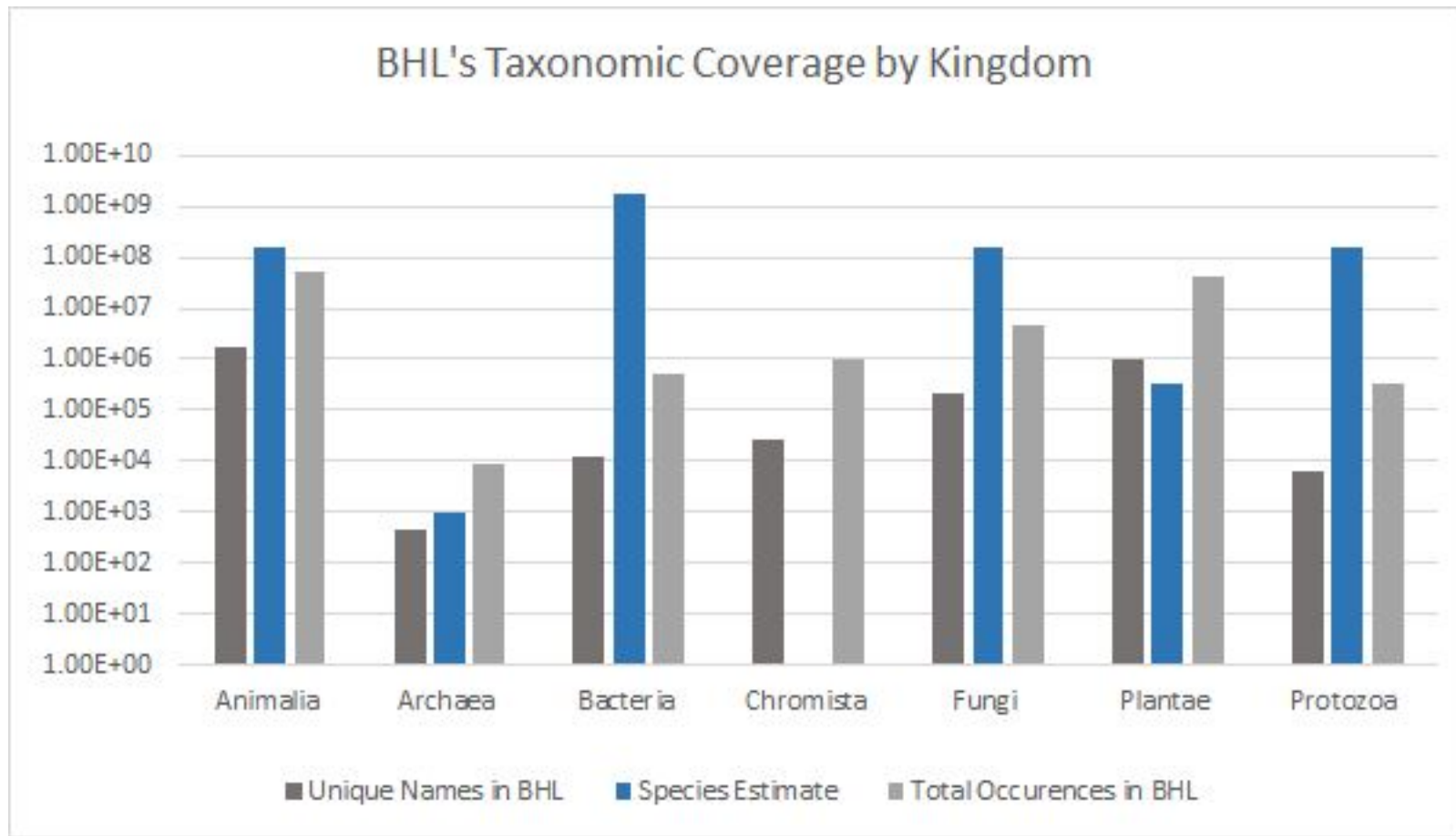
Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Medaeus(genus), Medaeus granulatus(species)

Animalia(kingdom), Arthropoda(phylum), Malacostraca(class), Decapoda(order), Xanthoidea(superfamily), Xanthidae(family), Miersiella(genus), Miersiella haswelli(species)

