

Interactive comment on “Technical Note: Open-paleo-data implementation pilot – The PAGES 2k special issue” by Darrell Kaufman and PAGES 2k special-issue editorial team

Anonymous Referee #1

Received and published: 8 March 2018

This is not technically a scientific paper, but more a technical note/review. However, it is very powerful and brings to light three key issues in today's research world and that is, 1) when and how should data that support a paper be made accessible to the research community, 2) how do you cite data, and 3) how long should that data be archived for.

This can be reviewed in two contexts: Firstly, the general principles of science and secondly the context of climate science.

1) The general principles of science. I am a classically trained scientist and it was instilled in me that a true scientist is one whose data and conclusions can be independently verified. That is, when a paper is going through the review process, the data that

C1

the paper is based on, should be made available for review so that it can be determined if the conclusions are based on sound data.

In the 1970s as computerisation slowly took over data acquisitions systems and technologies became more automated, data volumes began to explode and reached the point where they could no longer be published in research papers, mainly because we were still in the era of typesetting and the printing press. This led to researchers not making it accessible, mainly because it was extremely difficult and personalised data collections became common, which then became regarded as a researcher's competitive edge.

In today's age, with online databases and data systems, this is no longer the case, and data (both raw and highly processed) as well as derivative data products and models should be made available at the time of review to assess scientific integrity of the research paper. Whether the data should then be made more accessible is to some extent dependent on the community norms. Some medical and social science data cannot be made accessible, but the norm is that most research data, particularly that collected with public money should be.

Although this is Copernicus, I note that a similar discussion page is thriving on the AGU web site around the topic 'put it online'.

2) The context of climate science. The ideas in this paper are not unique to climate science - in fact many other disciplines are struggling to come to terms with what is proposed in this paper. There are no community norms, although there is evidence that some disciplines are now working together with the publishers to develop best practice guidelines and norms - e.g., in the Earth and environmental sciences (<https://eos.org/agu-news/enabling-fair-data-across-the-earth-and-space-sciences>)

As a scientist who has been struggling with this issue for decades, I think that this paper gives an excellent summary of the key ideas and relevant problems. It will become a reference paper for those that are actually working on trying to solve this issue as it

C2

clearly and succinctly documents the key issues. I can image that it will not be popular as is evidenced by what is appearing on the discussion forum.

Some of the issues the paper is raising are beginning to be addressed in some areas, but these are mostly unpublished and hence cannot be referenced. The advantage of this paper is that it concisely ties all the key issues into a coherent succinct paper that will raise greater awareness of the problems and be of enormous value to those that are now attempting to solve them.

I only found one error that I feel needs to be corrected and that is on line 13 on page 6 where the word 'providence' is used - I think this should be 'provenance'. It is the most minor of errors.

Interactive comment on Clim. Past Discuss., <https://doi.org/10.5194/cp-2017-157>, 2017.