MARS: A MIDDLEWARE THAT ASSISTS IN THE REGISTRATION OF PHYSICAL SAMPLE METADATA WITH SESAR

Julius Walton[1], Josh Gilley[1], Robert Niggebrugge[1], Aerin Parker[1], James F. Bowring[1], Kerstin Lehnert[2], Lulin Song[2], Sarah Ramdeen[2], Alexandra Hangsterfer[3]
[1]Computer Science, College of Charleston ,Charleston, SC; [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY; [3]Scripps Institution of Oceanography, La Jolla, CA

ESIP FUNDING FRIDAY

The MARS group was awarded a FUNding Friday Mini Grant at the 2020 ESIP Summer Meeting. The funds are being used to support the development of the MARS Map Maker tool which will broaden the functionality of MARS beyond the marine and lake core community. The MARS team will recruit testers from the ESIP community as it moves into production.

Background

It has been recognized that geoscience collections and the data they generate are national resources that can be relatively inexpensive to fund, but that access to such collections and data is not reliable (*National Research Council 2002*). This is mainly due to the lack of consistency across collections with regards to best practices when digitizing samples. NOAA'S National Centers for Environmental Information (NCEI) Index to Marine and Lacustrine Geological Samples (IMLGS) supports data from an international community of about 30 seafloor and lakebed sample collections and provides long-term archival of metadata on a database that uses a common vocabulary. This is one main community resource that helps to standardize sample data. Besides the need for a maintained and supported database, the need for globally unique and persistent identifiers for physical samples has been recognized as a best practice that ensures samples can be tracked through their life cycle and samples can be linked with data and publications. The System for Earth Sample Registration (SESAR) is the primary US Allocating Agent for such identifiers called IGSN Global Sample Number.

Communication Breakdown

Currently, the process of obtaining IGSNs requires collections managers to use SESAR's batch upload template. This template uses a standardized metadata profile, that is not an exact match to the metadata fields used by collection managers or to the metadata fields required by IMLGS. It is a substantial burden to have to reformat sample data numerous times to meet the metadata and vocabulary requirements of different data management tools.

The Solution: MARS
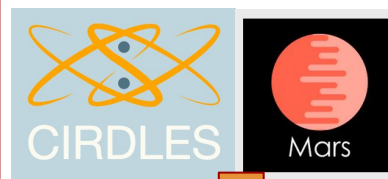

```
let map = {
    name: "SAMPLE",
    igsn: "IGSN",
```

```
latitude: "LAT",
longitude: "LON",
latitude_end: "END_LATMIN",
longitude_end: "END_LON",
elevation: "WATER_DEPTH",
elevation_end: "END_WATER_DEPTH",
primary_location_type: "PROVINCE",
cruise_field_prgrm: "CRUISE",
platform_name: "PLATFORM",
```

IGSN:	SIO0000PC
Sample Name:	M144_3-8
Other Name(s):	
Sample Type:	Core
Parent IGSN:	Not Provided

What is MARS?

MARS stands for **M**iddleware for **A**ssisting the **R**egistration of **S**amples. It is a tool that helps map physical sample metadata to the metadata required by SESAR. MARS was developed as a result of a collaboration among The College of Charleston (CoC), Scripps Institution of Oceanography and SESAR. MARS was initially developed to map Scripps' metadata fields, which are the same metadata fields used by NCEI's IMLGS, to SESAR's batch template. Scripps will use MARS to help seamlessly assign IGSNs to the thousands of sediment cores in its large archive collection.

How was MARS created?

MARS was created by Dr. Jim Bowring and his team of undergraduate students at CoC. They worked with SESAR to accurately map the Scripps metadata fields to SESAR and then developed a web-based service that allows Scripps to upload CSV metadata files to SESAR using the map - written in Javascript - to translate the metadata correctly. MARS connects to SESAR to quickly return IGSNs.

Interested in MARS?

Contact Dr. Jim Bowring at The College of Charleston
Email address: *bowringj@cofc.edu*

MapMaker & CIRDLES
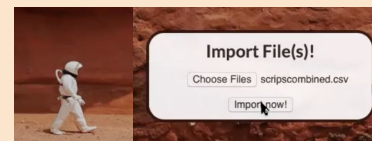
MARS MapMaker is a newly developed tool that will work to map physical sample metadata in *any* format to the format required by SESAR to register IGSNs. This tool will be available to an international group of people who manage institutional and/or personal sample collections.

MARS and MARS MapMaker are a part of CIRDLES (*Cyber Infrastructure Research & Development Lab for the Earth Sciences*), an undergraduate computer science software engineering initiative at The College of Charleston.

Informational Webinar

Join us to learn more about MARS!
ESIP will be hosting a webinar to introduce potential users to MARS, and to recruit testers for the MARS MapMaker.
Please **email: ahangsterfer@ucsd.edu** to be notified about how to join.

Reference National Research Council, 2002. *Geoscience Data and Collections: National Resources in Peril.* Washington, DC: The National Academics Press. https://doi.org/10.17226/10348.