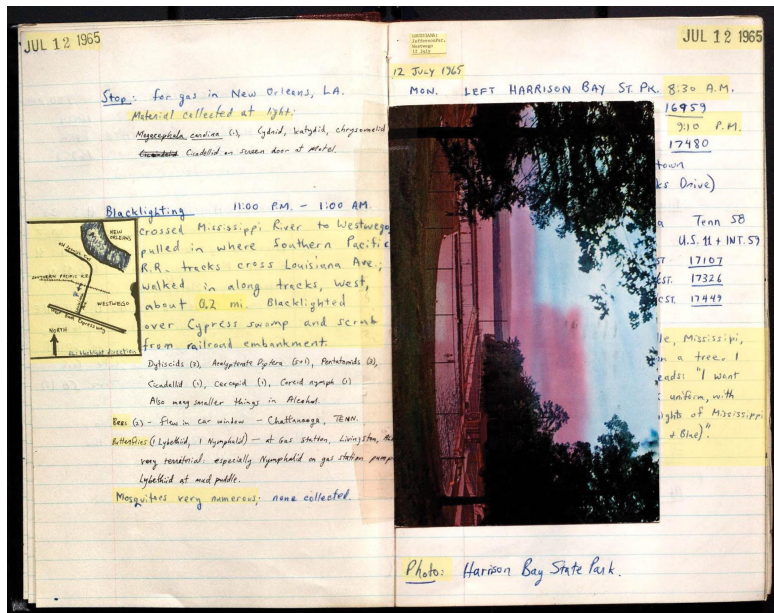
Animalia(kingdom), Chordata(phylum), Actinopterygii(class), Cypriniformes(order), Cyprinidae(family), Meda(genus)

Animalia(kingdom), Chordata(phylum), Actinopterygii(class), Cypriniformes(order), Cyprinidae(family), Meda(genus), Meda fulgida(species)

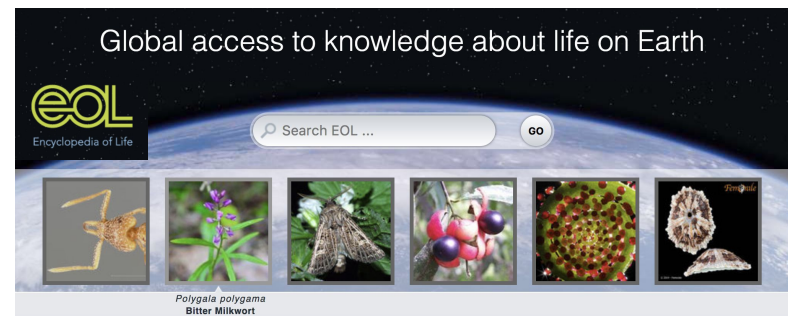
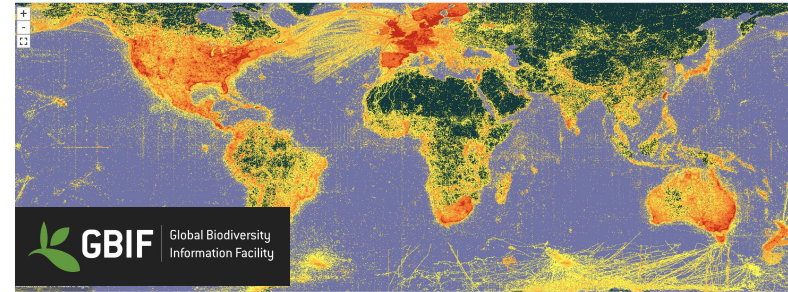
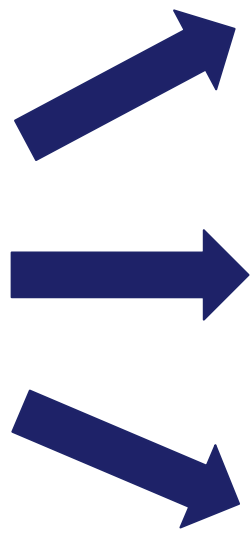
Content Analysis at BHL



