CIF21 DIBBs # 1443037 Collaborative Research: Cyberinfrastructure for Interpreting and Archiving U-series Geochronologic Data (DIBBs17, 11-12 Jan 2017)

James F. Bowring (PI)[1], Andrea Dutton[2] Noah M. McLean[3], Kenneth Rubin[4]

(1) Col of Charleston, Charleston SC, (2) Univ of Florida, Gainesville FL, (3) Univ of Kansas, Lawrence KS, (4) Univ of Hawaii, Honolulu HI

Uranium-series geochronology relies on the predictable isotopic decay of naturally occurring uranium and its daughter isotopes over time to date the crystallization age of speleothems and corals and to determine when the Earth's mantle melts to make magma. These varied applications require a wide range of isotopic measurements, made with a mass spectrometer, as well as a diverse suite of reference measurements used to calibrate them. Every laboratory performs these measurements differently and uses a different suite of reference materials. Our most significant challenge thus far has been to abstract from this variety of data and results a concise and parsimonious data dictionary and schema that represents all of this information in a way that minimizes redundant and thus potentially conflicting data. One complication, for example, is that many of the assumptions and physical constants used in reducing this isotopic data and producing U-series dates have changed over the last thirty years, thus making comparisons among results from different laboratories and scientists difficult. The older legacy data is still relevant and is often the only information we have from important regions or periods of time.

We have solved this challenge by developing a data dictionary and schema and compiling data from many different labs into a large database, published recently by PALSEA. We have worked to capture as much data as possible from each publication included in the dataset, and to regularize the data so that it conforms to the schema. We have also identified both a minimum set of data needed to reproduce the U-series dates and therefore the fundamental science behind them, as well as a larger set of recommended data that expert practitioners can use for quality control and assessment. The minimum and recommended data have been outlined in a recently submitted paper that argues that these should become data reporting standards for the larger U-series community.

Our biggest remaining challenge is completing and testing the current, collaboratively-developed open source version of the *U-series ET_Redux* software by this DIBBs team, then releasing it for an extended phase of wider community testing and input, and drumming up a critical mass of community interest. We anticipate feedback about both the function and design of the software, hopefully along with requests for help in integrating the software into laboratory information workflows. To help launch the software and engage the community, we will run a workshop at the large Goldschmidt Conference for geochemists in 2017.

Beyond the software release, by far the biggest challenge in science cyberinfrastructure is sustainability. While the funding for this project will run out in one year, the U-series community will not stop innovating new ways to make more precise and accurate measurements, and developing new methods that need to be integrated into existing knowledge and datasets. Our long-term task is to foster community interest and secure funds to provide this long-term support, since the need for this software does not expire with the end of our grant. We see outreach and expansion as critical to maintaining this momentum, and will look to create interoperability with other major cyberinfrastructure efforts. For instance, the popular paleoecology database Neotoma uses U-series dates to record the temporal as well as spatial trends in diversity through time. These records can be linked to U-series data in the database geochron.org and dynamically recalculated using *U-series ET_Redux* functionality.