

echinopscis (an extensible notebook for open science on specimens): overview & design principles

TaxonWorks Together (TWT) - 26th October 2023

Nicky Nicolson¹, Eve Lucas²

(1) Digital revolution, RBG Kew, (2) Accelerated taxonomy, RBG Kew



**TaxonWorks
Together 2023**

Context

- Transitioned from software development into research
- Interested in how we can use software development practices in research:
 - Automation
 - Version control
 - Dependency management
 - Continuous integration
- Institutional commitment to use digital technologies to accelerate the process of taxonomy

Community aim: digital extended specimen

- Integrate specimens and associated data across multiple research infrastructures
 - allowing the investigation of wider scale research questions
- How to get there? Activities at a range of scales:
 - Large scale: computational
 - Distributed: lightweight tools, link construction in context by researchers

Community aim: digital extended specimen

- Integrate specimens and associated data across multiple research infrastructures
 - allowing the investigation of wider scale research questions
- How to get there? Activities at a range of scales:
 - Large scale: computational
 - Distributed: lightweight tools, link construction in context by researchers



OBSIDIAN

- A personal knowledge manager: for creating & linking research notes
- Emphasises linking
- Data stored locally, using open formats
 - Markdown and optional structured data frontmatter
- Works offline
- Extensible architecture – plugins for data access and citation processing
- Active user and developer community



OBSIDIAN

- A personal knowledge manager: for creating & linking research notes
- Emphasises linking
- Data stored locally, using open formats
 - Markdown and optional structured data frontmatter
- Works offline
- Extensible architecture – plugins for data access and citation processing
- Active user and developer community



OpenRefine

Extend Obsidian for specimen research

- Access of relevant data
 - Specimens (GBIF)
 - Names (International Plant Names Index)
 - Collections (Global Registry of Scientific Collections)
 - People (Bionomia)
 - Literature (crossref)
- Creation of links, spatial and network exploration
- Citation in new work
- Open science working practices

echinopscis.github.io

 echinopscis

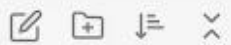
[Home](#) [Team](#) [Blog](#) [Project](#) ▾



"echinopscis" is an experiment in creating an "extensible notebook for open science" - a working environment that allows researchers to write, access data and create links between literature, specimens, names, institutions, people, traits etc.

Key principles:

- **Control of your data:** as a researcher, you remain in control of your data. The data is stored in text format, on your local machine. Text files are an open format, they will always be accessible without any need for specialised software.
- **Open to choose your working practices:** we've provided small pieces of functionality that can be combined in many different ways, enabling researchers to be "open to choose" how to organise their work.
- **Re-usable skills:** any skills necessary to work with this toolkit should be transferable to other open science tools and practices. If you invest in time exploring this prototype software, the things you learn (markdown formatting, bibliography / citation management, document production etc) could also be applied elsewhere in your work, or in other working environments.
- **Open science:** All code and documentation (and this project site) are managed on [github](#) - contributions are welcome.



echinopscis-sample-vault-0.1.7

pdfs

knapp_2010_ PDF

remote-data

bionomia

descriptions

gbif-occurrences

grscicoll

ipni

literature

scripts

templates

Glossary

Map

Start here



Start here

Welcome

This is the sample *vault* for "echinopscis" - an experiment in creating an extensible notebook for open science.

These instructions are a *page* in Obsidian, formatted in Markdown and located on your local computer. You can edit this page - toggle between view and edit mode using `ctr1 + E`.

By default these instructions are "pinned" so that they will always be visible as we create new pages. Any new pages that you create or open by navigating a link will open in a separate *pane* to the right.

About this demonstration

This vault has been configured with some tools which enable easy access, linking and visualisation of biodiversity informatics data.

The rest of this page gives a worked example of how to use these. Occasionally, some extra information about Obsidian and pointers to the Obsidian documentation has been included in this page in "tip" sections. These look like this - if you click on the coloured bar the contents will unroll.

