

Clavis

An open and versatile identification key format



Identifying taxa requires expert knowledge, which is becoming scarce. Observers need access to such knowledge, but identification keys are hard to use and even harder to make. An open, modern data format unambiguously capturing interlinked identification key data is needed to allow for interoperable solutions for editing and displaying digital keys.

Here, we present Clavis, an open data format schema for capturing knowledge required for taxon identification through digital keys, allowing for a level of detail beyond that of any current key format.

Background and key features

Clavis (**Clavis Lightweight And Versatile Identification Schema**) was developed after years of experience with tabular keys. Anything can be put into a sufficiently complex table, but the limitations of the flat, 2-axis structure made the contents increasingly repetitive, and the code to display it to the user needlessly complex. Clavis is a JSON (**J**ava**S**cript **O**bject **N**otation) schema, defining how to write and automatically validate any Clavis key. As keys are stored in JSON, they contain the required complexity in a compact manner, and are natively interpretable by all modern programming languages.

Clavis is fully open, and supports multilingualism, encoded geographical context, extended documentation and metadata, external services, non-binary characteristics, and more.

We deliberately present the Clavis format separately from any specific implementation. Solutions for editing and displaying keys have shorter life spans, and may be tailored to specific intended audiences. Our aim with presenting Clavis to the community separately is that it may serve as a solid, collaborative foundation for tools and data exchange in the long haul.

Multilingualism

All strings, references, urls and media elements support multiple versions for different languages. No need to maintain several copies of essentially the same key.

Machine readable geography

Keys, taxa, characters and states can be linked to GeoJSON polygons. Based on a location, only relevant taxa and characters may be shown, and with their appropriate local value and context.

Selection of notable features

Custom external services

Reference any external service for dynamic key content. This way, the interface can be aware of taxon name changes, retrieve multimedia objects, references to texts, etc.

Hierarchical taxonomies

Taxa are stored in a tree structure, where characters may be linked to the appropriate higher taxonomic level. Custom non-taxonomic divisions, e.g. "complex", "sex", are supported.

What's with the Pokémon?

To exemplify Clavis' features without the distraction of taxonomic debate, we use Pokémon in our manuscript. Check out the preprint:

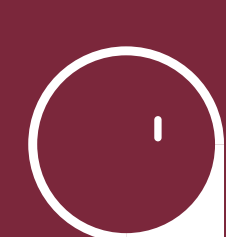
doi 10.1101/2022.05.26.493630



Wouter Koch

Senior adviser, NBIC
PhD candidate, NTNU

@wouterkoch



NORWEGIAN BIODIVERSITY
INFORMATION CENTRE



Norwegian University of
Science and Technology

Abstract doi
10.3897/biss.6.91225



All poster content by Wouter Koch
CC BY 4.0