Dear Dr. Martin -

Thank you for your request for clarifications of our Geoinformatics proposal, “Collaborative Research: Catalytic Track: Improving Accuracy and Efficiency of Multicollector Mass Spectrometry” (EAR-2149083, 2149026, 2149084).

Q1: Please clarify plans for how Tripoli would be developed to address the specific needs of intended users, such as laboratory managers. In particular, how would planned community engagement activities specifically inform development, distribution, and adoption of the software package to meet the needs of intended users, such as with regards to interface design and input/output handling?

R1: Our plan is to engage the community in the design/development process, with a bow to COVID-driven limitations. **First**, we will identify users by a) reaching out to users of the current Tripoli software, used by the TIMS isotope geochemistry and geochronology communities for 20 years with over 250 citations of doi:10.1029/2010GC003479 and 10.1029/2010GC003478; b) advertising for potential users through professional networks and social media; and c) engaging with mass spectrometer companies for suggested users. **Second**, we will set up a Slack account (or as preferred by users) to host and track discussions. **Third**, we will set up an open-source GitHub repository with a website and wiki to host the software and the development process including documentation.

Finally, we will schedule a repeating series of virtual meetings. COVID makes scheduling difficult, especially for those with care-taking responsibilities. Thus, our initial user engagement will be flexible, with approximately ten possible meeting times to facilitate small, focused discussions. These meetings will support brainstorming to determine needed capabilities, features, and support for lab workflows. Subsequent meetings will provide interactive evaluations of the software in development. We intend that this investment of time with lab managers and other power users will translate into a successful agile, iterative development process with continuous prototyping, risk analysis, implementation, testing, and adoption. User needs will also inform technique-specific features to be developed as plugin modules, an approach that has proved successful, for instance, in adapting ET_Redux for the U-Pb and U-Th communities. As the COVID situation changes, we will keep or adapt this approach for a series of future meetings during the middle and toward the end of the project, ensuring users have the opportunity to collaborate and contribute at every step of the software production process.

For input handling, our plan is to develop and publish a standard data schema so that Tripoli need not directly interact with mass spectrometer software packages. All commercially available mass spectrometers produce text files and Tripoli will provide built-in translators to the standard schema as part of its workflow automation. Some mass spectrometer software (e.g., Thermo) also produces binary output files with additional data and metadata, and we will work with the mass spectrometer companies (see letters of support) to extract that information when possible.

The same philosophy governs output – Tripoli will output a schema-compliant file that can be post-processed for use by any other software packages for data reduction and analysis. Tripoli will also provide output files in user-requested formats for technique-specific data reduction and/or lab-specific needs. One example is the XML file used by ET_Redux, and we anticipate a large number of these outputs for various purposes. User workshops at the beginning, middle, and end of the project will identify these, and our training efforts will focus on teaching users how to contribute to the Tripoli codebase and how to create new output formats.

The open-source model of development encourages “release early and often,” meaning that the engaged user community can participate almost in real-time. In particular, because of the modular nature of Tripoli’s architecture, the mathematical solutions and data reduction algorithms – the logical core – are separate from the user interface. Thus, specialized versions of the interface can be developed by others. Initially, control of the repository will fall to PI Bowring. Eventually, per R3 below, an oversight board will control this process. A future benefit of the modular design is that the core logic of Tripoli can be made available as a web service available to programs written in any language. PI Bowring is collaborating on a similar deployment for the software package Squid3 scheduled for release in 2022 that would serve as a blueprint for a similar solution for Tripoli.

Q2: Please clarify plans for how Tripoli would be developed to ensure compatibility with manufacturer specifications, including potential proprietary restrictions that may arise when interfacing with existing instrument software. How would instrument manufacturers be engaged in the software development process?

R2: To be successful, Tripoli must ensure compatibility with a wide range of mass spectrometer data types. All TIMS and MC-ICPMS mass spectrometers presently in use provide a standardized set of text output files whose formats vary by manufacturer but contain the relevant data needed for most calculations. As part of this project, Tripoli will transform these data into interactive visualizations and significantly improve the statistical data handling, all using the currently available open format data files produced and used in all mass spectrometer labs.

There is also significant room for improvement in the mass spectrometer data streams, most notably in interpreting real-time data from mass spectrometers to facilitate data reduction and user decisions during long MC-ICPMS and TIMS analyses. All three major TIMS and MC-ICPMS mass spectrometer companies (Isotopx, Thermo, and Nu, see letters of collaboration from each) have all agreed to work with us to develop data export solutions, based on professional relationships built over a decade of collaboration and/or recognition of successful past efforts. All three major instrument manufacturers are well-integrated with the isotope geochemistry and geochronology communities. Their employees, who often sponsor and actively participate in conferences and workshops, will be included in the design and implementation of Tripoli, and in particular, their software engineers will be consulted regularly as needs are articulated by the community.

Direct involvement of manufacturers' software engineers will preclude barriers from proprietary data formats, as they have direct access to the instrument data streams and can generate new output formats. We know of one case where the instrument manufacturer lost track of the documentation for a proprietary data file type (for the Element2 ICPMS), but a geochronology lab was able to decode it with help and permission from Thermo.

Q3: Please clarify plans for sustainability of the proposed Tripoli open-source software package beyond a potential NSF award. In addition to encouraging individual researchers to incorporate this software package into future research proposals, would any coordinated efforts (e.g., via user groups and/or manufacturing partners) be undertaken to identify roles for maintaining the software package into the future?

R3: Efforts to sustain Tripoli beyond the potential NSF award period will center on leveraging a technically proficient user base and relationships with industry partners. We plan to establish a shared industrial-academic oversight board committed to the sustainability and maintenance of the Tripoli product. Rotating members will be recruited from active user groups and industry partners, with care taken not to favor any particular manufacturer. As discussed in response R2, we have significant manufacturer buy-in, and we will continue to expand and deepen these partnerships. We will foster an active user group and strong knowledge base starting from the existing Tripoli community, as outlined in R1. Online tutorials, virtual workshops, and "hackathons" will further serve to engage future contributors beginning in Year 3 and beyond. The three PIs will sit on the board and work to identify and engage potential rotating members. As relatively early-career investigators, McLean and Burdick will ensure continuity on the board in the medium term. Since about the 1970s, technical developments in isotope geochemistry and geochronology have been driven by successful collaboration between users pushing the scientific envelope and manufacturers creating new cutting-edge instruments. We envision the oversight board furthering this collaboration, for instance by organizing communication of user community needs to instrument manufacturers as they release new software and hardware.

During the award period and then in the following maintenance phase, we will support the use of established open-source mechanisms. The collaborative ecosystem provided by Github "Issues" can be used to manage bugs, features, documentation, and tests, as demonstrated by PI Bowring's recent work developing software with the SHRIMP secondary ionization mass spectrometry community. This process also includes the maintenance of documentation and test data suites. At and after completion, success will be measured both by frequency of user interactions (contributing issues, code, documentation, and/or test data) as detailed in our proposal's "Essential Elements" section and by the promotion of capable contributors to be collaborators with manager-level permissions in the code repository.