[🔗 Tips and extra information about Obsidian >](#)

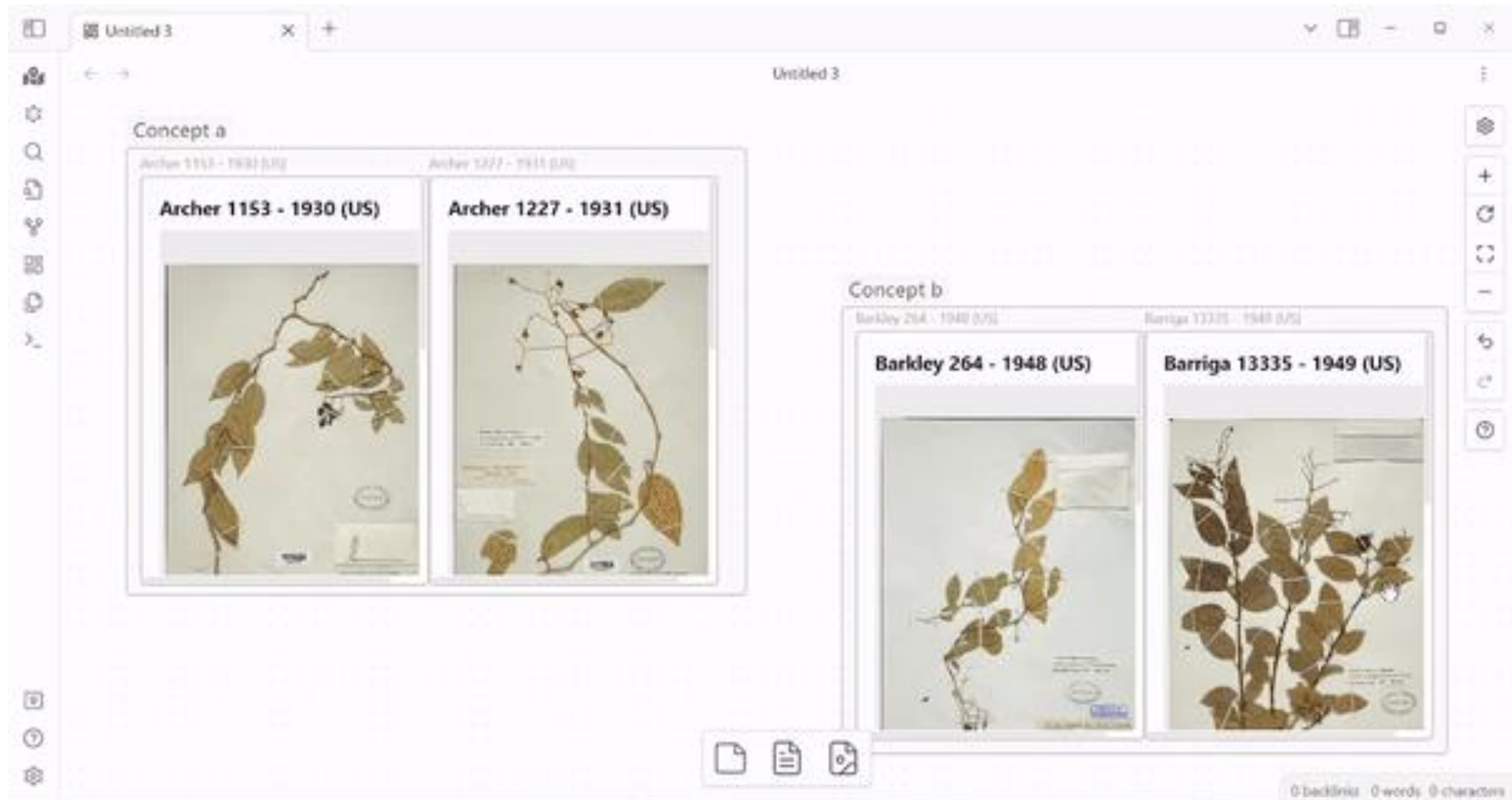
Functions are available via the command palette in Obsidian - a searchable list of the funct...

Arranging specimens: physical working space

- Our physical herbarium includes spaces for layout & comparison of large sets of specimens
- Working digitally means we can access specimen images from anywhere
- But how to effect this kind of interaction in digital environment?



