Coding the Column

Session 1: Feb 10, 2022 Session 2: April 21, 2022

**Hosts:** Leah LeVay (Session 1), Andy Fraass (Session 2)

**Technical Hosts:** Leah LeVay, Andy Fraass, Katie Jamson

**Speakers (Session 1):** Brian Huber (Mikrotax and database challenges), Johan Renaudie (Neptune Sandbox Berlin, NSB), Isabel Fenton (Triton), Heiko Pälike & Anna Joy Drury (EARTHSEQUENCING), Simon Goring (Neotoma), Mark Uhen (Paleobiology Database, PBDB), Shanan Peters (Macrostrat), Andy Fraass (eODP)

**Speakers (Session 2):** Johan Renaudie (NSB), Shanan Peters (Macrostrat), Jack Williams (Neotoma)

**Authors**: Andrew Fraass1

1 University of Victoria

**Impetus:** There is a need for scientific ocean drilling (SOD) to utilize next generation tools. Current capabilities are under utilized, and there are groups working on projects in neighboring or potentially collaborative spaces. Those groups, however, are not integrated, and do not regularly communicate. There, also, is a dearth of knowledge in the scientific ocean drilling community about what other scientific fields (even within Earth science, e.g., paleontology) have developed in this space.

**First Session [**[**Summary Jamboard**](https://jamboard.google.com/d/1pnBfZc3uyWNxbJ5VjaVAcKHr_-DJwHxgQcEjBjeAIbs/edit?usp=sharing)**]:** The first session was devoted to bringing together a large collection of speakers to discuss different database systems working or developing in the scientific ocean drilling space. The first talk after the introduction was devoted to a semi-historical look at work that has been done in the space over the career of a PI (Huber), challenges that have been found, and then the very successful current project (Nannotax and Mikrotax). The rest of the talks were divided between scientific ocean drilling projects primarily interested in microfossil (NSB, Triton) or high resolution data (EARTHSEQUENCING) from scientific ocean drilling sources. The second group of talks were from the broader Earth science community, representing quaternary deposits (Neotoma), stratigraphic (Macrostrat) and paleobiological data (PBDB), with a SOD focused group using both Macrostrat and PBDB finishing the session (eODP).

**Second Session [**[**Summary Jamboard**](https://jamboard.google.com/d/1hNn6PVT84G7-trXwZnCtBK7OpE0QvysufCoPlVaW3cI/edit?usp=sharing)**]:** The second session focused on the finer aspects of storing chronostratigraphic information. First, talks from NSB, Macrostrat, and Neotoma allowed the audience to get a better grasp on the differing philosophies behind the way these systems store and use information. After that, we had group discussions centered around several questions: “What are the required capabilities of a stratigraphic system?”, “What are Required/Desired/Optional data?”, “What would you trust out of a system, data, interpretations, more?”.

**Discussions:**

* Funding is an issue. There is a need to build and fund a system that reflects the international character of the IODP community.
* There are multiple tiers of users: researchers who only want an age model, those who will use and tweak an age model, and those who will want to contribute and refine within the system.
* There are limitations to the existing systems and a lack of useful user interfaces.
* Depth is the least interpreted reference frame, though even it is an inferred property in marine systems. It is useful to have the model clearly indicated.
* An individual Hole, and depth within it, should be the single reference frame. This helps avoid the issue of splicing, which is more complicated, though would be extremely useful to tackle.
* Required data: Legacy data can have less meta-data associated with it (e.g., less common to report sampling resolution). Modern data should have this data required as a part of the data collection process, but legacy data should still be incorporated as to not do so would lose a great deal of data.

**Key takeaway points:**

* Depth scales and splices remain a tricky issue within many systems.