Climate
of the Past

Open Access

Discussions

# *Interactive comment on* "Technical Note: The Linked Paleo Data framework – a common tongue for paleoclimatology" *by* N. P. McKay and J. Emile-Geay

**I. Hessler**

ines.hessler@gmail.com

Received and published: 4 November 2015

Comment(s) "McKay and Emile-Geay, Technical Note: The Linked Paleo Data framework – a common tongue for paleoclimatology, Climate of the Past Discussions 11, 4309-4327, 2015"

Overall I think the article presented by McKay and Emile-Geay reflects a great initiative that intends to facilitate research data sharing, discovery and reuse within the paleoclimate community. The authors are correct in stating that there is currently no universal way to describe, store and share paleoclimate data, which, unfortunately, also applies to most other research disciplines. However, the question is if this is due to the lack

of suitable data formats, metadata standards and/or available infrastructure or due to other aspects such as competition, giving other tasks precedence or simply unawareness of how and where to share research data. As I am not a reviewer of the article by McKay and Emile-Geay, I simply would like to leave some food for thoughts.

1. In the introduction chapter the terms data standard, metadata standard, data container and format are used fairly inconsistently. At times it is hard to follow if the authors talk about data, metadata or data formats. A brief definition of the individual terms may be appropriate to ensure a consistent understanding of their meaning. 2. While talking about universally readable data formats the authors mention netCDF format, which is a widely applied and accepted 'self-describing, machine-independent data format that supports the creation, access, and sharing of array-oriented scientific data'. Therefore, the authors choice of the JSON-LD data format strikes me as unusual as this format, at least to my understanding, is not widely known and applied. Introducing a largely unknown/unused format to the community might result in it not being readily adopted. 3. Regarding the proposed metadata standard I am wondering about the necessity to introduce yet another standard. On the one hand netCDF offers a metadata convention for e.g. Climate and Forecast (CF) data that is easily incorporated into the netCDF files (netCDF CF 1.6) itself. On the other hand there are several metadata standards specifically designed for geo information (e.g., ISO19115) that have several multi purpose fields that can house otherwise non assignable information. 4. Regarding the unique identifier I would recommend looking into Digital Object identifier (DOI) that could either be associated to the individual data set or to the data collection. Most countries have central agencies, universities or research institutions that provide DOI minting services. 5. To link information from data files, metadata, authors, publications and grants I would recommend looking into the infrastructures that are already in place in universities and/or libraries. These often fairly sophisticated systems have been set up for exactly such purposes. In general it might be a good idea to involve a (local) liaison librarian, a member from the universities eResearch group (if existent) or research office, as they are often familiar with issues related to data sharing, storing,

discoverability, linkage and reusability, and are potentially a good source of information. I also would recommend looking into Research Data Alliance as they intent to 'build the social and technical bridges that enable open sharing of data' across various disciplines.

———————————————

Interactive comment on Clim. Past Discuss., 11, 4309, 2015.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper