

CW25 Hack Day

Rethinking Publication - Versionable Papers



Team Magnolia



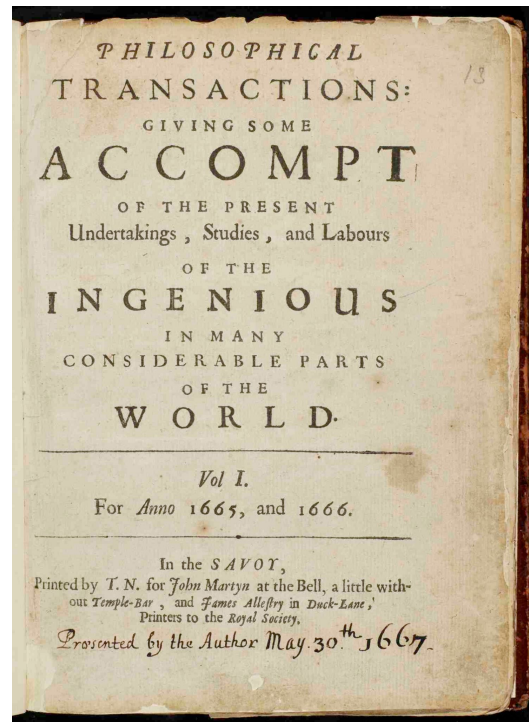
Jack Atkinson, Liam Pattison, Mateus Harrington,
Matthew Gillman, Neil Chue Hong, Nicky Nicolson,
Richard Acton, Tom Bland, Scott Archer Nicholls

Motivation

- Researchers use code to generate and/or analyse results
- Sometimes the code and/or data change
 - These may be a citable resource...
 - But may not justify a new paper

Solution

- (Semantic-like) versioning of papers: the updated version would be made available along with previous.
- A tool to automatically generate a new version of the paper with the new code and/or new data.



Philosophical Transactions Volume 1
by [Henry Oldenburg](#) licensed under
[CC BY](#)

How we worked

Attempted Agile (thanks Scott!)

2 teams

- Prototype Pipeline for example
- Exploration of broader concepts

4 Sprints with retrospectives between

- Exploratory
- Progress x2
- Pulling together and polishing

New [Github repo](#) made, with projects page for management

We did not use any GenAI tools for this Hack

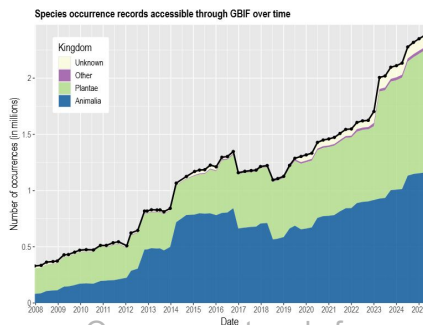


Github Repo

Goble, C. (2024, February 16). FAIR Digital Research Objects: Metadata Journeys. University of Auckland Seminar, Auckland. Zenodo. <https://doi.org/10.5281/zenodo.10710142>

Case study - assessing species occurrence

- We protect species by “redlisting” them
- *Trachelophorus giraffe* from Madagascar
- We use observation and collection points to calculate metrics like extent of occurrence
- These are shared in a data aggregator:
<https://gbif.org> - constantly updated