BHL Collections as Data

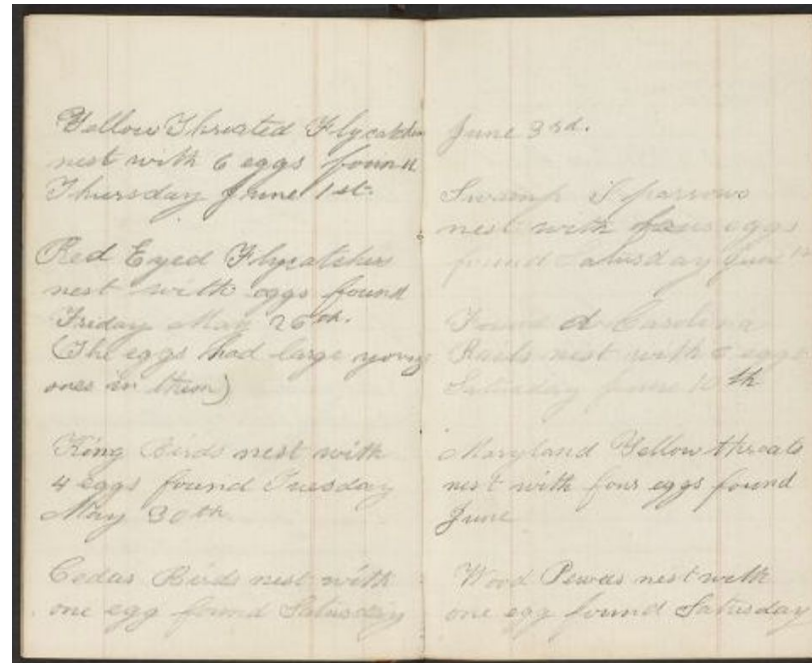


Contained In: Robert E. Silberglied Papers, 1960-1982.
<https://biodiversitylibrary.org/page/46042466>



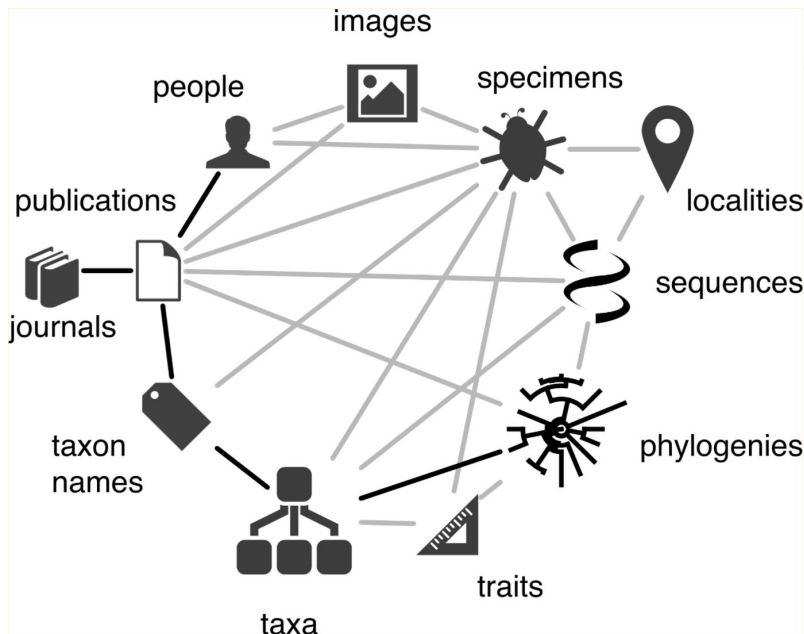
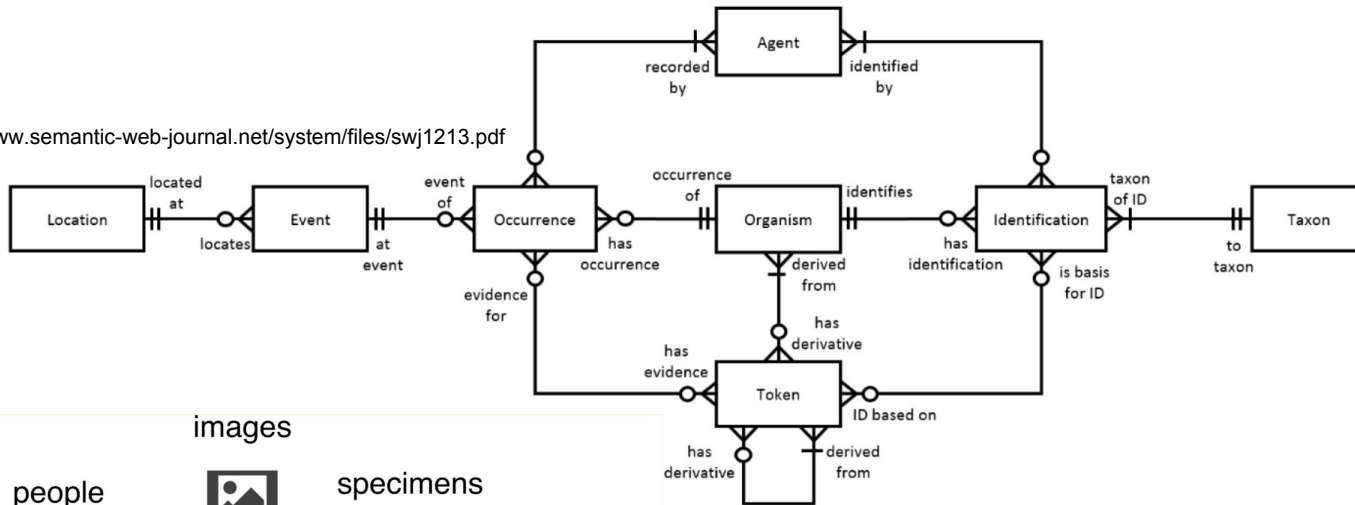
Handwritten Data