Arranging specimens: virtual prototype



Roadmap

1. **Personal research environment** based on Markdown authoring and linking (now working on “seeding” – pre-populating data)
2. **Web publication** using static site generators (conceptual similarities with the [GBIF hosted portal](#) and TaxonPages work)
3. **Document production:** with structured bibliographic/specimen references
4. **Dataset production:** mobilisation of content and links into DarwinCore archives for aggregator harvesting

MSc project assessing different options for digital curation, at the Royal Botanic Garden, Edinburgh

Royal Botanic Gardens
Kew



Royal
Botanic Garden
Edinburgh



Curation of *Camellia*

Exploring and testing methods of taxonomic curation and identification
of herbarium specimens

Celia Aceae
2023



Taxonomic Literature (TL-2) – multi-volume work presented in Obsidian



Anton Savchenko
@notgaudi

Release alert 🙌

I present a digital version of Taxonomic Literature ed. 2, commonly known as TL-2. It can be viewed online or downloaded as an Obsidian database and used offline on you PC. See it here: tl2.io

The screenshot displays the 'Fries, Elias Magnus {std. Fr.}' entry in the TL-2 online database. The interface includes a search bar on the left with a list of authors, a main content area with detailed information, and a right-hand sidebar with 'ON THIS PAGE' and 'Code' sections.

TL-2 online

Search page or heading

Acosta, Isaac (std. Acosta Solis)

Acosta, Christobal (Christobal de Costa) (std. C. Acosta)

Adam, Jean (std. J. Adam)

Adametz, R. Leopold (std. Adametz)

Adami, Ernst Daniel (std. Adami)

Adamovik, Lujp (std. Adamovik) (Suppl.)

Adamovik, Lujp (std. Adamovik)

Adams, (Adams), (Johann) Michael Friedrich (std. M.F. Adams) (Suppl.)

Adams, George Jr. (std. G. Adams, Jr.)

Adams, George Sr. (std. G. Adams, Sr.)

Adams, John (std. J. Adams)

Adams, Michael Friedrich (std. M.F. Adams)

Adamski, Maximilian (Maximilian Joseph ...) (std. Adamski)

Adamsen, Robert Stephen (std. Adamsen) (Suppl.)

Adamsen, Michel (std. Adams.) (Suppl.)

Adamsen, Michel (std. Adams.) (Suppl.)

Ade, Alfred (std. Ade)

Aderhold, Rudolf Ferdinand Theodor (std. Aderhold)

Aderholdt, August Eduard (std. Aderholdt)

Adler, Wilhelm (std. W. Adler)

Fries, Elias Magnus {std. Fr.}

DOI link: this entry starts at [page 878](#) of TL-2 Vol. 1.

Supporting pages: [abbreviations](#), [layout key](#).

See also [second entry](#) for this author

Author

Fries, Elias Magnus (1794-1878), Swedish botanist at Uppsala, one of the founders of taxonomic mycology. (Fr.)

Herbarium and types

Types, acquired in 1882. Some types may also be at C. The series of exsiccatae issued by Fries are listed below. - The original drawings and paintings used by Fries are at the 'Reichsmuseum Stockholm' fide NL. Exsiccatae. [Collection B](#)

1. *Scleromyces Sveciae* (dec. i-xiv, Lund 1819-1834), sets at B (complete), BM, C, E, FH, MSTL, PC, S.
2. *Lichenes sueciae exsiccati* (fasc. i-iv, vii-ix, xiii, nos. 1-128 1824-1827; fasc. v-vi, x, [xiv], nos. 121-180, 271-300, 391-42 xii, nos. 331-360 by Th. M. Fries; fasc. xi never issued), see DUKE, FH, G, GB, GOET, HAL, L, LD, M, MSC, O, PC, S, UP
3. *Herbarium normale plantarum rariorum et criticarum* S Lund, Uppsala; details see KR p. 215), sets at B, BM, EB, C, S, W, YAG. [Sets i-ii publ. by H. H. Ringius].
4. *Hieracia europaea exsiccata* [with F. Lagger] (nos. 1-161, 1 60 nos. Uppsala ed. 1 1871, ed. 2 1872), sets at B, C, L, LE, J

Ref.: IH 2: 209; KR 214-215.

