



Smithsonian Libraries

Linked Open Data and Systematic Taxonomy

A tale of two publications
In three acts

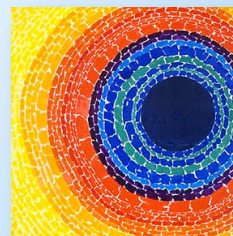
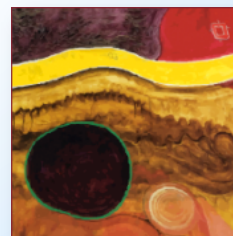
Joel Richard
Smithsonian Libraries
richardjm@si.edu

CENDI/NKOS , December 6, 2012



Who are the Smithsonian Libraries?

- 20 Libraries in the U.S. and Panama
- Supports research of staff and the public
- Strong effort to digitize pre-1923 texts
- Index Animalium and Taxonomic Literature II are two examples





Disclaimer

We are still learning.

We are still building.



Act I: The Players

(or, identifying the data with which
we are working and their meaning
and usefulness to the scientific
community.)



Taxonomic Literature II

Essential Reference
Tool for Botanists

Botanists/Authors
and Publications
from 1753–1940



Multiple indexes, “unique identifiers”

It is a “database in book form”



Darwin, Charles Robert (1809-1882). British evolutionary biologist. (*Darwin*).

Darwin, Charles Robert (1809-1882), British evolutionary biologist. (*Darwin*).

Ref.: 111 1: 152.

The literature on Darwin and his work is immense. No attempt is made here

1313. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London (John Murray) 1859. Oct. (*Origin sp.*)

literature on Darwin see Smit p. 904-913.

Main bibliography: R. B. Freeman, The works of Charles Darwin. An annotated bibliographical handlist. London 1965, x, 81 p.

Autobiographies and letters: Francis Darwin, Life and letters of Charles Darwin, London 1887, 3 vols. and Charles Darwin, his life told in an autobiographical chapter and in a selected series of his published letters, London 1892.

Biography: e.g. G. R. de Beer, Charles Darwin, evolution by natural selection, London 1963; also as Charles Darwin, a scientific biography, New York 1965 and in DSB 3: 565-577. 1971. (bibl.)

Ideas: e.g. M. T. Ghiselin, The triumph of the Darwinian method, Berkeley and Los Angeles 1969.

HANDWRITING: e.g. in Francis Darwin, The autobiography of Charles Darwin, Dover ed. 1958.

EPONYMY: *Darwiniella* Spegazzini (1887); *Darwiniothamnus* G. Harling (1962).

Note: *Darwinia* Rafinesque (1817) and *Darwinia* Rudge (1815) are dedicated to Erasmus Darwin (1731-1802), English physician, grandfather of Charles Darwin.

POSTAGE STAMPS: Poland 20 gr. (1959) yv. 997; Roumania 55 b. (1959) yv. 1621; Soviet Union 40 k. (1959) yv. 2144; Czechoslovakia 3k. (1959) yv. 1045; DDR 10 p. (1958) yv. 349; Ecuador 20 c. (1936) yv. 335.

1313. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London (John Murray) 1859. Oct. (*Origin sp.*)

Publ.: 24 Nov 1859, p. [i]-ix, [1]-502. – Facsimile edition with an introduction by Ernst Mayr, Cambridge, Mass. 1964. – See Freeman 112-207, 409-441 for all editions.



Smithsonian Libraries

Donate

Keyword Search

Collections

News & Events

Donate

Research Tools

About

Contact

Taxonomic Literature II

[TL-2 Home](#)
[Search](#)
[Read Online](#)
[History](#)
[Acknowledgements](#)
[Downloads](#)
[Help](#)

[Back to Digital Library](#)

.: Search TL-2

Advanced Search

Keyword Search

Use the criteria below to search the TL-2 database. You may enter up to four search terms and a combination of AND and OR to achieve more specific results.

Surname

Darwin

Search

<< first < prev 1 **2** 3 4 next > last >>

Author Name

[Darwin, Charles Robert](#) (1809-1882)

TL-2 Number

[1313](#)

Title

On the origin of species

Year(s)

1859

Author Name

[Darwin, Robert Waring](#) (1724-1816)

TL-2 Number

[1314](#)

Title

Principia botanica

Year(s)

1787

Author Name

[*Darwin, Charles Robert](#) (1809-1882)

TL-2 Number

[29.064](#)

Title

On the two forms, or dimorphic condition, in the species of Primula

Year(s)

1862

Joel Richard, richardjm@si.edu

CENDI/NKOS, December 6, 2012

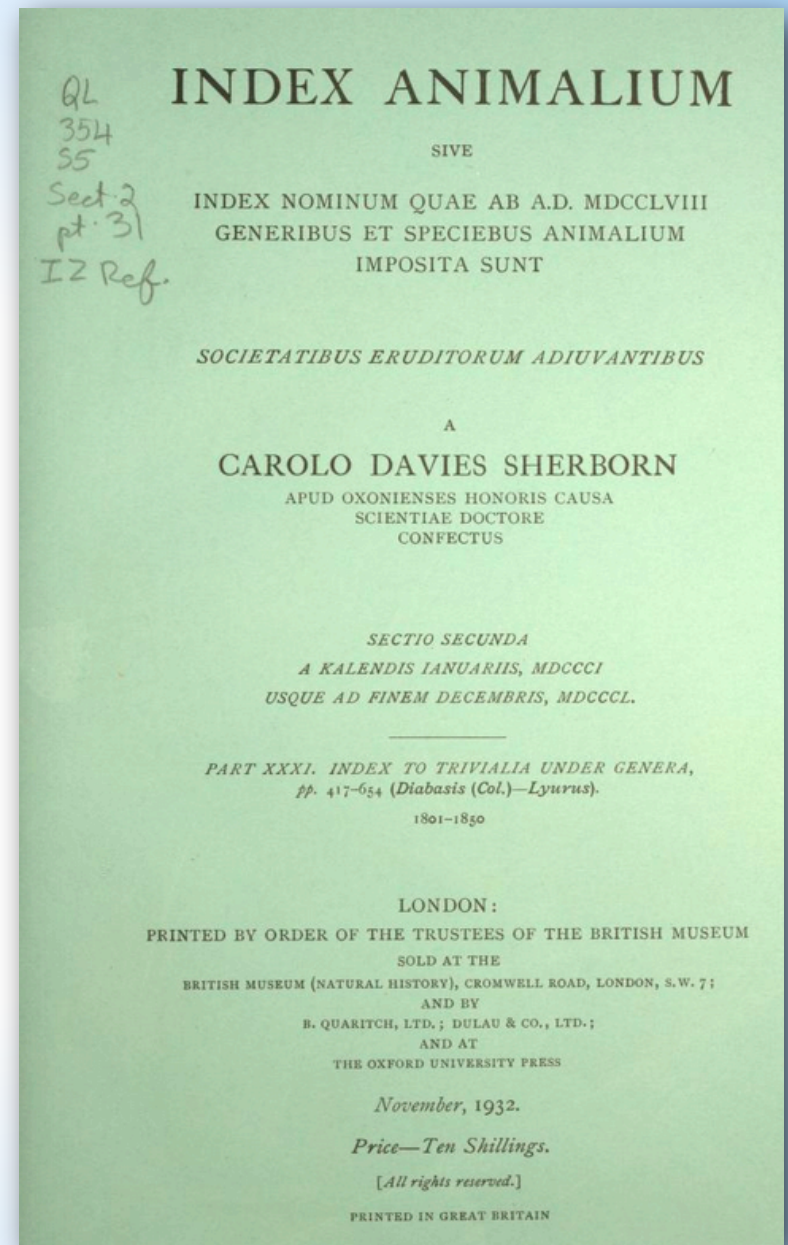


Index Animalium

Genus name, author
& citation for
430,000 animals

Covers Publications
from 1758–1850

Also a database, but
many challenges
still exist in the data.





- Plautus M. T. Brünnich, Fund. Zool. 1772, 78.—A.
 plautus Hesperia, J. C. Fabricius, Ent. Syst. III (1) 1793, 291.
 Plebejus *v.* Papilio.
 plebeja Bibio (L.), J. C. Fabricius, Syst. Ent. 1775, 757.—Musca.
 plebeia Buprestis, J. C. Fabricius, Gen. Ins. 1777, 236.
 plebejum Cerambyx, J. F. Gmelin in Linn. Syst. Nat., ed. 13, I. 1789, 185.
 plebeja Cicada, Scopoli, Ent. Carn. 1763, 117.
 plebeius Curculio, A. F. de Fourcroy, Ent. Paris. 1785, 131.
 plebeia Donax, R. Pulteney, Cat. Birds &c. Dorset, 1799, 32.
 plebeia Melolontha, A. G. Olivier, Entom. I. 1789, 25.
 plebeja Musca, Linnaeus, Syst. Nat., ed. 10, 1758, 589; ed. 12, 1767, 97.
 plebeia Phalaena, A. F. de Fourcroy, Ent. Paris. 1785, 285.
 plebeja Phalaena (Noctua), Linnaeus, Faun. Suec., ed. 2, 1761, 320.
 plebeius Polynemus, P. M. A. Broussonet, Ichth. 1782.
 plebeja Sphinx, J. C. Fabricius, Gen. Ins. 1777, 273.
 plebejus Stenocorus, J. C. Fabricius, Suppl. Ent. Syst. 1798, 147.
 plebeia Stratiomys (Linn.), O. F. Müller, Z. Dan. Prod. 1776, 177.—Mus.
 plebeja Tettigonia (Scop.), G. A. Harrer, Besch. Schaeffer Icones, 1784.
 Plecostomus L. T. Gronovius, Zooph. I. 1763, 127.—P.
 plecostomus Acipenser, Linnaeus, Syst. Nat., ed. 10, 1758, 238.
 plecostomus Loricaria, Linnaeus, Syst. Nat., ed. 12, 1766, 508.
 plecostomus Plecostomus, F. C. Meuschen, Mus. Gronov. 1778, 39.
 plecta Noctua (L.), J. C. Fabricius, Syst. Ent. 1775, 605.—Phalaena.
 plecta Phalaena (Noctua), Linnaeus, Faun. Suec., ed. 2, 1761, 321.
 "Plectorhinchus, Lacép. 1800," Ag.; is 1802.
 pleiades Madrepora, J. Ellis & D. Solander, Nat. Hist. Zooph. 1786, 169.



plexippus Papilio (Danaus), Linnaeus, Syst. Nat., ed. 10, 1758, 471; ed. 12, 1767, 767.

- Pleuronectes Linnaeus, Syst. Nat., ed. 10, 1758, 268; ed. 12, 1766, 455; [Artedi, Ichth. 1738, gen. xiv, p. 16].—P.
 pleuronectes Amusium (Gm.), J. F. Bolten, Mus. Bolten. (2) 1798, 165.—Ostrea.
 pleuronectis Ascaris, O. F. Müller, Z. Dan. Prod. 1776, 214.
 pleuronectes Cercaria, O. F. Müller, Verm. I (1) 1773, 70.
 pleuronectis Echinorhynchus, J. F. Gmelin in Linn. Syst. Nat., ed. 13, I. 1790, 3047.
 pleuronectes Monoculus, J. F. W. Herbst, Schrift. Berlin. Ges. Nat. Freunde, III. 1782.
 pleuronectes Ostrea, Linnaeus, Syst. Nat., ed. 10, 1758, 696; ed. 12, 1767, 1145. [102.
 pleuronectes Pecten (L.), G. Cuvier, Tab. Elém. 1798, 420.—Ostrea.
 pleuronectes Proboscidea (Müll.), J. G. Bruguière, Ency. Méth. (Vers Int. 1791), 97.—
 pleuronectis Scolex, O. F. Müller, Zool. Dan. II. 1784, 53. [Ascaris.
 Pleurorinchus B. Nau, Schrift. Berlin. Ges. Nat. Freunde, VII. 1787, 474.—Ver.
 Pleurotoma Lamarck, Mém. Soc. H. N. Paris, 1799, 73.—G.
 plexippus Papilio (Danaus), Linnaeus, Syst. Nat., ed. 10, 1758, 471; ed. 12, 1767, 767.
 plica Argoderma (L.), J. X. Poli, Test. Sicil. II. 1795, 254 & 260.—Ostrea.
 plica Lacerta, Linnaeus, Syst. Nat., ed. 10, 1758, 208; ed. 12, 1766, 367.
 plica Ostrea, Linnaeus, Syst. Nat., ed. 10, 1758, 697; ed. 12, 1767, 1145.
 plica Pecten (L.), F. C. Meuschen, Mus. Gev. 1787, 416.—Ostrea.
 plica-auris Vespertilio, M. T. Brünnich, Dyr. Hist. 1782, 56.
 plicarius Bulimus (Born), J. G. Bruguière, Ency. Méth. (Vers) (1) 1789, 328.—Helix.
 plicaria Helix, *err. pro* plicata, Born, I. v. Born, Test. Mus. Caes. Vindob. 1780, 389.
 plicarius Murex (L.), F. C. Meuschen, Mus. Gev. 1787, 330.—Voluta.
 plicaria Patella, J. F. Gmelin in Linn. Syst. Nat., ed. 13, I. 1790, 3708.
 plicaria Voluta, Linnaeus, Syst. Nat., ed. 10, 1758, 732; ed. 12, 1767, 1193.
 plicataria Mactra, Linnaeus, Syst. Nat., ed. 12, 1767, 1125.
 plicatella Anomia, Linnaeus, Syst. Nat., ed. 10, 1758, 702; ed. 12, 1767, 1152.



Act II: The Linking

(or, identifying those data elements to be linked, inherent challenges of parsing OCR text, and identifying linkable remote data sources)



Linkable Data Elements

Darwin, Charles Robert (1809-1882), British evolutionary biologist. (*Darwin*).

1313. *On the origin of species* by means of natural selection, or the preservation of favoured races in the struggle for life. London (John Murray) 1859. Oct. (*Origin sp.*)

plexippus Papilio (Danaus), Linnaeus, Syst. Nat., ed. 10, 1758, 471 ed. 12, 1767, 767.



<http://library.si.edu/t12/author/darwin>
RDF Type = foaf:Person

Darwin, Charles Robert (1809-1882), British evolutionary biologist. (*Darwin*).

HERBARIUM and TYPES: plants collected on the voyage of the Beagle (1831-1836): CGE, K, MANCH, MO, P, TCD (alg.).
Ref.: IH 1: 152.

NOTE: The literature on Darwin and his work is immense. No attempt is made here even to present a review of the most important biographies and studies. In a work on taxonomic literature, however, Darwin should have his place, even if only symbolic, because of his immense influence also on the development of ideas in taxonomy leading to the development of modern evolutionary biology. For a list of the main recent literature on Darwin see Smit p. 904-913.

Main bibliography: R. B. Freeman, The works of Charles Darwin. An annotated bibliographical handlist. London 1965, x, 81 p.

Autobiographies and letters: Francis Darwin, Life and letters of Charles Darwin, London 1887, 3 vols. and Charles Darwin, his life told in an autobiographical chapter and in a selected series of his published letters, London 1892.

Biography: e.g. G. R. de Beer, Charles Darwin, evolution by natural selection, London 1963; also as Charles Darwin, a scientific biography, New York 1965 and in DSB 3: 565-577. 1971. (bibl.)