1	<i>Parus atricapillus</i>	March 4 ^① - 5 ^② - 6 ^① - 7 ^② - 8 ^⑤ - 9 ^② - 10 ^② - 12 ^② 13 ^④ - 15 ^⑧ - 18 ^② - 21 ^② - 22 ^② - 24 ^② - 26 ^② - 27 ^② - 30 ^②
2.	<i>Corvus americanus</i>	March 4 ^④ - 5 ^④ - 7 ^④ - 8 ^⑩ - 9 ^④ - 10 ^④ - 13 ^⑤ - 15 ^④ 16 ^④ - 20 ^① - 21 ^② - 22 ^⑤ - 24 ^⑤ - 26 ^⑩ - 27 ^⑩ - 28 ^⑤ - 31 ^{④③} * = "gobble"
3.	<i>Cyanocitta cristata</i>	March 4 ^① - 5 ^① - 8 ^① - 9 ^② - 10 ^② - 13 ^① - 15 ^② 20 ^① - 21 ^① - 22 ^① - 24 ^① - 26 ^① - 27 ^② - 28 ^④ - 30 ^②
4.	<i>Banius borealis</i>	March 5 ^(Baldi #1) - 9 ^② - 15 ^① - 16 ^① - 17 ^①

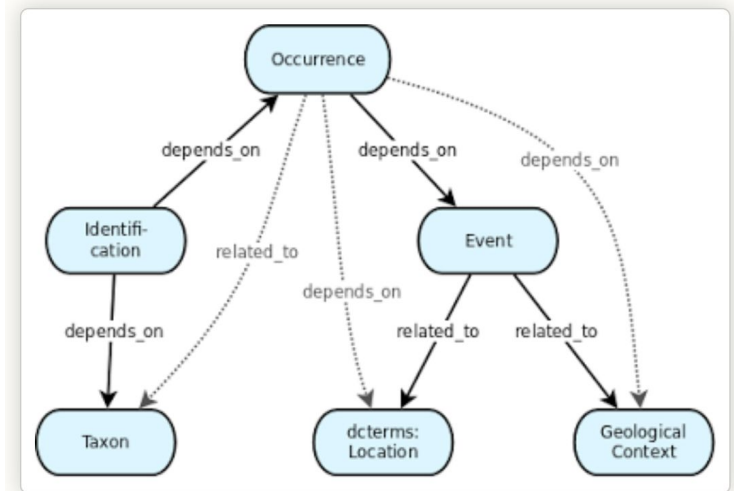


Darwin Core as Linked Data

<http://www.semantic-web-journal.net/system/files/swj1213.pdf>



<http://iphylo.blogspot.com/2016/03/towards-biodiversity-knowledge-graph.html>



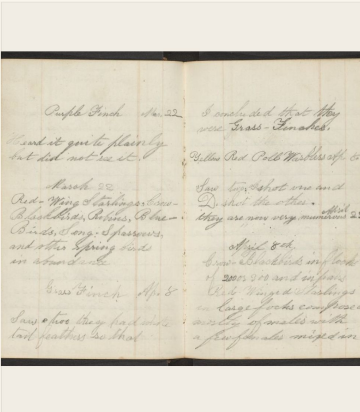
<http://biscicol.blogspot.com/2013/03/biscicol-triples-and-darwin-core.html>

Manuscript Transcriptions

FromThePage

Overview Transcribe Versions Settings

Facsimile



Transcription

[Verso]
Purple Finch [margin]Mar. 22/[margin]
Heard it quite plainly
but did not see it.

March 22
Red-Wing Starlings, Crow-
Blackbirds, Robins, Blue-
Birds, Song Sparrows,
and other spring birds
in abundance.