- Urban, Gesch. Berlin-Dahlem 275, 286, 347. 1916.
- Holm and Namnföldt, Friesia 7: 10-39. 1962 (on *Scleromyces* York Bot. Gard. 19(1): 125-126. 1969.
- Pfister, Mycotaxon 3(1): 185-192. 1975 (on FH set of *Scler. suec*)

ON THIS PAGE

Author

- Herbarium and types
- Bibliography and biography
- Composite works
- Eponymy (genera)
- Handwriting
- Publications

- n.1862. Novitae forae sueciae
- n.1863. Observationes mycologicae...
- n.1864. Symbolae gasteromycorum
- n.1865. Flora hallandica
- n.1866. Flora hallandica
- n.1867. Systema mycologium
- n.1868. Systema orbis vegetabilis
- n.1869. Stripes agri femionensis
- n.1870. Elenchus fungorum
- n.1871. Novitae forae sueciae.

Code: B, C, E, FH, MSTL, PC, S.

Organization: ZE Botanischer Garten und Botanisches Museum, Freie Universität Berlin

Location: Berlin, Berlin, Germany

Total number of specimens: 3800000

Current status: Active

Date founded: 1815

Taxonomic coverage: All groups

Geographic coverage: Worldwide, especially central Europe, Mediterranean area, southwestern Asia, Africa, Cuba, and

europaea
sueciae
am
generis
zariiske
romysetum
ria
Sver de i
svampar
as
n.1861. Fungi germaniae
n.1862. Epilichis systematis mycologi

Integration: widescale adoption of Reconciliation API

- **“strings to things”** – the echinopscis demo shows multiple workflows where a user translates a piece of text (like a specimen reference) into an entity managed by an external authority (like a specimen record)
- The [Open Refine Reconciliation API](#) is now documented as a standard and can be implemented:
 - by the data provider, those that maintain an authoritative dataset
 - Bionomia: <https://bionomia.net/reconcile>
 - Catalogue of Life: <https://github.com/CatalogueOfLife/backend/issues/1265>
 - Datasette (generic): <https://github.com/drkane/datasette-reconcile>
 - by data consumers: developers of taxonomic / data manipulation tools:
 - Open Refine: <https://openrefine.org/docs/manual/reconciling>
 - Python: <https://github.com/jvfe/reconciler>
 - Could we build an Obsidian plugin?

Integration: navigation of linked data

- “**entity explosion**” is a browser plugin that shows links from Wikidata based on the currently loaded page
- **Should be possible to develop an Obsidian equivalent**, that uses an ID stored in page frontmatter to run a Wikidata query and show useful links

The screenshot shows a web browser window displaying the IPNI record for Richard Kenneth Brummitt (1937-2013). The browser address bar shows the URL `ipni.org/a/1221-1`. The IPNI logo and search bar are visible at the top. The main content area displays the following information:

- Standard Form:** Brummitt
- IPNI Life Sciences Identifier (LSID):** urn:lsid:ipni.org:authors:1221-1
- Alternative Abbreviations:** Brummitt From Meikle
- Alternative Names:** Brummitt, Dick
- Area of Interest:** Spermatophytes
- Example of Name Published:** Airyantha Brummitt in Kew Bull. 22: 375. 1968.
- Countries:** United Kingdom

Overlaid on the right side of the page is a Wikidata overlay window. It features a dropdown menu for **language** set to `[en] English` and another dropdown for **entity** set to `Richard Kenneth Brummitt (Q372403)`. Below these, the Wikidata entry for Richard Kenneth Brummitt is displayed, including:

- Wikidata: [Q372403](#)
- Wikipedia: [Richard Kenneth Brummitt](#)
- Wikispecies: [Richard Kenneth Brummitt](#)
- academic degree: **Doctor of Philosophy**
- country of citizenship: **United Kingdom**
- field of work: **botany**
- given name: **Richard**
- given name: **Kenneth**
- instance of: **human**
- languages spoken, written or signed: **English**
- occupation: **botanist**
- place of birth: **Liverpool**
- sex or gender: **male**
- date of birth: `1937-05-23T00:00:00Z`

The Windows taskbar at the bottom shows the search bar with the text "Type here to search" and various application icons. The system tray on the right indicates the time as 13:46 on 26/10/2023.

Explore

Background, demo videos, installation information, roadmap and ideas for contributions:

echinopscis.github.io

Source code and documentation:

github.com/echinopscis



n.nicolson@kew.org



[nickynicolson](https://github.com/nickynicolson)



[nickynicolson](https://twitter.com/nickynicolson)