Ideas: e.g. M. T. Ghiselin, The triumph of the Darwinian method, Berkeley and Los Angeles 1969.

HANDWRITING: e.g. in Francis Darwin, The autobiography of Charles Darwin, Dover ed. 1958.

EPONYMY: *Darwinia* Spegazzini (1887); *Darwiniothamnus* G. Harling (1962).

Note: *Darwinia* Rafinesque (1817) and *Darwinia* Rudge (1815) are dedicated to Erasmus Darwin (1731-1802), English physician, grandfather of Charles Darwin.

POSTAGE STAMPS: Poland 20 gr. (1959) yv. 997; Roumania 55 b. (1959) yv. 1621; Soviet Union 40 k. (1959) yv. 2144; Czechoslovakia 3k. (1959) yv. 1045; DDR 10 p. (1958) yv. 349; Ecuador 20 c. (1936) yv. 335.

1313. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London (John Murray) 1859. Oct. (*Origin sp.*)

Publ.: 24 Nov 1859, p. [i]-ix, [1]-502. - Facsimile edition with an introduction by Ernst Mayr, Cambridge, Mass. 1964. - See Freeman 112-207, 409-441 for all editions.

foaf:lastName, foaf:familyName

foaf:firstName, foaf:givenName

foaf:name, skos:prefLabel

bio:birth

bio:death

skos:definition

t12:personAbbreviation

<http://library.si.edu/t12/title/origin...>
RDF Type = bibo:Book

t12:titleNumber

dc:title

event:place

dc:publisher

t12:titleAbbreviation

dc:created



Challenges with Our Data

- Errors in the Corrected OCR
- Challenges in Parsing Citations
- The 80/20 rule: manually making connections unable to be made by automated means
- Finding suitable sources of data to link to. (DBPedia? VIAF? EOL? Others?)



Linked Data Sources

Low-Hanging Fruit:

- DBPedia
- OCLC WorldCat
- Biodiversity Heritage Library
- Virtual International Authority File
- Encyclopedia of Life
- Library of Congress Subject Headings
- GeoNames
- Open Library



Act III: The Sum of the Parts

(or, our goals and desires for this
data, what it means to the linked
data world and the scientific
community in general)



What's the point?

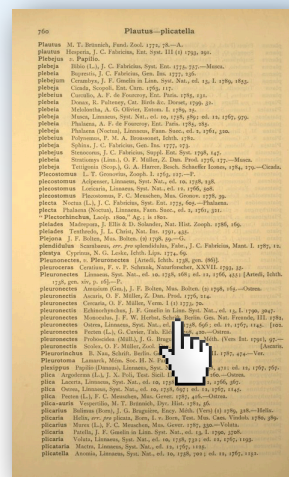
- This data may already exist online.
- It may also not always be as accurate as needed for science.
- We are in a position to be the authoritative source for this information.
- Linked Data allows it to be easily reused and shared.



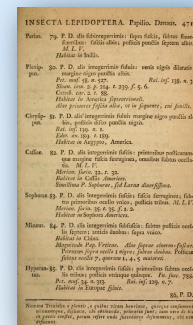
Danaus plexippus



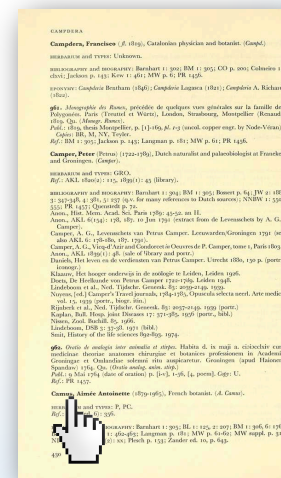
Index Animalium



Systema Naturae, etc



Taxonomic Lit. II



Your Local Library



Aimeé Antoinette
Camus
(botanist)