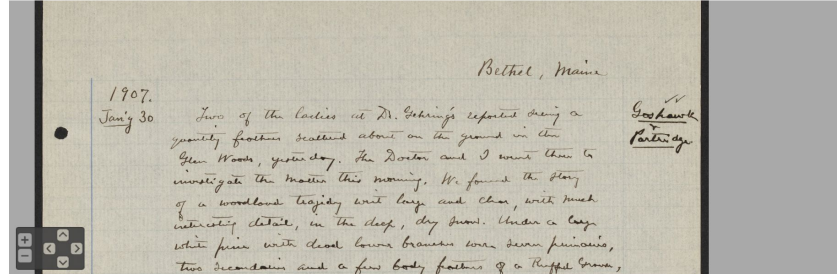
Grass Finch [margin]Apr. 8/[margin]
Saw two they had white
tail feathers so that

[Recto]
I concluded that they
were Grass-Finches.

Yellow Red Poll Warbler [margin]Apr.8/[margin]
Saw two; I shot one and
D. shot the other.
they are now very numerous [margin]April 22/[margin]

April 8th
Crow-Blackbirds in flocks
of 200 or 300 and in pairs

DigiVol Expedition My Profile



1. Verbatim Text

Bethel, Maine
1907.
Jan'y 30
Two of the ladies at Dr. Gehring's reported seeing a quantity of feathers scattered about on the ground in the Glen Woods, yesterday. The Doctor and I went there to investigate the matter this morning. We found the story of a woodland tragedy writ large and clear, with much interesting detail, in the deep, dry snow. Under a large white pine with dead lower branches were seven primaries, two secondaries and a few body feathers of a Ruffed Grouse, scattered over a space of five or six square yards. All

2. Where a species or common name appears in the text please enter any relevant information into the fields below

1. Date ? 1907-01-30 Locality Bethel, Maine. Signs unde Scientific Name ? Common Name goshawk

2. Date ? 1907-01-30 Locality Bethel, Maine. Signs unde Scientific Name ? Common Name partridge Disable

SMITHSONIAN DIGITAL VOLUNTEERS: TRANSCRIPTION CENTER SIGNUP LOGIN

HOME > PROJECTS > SMITHSONIAN INSTITUTION ARCHIVES > ALEXANDER C. ANDERSON - WESTERN UNION TELEGRAPH EXPEDITION REPORT, 1865
VIEWING PAGE 4 OF 18

[[begin page]]
is very rugged. For this reason we follow the line indicated in the map attached to my published Pamphlet. Passing by the Crow Encampment and skirting the Similkameen at the commencement of the level valley. The Okanagan should be crossed above the junction either below the lower lake or between the two lakes, both sides being alike level and under the same disadvantages for wood, save that on the east side it might be obtained more easily from the back country.

There are no difficulties along the Similkameen, the greater portion of the distance consisting of level flats interrupted occasionally by rocky hills sides covered with wood. These at low water are avoided by travellers by fording so as to reach the alternate flats on the opposite sides. Of course a telegraph line would continue along the same side (the left) on which the stream is struck and along it the Government road is now, I presume, continuously improved.

The prevailing lumber is the red pine, which you will know by its red, scaly bark, and long pendent fronds. This (the *Pinus Ponderosa*) is as its name intimates, very heavy and is moreover very subject to decay in the open air. For this reason though useful in the form of boards for indoor work it is entirely unfit for posts.

4 (circled)
[[end page]]

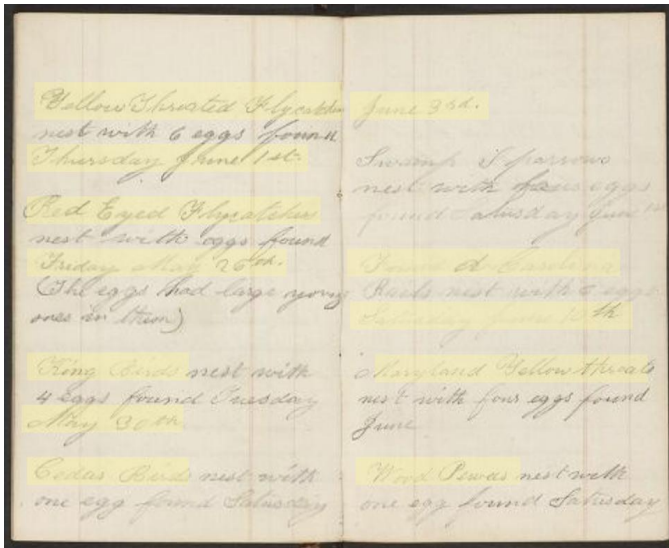
DigiVol



Smithsonian Transcription Center

Species Occurrence Events

BHL



Darwin Core - Archive Occurrence

```

taxonID, vernacularName, language
123, "sperm whale", "english"
123, "cachalot", "french"
124, "gray whale", "english"
    
```



Wieczorek, John et al. "Darwin Core: An Evolving Community-Developed Biodiversity Data Standard." *PLoS ONE* 7.1 (2012)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	
occurrenceID	basisOfRecord	eventDate	endDayOfYear	year	month	day	verbatimEventDate	eventRef	scientificName	higherClassification	kingdom	phylum	class	order
http://arctos.database.museum/guid/CUMV:PreservedSpecimen	PreservedSpecimen	1926-04		1926	4	04	1926		day of m. Ambystoma maculatum	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Cau
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1942-04-17	107	1942	4	17	4/17/42		Desmognathus fuscus	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Cau
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1942-04-17	107	1942	4	17	4/17/42		Gyrinophilus porphyriticus	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Cau
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1942-04-17	107	1942	4	17	4/17/42		Eurycea bislineata bislineata	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Cau
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1942-04-17	107	1942	4	17	4/17/42		Plethodon cinereus	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Cau
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1953-09-27	270	1953	9	27	27-Sep-53		Rana sylvatica	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Anu
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1979-06-02/1979-06-07					6/2/79		Eleutherodactylus eneidae	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Anu
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	1981-06-01	152	1981	6	1	1-Jun-81		Masticophis flagellum piceus	Animalia; Chordata; Rept	Animalia	Chordata	Reptilia	Squ
http://arctos.database.museum/guid/CUMV:Amphip:PreservedSpecimen	PreservedSpecimen	2011-06-23	174	2011	6	23	23-Jun-11		Rana (Lithobates) clamitans	Animalia; Chordata; Amp	Animalia	Chordata	Amphibia	Anu

<https://github.com/gbif/ipt/wiki/occurrenceData#templates>

Bottom: <https://biodiversitylibrary.org/page/40222549>

Collect Data

common african tree

4-segmented tarsus

California

at 6:30 at dusk



{{element | formal element | written element}}

{{element | written element}}

{{taxon | Sialia sialis | Blue Bird}}

{{location | California, USA | California}}

{{taxon | Cow bird}}

Export and relate terms

{{taxon | Northern Flicker}}, {{location | Cambridge, MA}}, {{April 15, 1865}}
{{taxon | Blue Jay}}, {{location | Cambridge, MA}}, {{May 1865}}
{{taxon | Song Sparrow}}, {{location | Cambridge, MA | Mt. Auburn St.}}

1	Work	Page_Number	Page_URL	Dates	Locations	Taxon
2	Diaries of William Brewster, 1865-1919 (inclusive)	2	https://biodiversitylibrary.org/page/40222557	1865 April 15	Cambridge, MA	Bluebird
3	Diaries of William Brewster, 1865-1919 (inclusive)	6	https://biodiversitylibrary.org/page/40222553	1865 April 15	Cambridge, MA	Bluebird
4	Diaries of William Brewster, 1865-1919 (inclusive)	6	https://biodiversitylibrary.org/page/40222553	Monday, April 19, 1865	Cambridge, MA	White Bellied Nuthatch
5	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	1865 April	Cambridge, MA	Robin's
6	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	29 April 1865	Cambridge, MA	Song Sparrows
7	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	13 May 1865	Cambridge, MA	Golden Robin
8	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	10 May 1865	Cambridge, MA	Blue Jay
9	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	7 June 1865	Cambridge, MA	Blue Jay
10	Diaries of William Brewster, 1865-1919 (inclusive)	7	https://biodiversitylibrary.org/page/40222552	1 June 1865	Cambridge, MA	White eyed Vireo