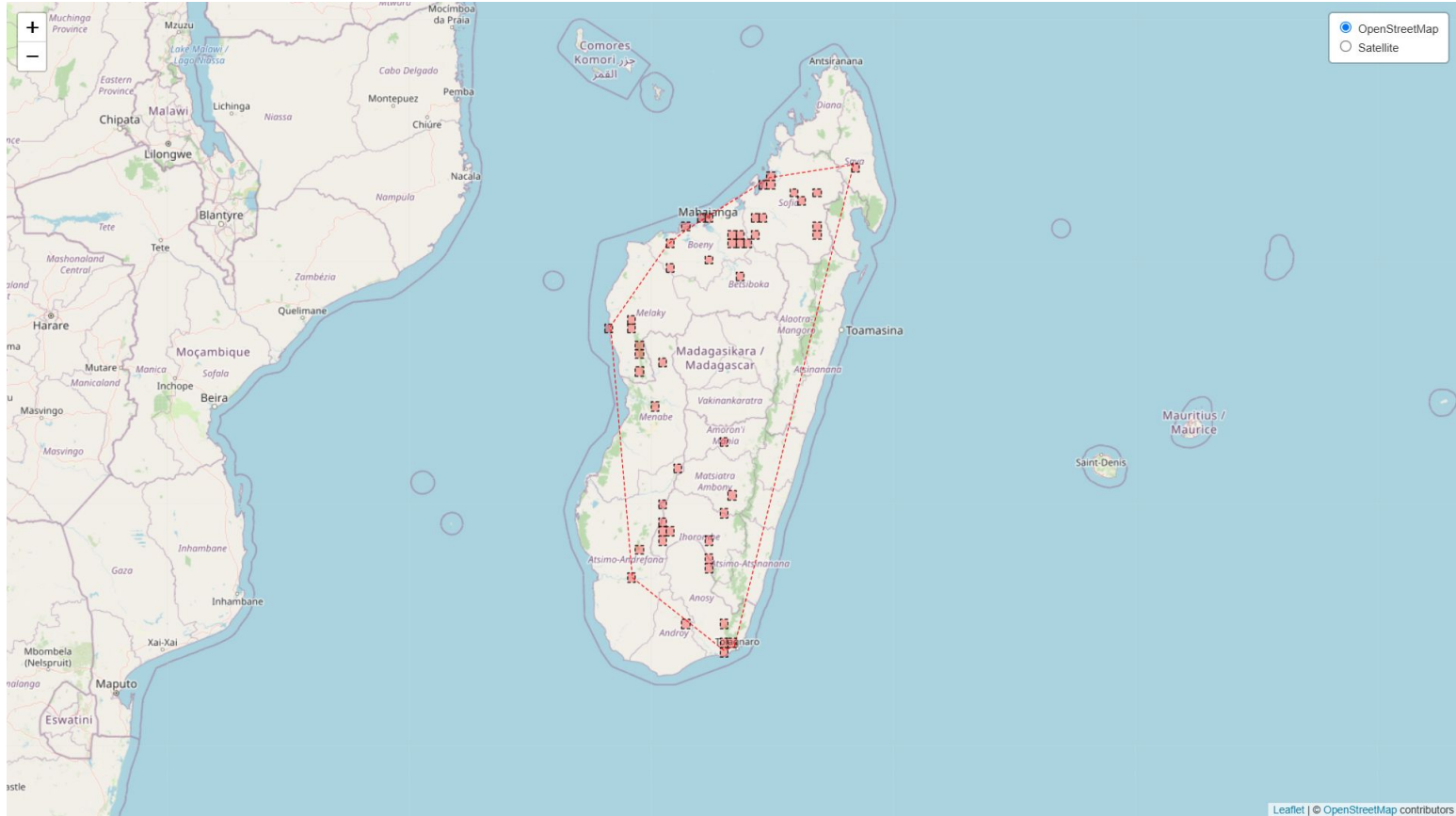


Occurrence trends for
Madagascar: data from gbif.org



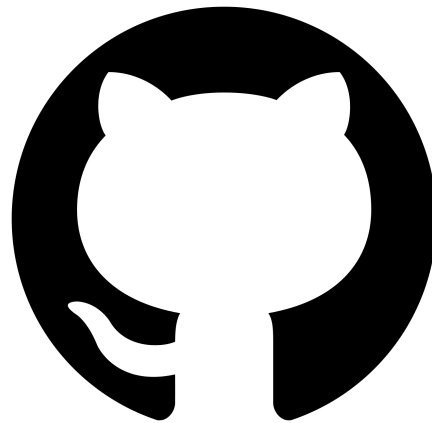
Giraffe-necked weevils (*Trachelophorus giraffae*)
by [Frank Vassen](#) licensed under [CC BY 2.0](#)

New data triggers new extent of occurrence calculation



Technical implementation

- [Targets](#) (R package) for Make-like pipeline
- As data, targets will re-evaluate relevant analysis and render [Quarto](#) file html/pdf to produce updated paper
- CI to build Docker image from renv.lock file (defines R environment) and publish paper with appropriate versioning



Demo

<https://github.com/jatkinson1000/versionable-papers/tree/targets-skeleton>

Future outlook

- Strong fit to Jack's [Fellowship plans](#) on improving interaction between software and publication
- Integrating [RO-Crate](#) and DOIs for all elements
- [Semantic Versioning](#) - understanding how to implement an incremental research pipeline and what constitutes a micro / minor / major update
- Versioning of research should be linked with work on [ReSA Authorship Task Force](#)
- Hack has created a lot of valuable input for the [UKRI Research Data Policy consultation](#)
- Conversations with [All Contributors](#) team to ensure that everything is captured, e.g. the people collecting the specimens

Thank you for listening