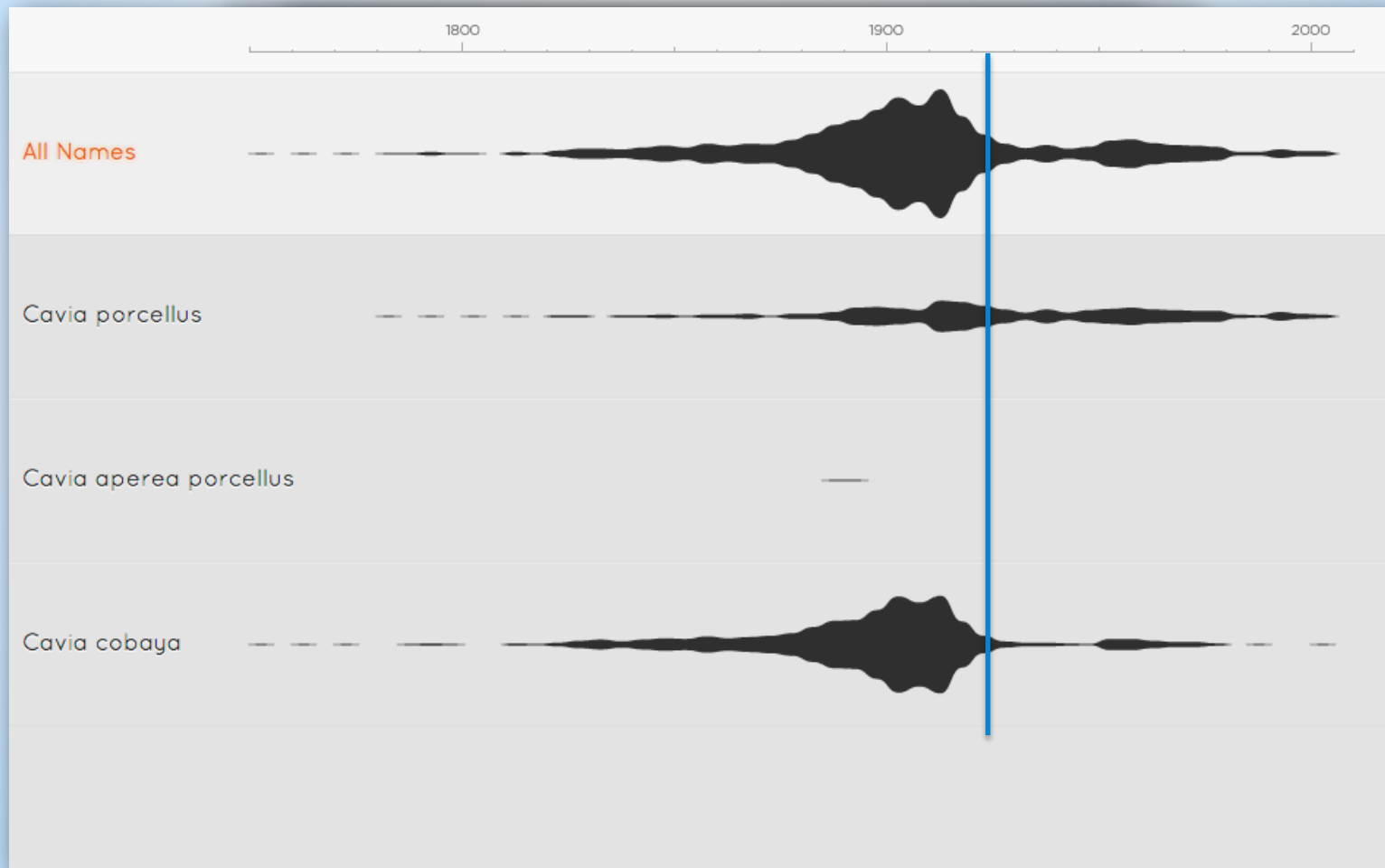




One Example of Reuse

Ryan Schenk

<http://synonyms.com/>





Thank you!

Joel Richard

RichardJM@si.edu

<http://library.si.edu/staff/richardjm>

<http://slideshare.net/joelrichard>