Reconcile Data

OpenRefine

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
4	4	<ul style="list-style-type: none">Gold- Finches (1 rows)Gold-finches (1 rows)Goldfinches (1 rows)gold finches (1 rows)	<input type="checkbox"/>	Gold- Finches
3	4	<ul style="list-style-type: none">maryland yellow throat (2 rows)Maryland Yellowthroat (1 rows)Maryland Yellowthroats (1 rows)	<input type="checkbox"/>	maryland yellow throat
2	4	<ul style="list-style-type: none">Golden Winged Woodpecker (3 rows)Golden Winged Woodpeckers (1 rows)	<input type="checkbox"/>	Golden Winged Woodpecker
2	3	<ul style="list-style-type: none">Crow-Blackbird (2 rows)Crow Blackbird (1 rows)	<input type="checkbox"/>	Crow-Blackbird
2	2	<ul style="list-style-type: none">Black Cap titmouse (1 rows)Black cap titmice (1 rows)	<input type="checkbox"/>	Black Cap titmouse
2	2	<ul style="list-style-type: none">Cat Birds (1 rows)Catbirds (1 rows)	<input type="checkbox"/>	Cat Birds
2	5	<ul style="list-style-type: none">Red Winged Starling (4 rows)Red-winged starling (1 rows)	<input type="checkbox"/>	Red Winged Starling

1865 April 15	Cambridge, MA	Bluebird	Paradisaea rudolphi (Finsch, 1885) Choose new match
Monday, April 19, 1965	Cambridge, MA	White Bellied Nuthatch	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new topic
1865 April	Cambridge, MA	Robin's	Etmopterus robinsi Schofield and Burgess, 1997 Choose new match
29 April 1865	Cambridge, MA	Song Sparrows	Melospiza Choose new match

taxize

```
tax_name(query = "Helianthus annuus", get = "family", db = "itis")
tax_name(query = "Helianthus annuus", get = "family", db = "ncbi")
tax_name(query = "Helianthus annuus", get = c("genus","family","order"),
  db = "ncbi")
```

Darwin Core - Archive

1	occurrenceID	basisOfRecord	eventDate	er ye m	day	verbatimEventDate	scientificName	kingdom	phy cla	oi far	geni spe	int taxonRank	identifiedBy	dat no	decimalLatitude	decimalLongitude
2	http://biodivlibrary.org/occurrence/186501	HumanObservation	1865-04-15			1865 April 15	<i>Sialia sialis</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
3	http://biodivlibrary.org/occurrence/186502	HumanObservation	1865-04-15			1865 April 15	<i>Sialia sialis</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
4	http://biodivlibrary.org/occurrence/186503	HumanObservation	1865-04-19			Monday, April 19, 1965	<i>Sitta carolinensis</i> Latham 1790	Animalia				Species	William Brewster		42.3736158	-71.1097335
5	http://biodivlibrary.org/occurrence/186504	HumanObservation	1865-04			1865 April	<i>Turdus migratorius</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
6	http://biodivlibrary.org/occurrence/186505	HumanObservation	1865-04-29			29 April 1865	<i>Melospiza melodia</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
7	http://biodivlibrary.org/occurrence/186506	HumanObservation	1865-05-13			13 May 1865	<i>Tarsiger chrysaeus</i> Hodgson 1845	Animalia				Species	William Brewster		42.3736158	-71.1097335
8	http://biodivlibrary.org/occurrence/186507	HumanObservation	1865-05-10			10 May 1865	<i>Cyanocitta cristata</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
9	http://biodivlibrary.org/occurrence/186508	HumanObservation	1865-06-07			7 June 1865	<i>Cyanocitta cristata</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
10	http://biodivlibrary.org/occurrence/186509	HumanObservation	1865-06-01			1 June 1865	<i>Vireo griseus</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
11	http://biodivlibrary.org/occurrence/186510	HumanObservation	1865-04-15			15 April 1865	<i>Sialia sialis</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
12	http://biodivlibrary.org/occurrence/186511	HumanObservation	1865-04-19			19 April 1865	<i>Sitta carolinensis</i> Latham 1790	Animalia				Species	William Brewster		42.3736158	-71.1097335
13	http://biodivlibrary.org/occurrence/186512	HumanObservation	1865-04-24			24 April 1865	<i>Turdus migratorius</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
14	http://biodivlibrary.org/occurrence/186513	HumanObservation	1865-04-25			25 April 1865	<i>Cyanocitta cristata</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
15	http://biodivlibrary.org/occurrence/186514	HumanObservation	1865-04-29			29 April 1865	<i>Melospiza melodia</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
16	http://biodivlibrary.org/occurrence/186515	HumanObservation	1865-05-05			5 May 1865	<i>Colaptes auratus</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
17	http://biodivlibrary.org/occurrence/186516	HumanObservation	1865-05-08			8 May 1865	<i>Haemorhous purpureus</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
18	http://biodivlibrary.org/occurrence/186517	HumanObservation	1865-05-03			3 May 1865	<i>Turdus merula</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
19	http://biodivlibrary.org/occurrence/186518	HumanObservation	1865-05-12			12 May 1865	<i>Dumetella carolinensis</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
20	http://biodivlibrary.org/occurrence/186519	HumanObservation	1865-05-12			12 May 1865	<i>Spizella passerina</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
21	http://biodivlibrary.org/occurrence/186520	HumanObservation	1865-05-15			15 May 1865	<i>Molothrus ater</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335
22	http://biodivlibrary.org/occurrence/186521	HumanObservation	1865-05-18			18 May 1865	<i>Tachycineta bicolor</i> (Vieillot, 1808)	Animalia				Species	William Brewster		42.3736158	-71.1097335
23	http://biodivlibrary.org/occurrence/186522	HumanObservation	1865-05-20			20 May 1865	<i>Setophaga aestiva</i>	Animalia				Species	William Brewster		42.3736158	-71.1097335

