



## Open-paleo-data implementation: PAGES 2k and SISAL databases and others

Belen Martrat (1), Darrell Kaufman (2), and Marie France Loutre (3)

(1) Institute of Environmental Assessment and Water Research (IDÆA-CSIC), Department of Environmental Chemistry, Barcelona, Spain (belen.martrat@idaea.csic.es), (2) School of Earth Sciences & Environmental Sustainability, Northern Arizona University, Flagstaff, USA (darrell.kaufman@nau.edu), (3) PAGES International Science Office, Bern, Switzerland (marie-france.loutre@pages.unibe.ch)

Data Stewardship is one of the current PAGES Integrative Activities. It seeks to advance best practices for archiving paleodata such that they are easily and accurately reused for future and potentially unforeseen purposes. Many PAGES Working Groups are engaged in major data syntheses. Their goal is to develop community endorsed products to investigate global-scale issues and to benchmark model simulations. Creating a community wide product is challenging because it should involve broad international participation by regional and topical experts, assure that the data are accurate, maintain a high degree of uniformity and completeness, promote transparency, accountability and reproducibility in the procedures used to assemble the product, and assign proper credit to both data producers and database developers. Objective criteria must be crafted to select a particular subset of the vast assortment of paleodata in a way that maximizes the potential scientific benefit and minimizes potential selection biases, including which of the older data recourses should be rescued and which new data are missing from public archives. The essential metadata must be identified and extended metadata fields must be developed to encapsulate information needed to correctly reuse the data, including the evidence that underpins the interpretation and cautionary notes about alternative or evolving interpretations or shortcomings of individual records. A flexible database structure is essential to contain additional fields of unlimited variety, and a systematic and universal versioning scheme is required to track revisions and cross-reference individual records. We will present examples of strategies used to address these challenges in context of the paleoclimate proxy datasets that have been assembled for the PAGES 2k network (Emile-Geay et al., 2017; Kaufman et al., 2018), the SISAL working group (Atsawawaranunt et al., 2018) and others.

Atsawawaranunt, K., Comas-Bru, L., Amirnezhad Mozhdehi, S., Deininger, M., Harrison, S. P., Baker, A., Boyd, M., Kaushal, N., Ahmad, S. M., Ait Brahim, Y., Arienzo, M., Bajo, P., Braun, K., Burstyn, Y., Chawchai, S., Duan, W., Hatvani, I. G., Hu, J., Kern, Z., Labuhn, I., Lachniet, M., Lechleitner, F. A., Lorrey, A., Pérez-Mejías, C., Pickering, R., Scroxton, N. and Members, Sisal Working Group, 2018, The SISAL database: a global resource to document oxygen and carbon isotope records from speleothems, *Earth Syst. Sci. Data*, 10, 3, 1687-1713, DOI: 10.5194/essd-10-1687-2018

Emile-Geay, J., McKay, N.P., Kaufman, D.S., Routson, C.C., von Gunten, L., Wang, J., Anchukaitis, K.J., Abram, N.J., Addison, J.A., Curran, M.A.J., Moy, A.D., Evans, M.N., Henley, B.J., Hao, Z., Ge, Q., Shao, X., Martrat, B., McGregor, H.V., Neukom, Raphael, J., Rixt de, G., Martin, P., Gregory T., Stenni, B., Thirumalai, K., Werner, J.P., Xu, C., et al., 2017, A global multiproxy database for temperature reconstructions of the Common Era, *Scientific Data*, 4, 170088, DOI: 10.1038/sdata.2017.88

Kaufman, D., Abram, N. J., Evans, M. N., Francus, P., Goosse, H., Linderholm, H., Loutre, M.-F. , Martrat, B. , McGregor, H. V., Neukom, R., St. George, S., Turney, C. and von Gunten, L. , 2018, Technical note: Open-paleo-data implementation pilot – the PAGES 2k special issue, *Clim. Past*, 14, 5, 593-600, DOI: 10.5194/cp-14-593-2018