FIRST STEPS TOWARD A TERRESTRIAL ANALOG DATA PORTAL. L. P. Keszthelyi, M. A. Hunter, T. M. Hare. U.S. Geological Survey Astrogeology Science Center, 2255 N. Gemini Dr., Flagstaff, AZ 86001 (laz@usgs.gov).

The Problem: The many NASA-funded terrestrial analog studies have produced immense volumes of data. The exact amount is unknown but is guaranteed to far exceed the few Petabytes of planetary mission data in the PDS. A series of recent directives from the Federal Government now require that all of this data be made accessible to the U.S. taxpayers in open and machine readable formats [1-5]. NASA grants to academic and private institutions require adherence to these regulations as part of the now mandatory data management plans. These legal requirements not only apply to work done subsequent to their enactment – they apply retroactively to all Federally funded research. Clearly there is no plausible way to come into compliance with these regulations overnight. However, the enforcement of these rules is increasing and it is essential that rapid progress be made in providing researchers with practical tools to meet these regulations.

Currently, there is no effective infrastructure to support researchers who wish to do the right thing but many are trying to create their own data delivery systems. Therefore the majority of the data lies hidden in researchers' offices or, rarely, on hard-to-find websites [6]. International consortiums are making rapid progress in creating standards and community resources for specific types of data. Online databases for petrology and planktonic foraminifera are two successful examples of these types of focused community services [7,8].

Unfortunately, these narrow discipline-based databases do not serve the NASA planetary analog community well. One of the hallmarks of planetary science is its interdisciplinary nature and planetary analog work shares this characteristic. Rock samples, geophysical data, satellite images, weather data, ground-based and aerial spectroscopy and imaging, topography, and more are all vitally relevant. Astrobiology adds requirements for biologic and ecological data sets. Developing the infrastructure to house and deliver such a diversity of data would seem to be a Herculean, if not Sisyphean, effort.

Functional Requirements: The scope of the problem is perhaps best summarized as an (unprioritized) list of high-level requirements for Federally-funded planetary analog data management.

1. Data includes diverse physical and digital materials
2. Data entry into the system must be easy
3. Data must be preserved indefinitely
4. Data must be readily discoverable (by humans and by machines)
5. Data must be readily accessible (open to all)
6. Data must be easy to use (in current formats)

A Solution? Fortunately, there is another organization facing a very similar problem of data archival and accessibility: the United States Geological Survey. It is not just that misery loves company; the USGS is already well on the way to creating a coordinated set of repositories to house data including formal publications; time series data on everything from hydrology to seismology; digital records of air, rock, water, and biologic samples; the full array of aerial and orbital remote sensing; topography derived from a variety of techniques; and even physical samples or rock and ice. As the world's largest Earth science organization, the USGS has the resources to produce an effective and cohesive data management system. And it is must abide by the same Federal strictures that NASA-funded projects must.

For FY19, NASA directed the USGS Astrogeology Science Center to conduct an exploratory project to determine if there was a practical way to leverage the USGS investment in data management for use by the NASA planetary analog community. When this project was conceived in early 2018, the idea was to focus on the National Map developed and maintained by the USGS [9]. However, this idea was overtaken by a new effort within the USGS – the development of the USGS ScienceBase [10].

ScienceBase = A Solution! We find that the USGS ScienceBase Catalog will meet all the legal and functional requirements for NASA terrestrial analog data management. The diversity of USGS digital and physical data types encompasses the full range of data types used in terrestrial analog work. The system is being designed to allow research scientists to upload their own data with little effort. The data will be preserved by USGS Core Science Systems (CSS) for as long as the Federal government continues to exist. In addition to satisfying all of the core technical requirements we identified for NASA-funded work, ScienceBase also links users to troves of additional potentially useful data that are being generated and maintained by science centers across the USGS, as well as many state agencies.

ScienceBase is built on open standards for metadata (i.e., Dublin Core model) and supports sharing geospatial data via Open Geospatial Consortium (OCG) standards, in addition to common commercial formats, and can link to data hosted by other institutions. ScienceBase items utilize persistent and unique URLs, and are accessible through a well-documented application programming interface (API). Current developments by the ScienceBase team promise to enhance this core functionality to optimize front-end applications, like as this portal, with custom extensions. Examples of groups

already leveraging ScienceBase for their own data repository and discovery needs can be found at https://sciencebase.gov/about/case-studies.

We have reached the conclusion that the USGS ScienceBase Catalog would be a highly cost-effective and functional solution for NASA to utilize for making planetary analog data compliant with Federal law and far more useful to the community than it is today.

Next Steps: Whether and how NASA leverages the USGS ScienceBase Catalog and associated infrastructure is not yet decided. However, we have developed a proposed path that moves at a fast but realistic pace. The outline of this plan is enumerated below:

1. Test bulk ingestion of metadata records; conduct knowledge inventory of analog sites (FY19-20)
2. Demonstrate a portal that presents a mix of planetary analog data; coordinate with CSS on needs (FY20-21)
3. Disseminate procedures for validating and contributing data from outside USGS (FY22 onward)
4. Expand the portal with community input (FY23 onward)

References: [1] OMB (2009) Open Government Directive M10-06. [2] White House (2013) Executive Order 13642. [3] OMB (2013) Memorandum M-13-13. [4] OMB (2016) Circular A-130. [5] OSTP (2013) Memorandum. [6] U.S. (2014) Open Data Action Plan. [7] PetDB. [8] Siccha, M. and M. Kucera (2017) Scientific Data, 4, 170109. [9] The National Map. [10] ScienceBase Catalog.