Validate!

GBIF Data Validator

VALIDATION REPORT | 29 NOVEMBER 2017

brewster_dwc-a.csv

Report produced by GBIF data validator

This an early access version. Please report issues and feedback [here](#).

System health

SUMMARY DARWIN CORE EXTENSIONS NEW VALIDATION

- The file can be indexed by GBIF

Some issues were detected by the validator:

GBIF Occurrence Interpretation Geodetic datum invalid Taxon match higherrank Coordinate rounded Geodetic datum assumed WGS84

File Format: Tabular File (.csv, .tsv)
Media Type: text/plain
Core Row Type: Darwin Core Occurrence
Extensions: 0

This report has been written to <https://www.gbif.org/tools/data-validator/1511367174936> It was generated a few seconds ago and will be deleted after one month. Until then you can revisit the report at your convenience.

Validate!

gnr_resolve()

```
> gnr_resolve(species, canonical = TRUE, data_source_ids = eol)
  user_supplied_name      submitted_name
1      Sialia sialis      Sialia sialis
2      Sitta carolinensis Latham 1790      Sitta carolinensis latham 1790
3      Turdus migratorius      Turdus migratorius
4      Melospiza melodia      Melospiza melodia
5      Tarsiger chrysaeus Hodgson 1845      Tarsiger chrysaeus hodgson 1845
6      Cyanocitta cristata      Cyanocitta cristata
7      Vireo griseus      Vireo griseus
8      Colaptes auratus      Colaptes auratus
9      Turdus merula      Turdus merula
10     Dumetella carolinensis      Dumetella carolinensis
11     Spizella passerina      Spizella passerina
12     Molothrus ater      Molothrus ater
13     Tachycineta bicolor (Vieillot, 1808)      Tachycineta bicolor (vieillot, 1808)
14     Tachycineta bicolor (Vieillot, 1808)      Tachycineta bicolor (vieillot, 1808)
15     Setophaga aestiva      Setophaga aestiva
16     Contopus virens      Contopus virens
17     Troglodytes aedon      Troglodytes aedon
18     Turdus chrysolaus      Turdus chrysolaus
19     Coccyzus erythrophthalmus      Coccyzus erythrophthalmus
20     Dolichonyx oryzivorus      Dolichonyx oryzivorus
21     Vireo griseus griseus (Boddaert, 1783)      Vireo griseus griseus (boddaert, 1783)
22     Conopias albobittatus parvus (Pelzeln, 1868)      Conopias albobittatus parvus (pelzeln, 1868)
23     Vireo olivaceus      Vireo olivaceus
24     Aquila chrysaetos (Linnaeus, 1758)      Aquila chrysaetos (linnaeus, 1758)
25     Aquila chrysaetos (Linnaeus, 1758)      Aquila chrysaetos (linnaeus, 1758)
26     Bombycilla cedrorum      Bombycilla cedrorum
27     Melospiza georgiana      Melospiza georgiana
28     Porzana carolina      Porzana carolina
29     Geothlypis trichas      Geothlypis trichas
30     Coccyzus americanus      Coccyzus americanus
31     Spinus tristis      Spinus tristis
```



Thank You!

Questions?

Katie Mika and Alicia Esquivel

12/6/17 | SWIB17

kmika@fas.harvard.edu
esquivelndsr@gmail.com

Stay Connected with BHL!

Follow **@BioDivLibrary** on social media



 **BHL**
Biodiversity Heritage Library