Towards a Terrestrial Analogs Data Portal: Use Cases and Requirements

David M.H Baker[1], Mary (Elise) Rumpf[2], Marc A. Hunter[2], Laszlo P. Keszthelyi[2], Jacob A. Richardson[1,3], Patrick L. Whelley[1,3], Nicholas M. Schmerr[3], Kelsey E. Young[1].
[1]NASA Goddard Space Flight Center, Greenbelt, MD 20771 (david.m.hollibaughbaker@nasa.gov), [2]U.S. Geological Survey Astrogeology Science Center, Flagstaff, AZ 86001, [3]Univ. of Maryland, College Park, MD 20742

Problem: No dedicated repository for terrestrial analog field data

A recent survey conducted by the USGS Astrogeology Science Center [1] and the final report of the Planetary Data Ecosystem (PDE) Independent Review Board (IRB) [2] emphasize a need to establish and maintain the ability to serve, preserve, and link to the diverse and extensive terrestrial analog field data produced by NASA programs.

References:
[1] Rumpf, M.E. et al. (2019) *AGU Fall Meeting* 2019. Abstract #EP21E-2211.
[2] NASA (2021) Planetary Data Ecosystem Independent Review Board Final Report, https://science.nasa.gov/researchers/science-data.
[3] Keszthelyi, L. et al. (2019) *4th Planetary Data Workshop* 2019. Abstract #7023.
[4] ScienceBase Wiki, https://my.usgs.gov/confluence/display/sciencebase/ScienceBase, Accessed April 2021.

Many researchers have limited access to terrestrial analog field data

Much field data are with individuals or in disparate locations and may be lacking sufficient documentation or metadata.

Community desires a place to share and access field data relevant to planetary science

USGS Astrogeology conducted a survey of the planetary analog community in 2019 [1]:

- Only 14% of respondents had placed their analog samples or data in a long-term publicly accessible archive .
- 92% of scientists interested in analog studies used data and results from other researchers.
- However, 77% reported difficulties in discovering this data.

A Solution: Terrestrial Analogs Data Portal (TADP)

https://www.sciencebase.gov/catalog/?community=Terrestrial+Analogs+for+Planetary+Science
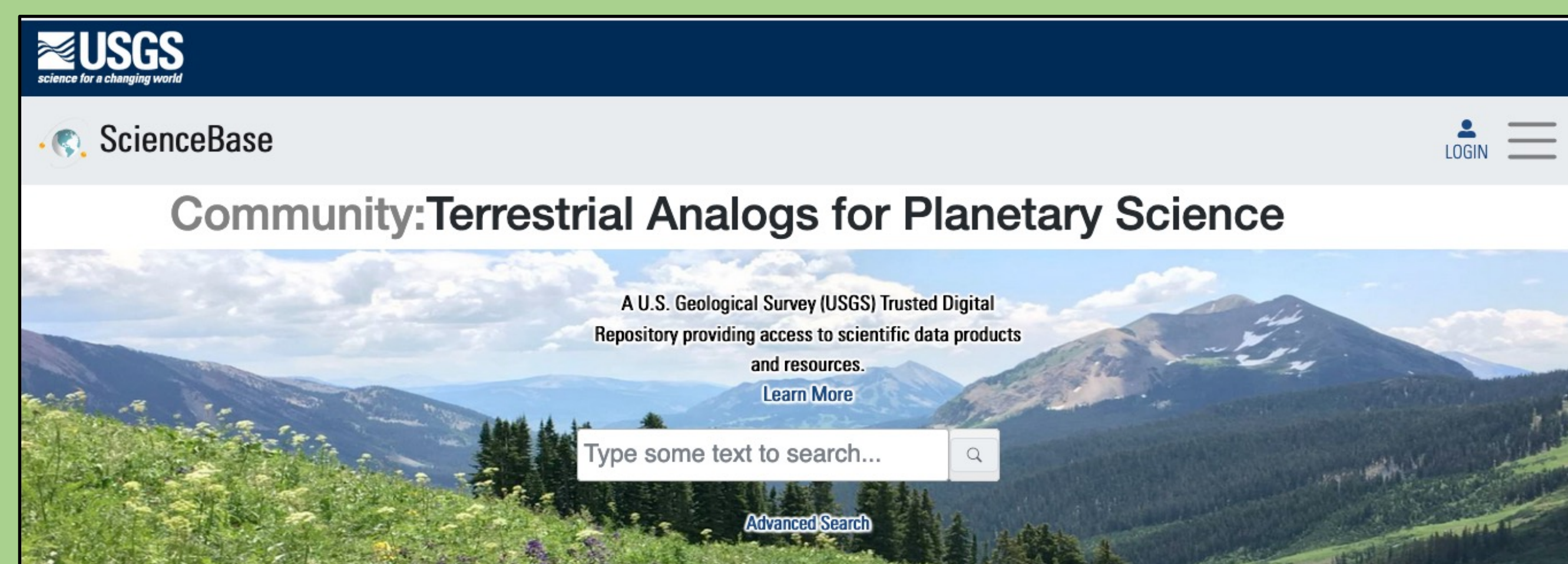
Key Features [ref 3,4]:
☐ Founded on USGS ScienceBase Catalog
- Built on Dublin Core Metadata Element
☐ Shares geospatial data via Open Geospatial Consortium (OGC) standards and commercial formats
☐ Long-term preservation provided by the USGS Core Science Systems; meets NASA legal/functional reqts.
☐ Custom extensions (e.g., ArcGIS REST Service).
☐ USGS Trusted Digital Repository
☐ Persistent and unique URLs

Repository Requirements: Data Types

- Digital data from NASA-funded terrestrial analog field work.
- Accommodate a diversity of data types, formats and volumes obtained from various field instruments and generated through various post-processing.
- Documentation and metadata (from field notes, post-processing) are critical for data useability.

Repository Requirements: Features

- Metadata standards
- Common dictionaries
- Geospatial search and discovery
- Linkages to external repositories, lab data, samples and wider Planetary Data Ecosystem
- Low barrier to archiving
- Long term preservation; DOIs

A field analog community group is essential to develop and promote best practices for metadata, standards, data formats, and documentation.

Use Case: NASA GIFT

Leads: Kelsey Young and Amy McAdam, NASA GSFC
- Field campaigns to multiple sites: Hawaii, Iceland, Medicine Lake, and future sites.
- Field data types: LiDAR, GPR, hXRF, UAS orthoimages and DTMs, dGPS, bore holes, photos, field notes.
- Field samples and post-field lab analysis.

TADP Desire: Hosting and serving of data (no adequate local repository) for public accessibility.
- Deposit data from multiple field campaigns and investigators; some over several years and into future.
- Currently working to prepare Iceland field data for depositing in TADP and providing public access.

General Data Work Flow: Getting Data in TADP

A. Team collects data in field and post-processed.
B. Post-field, team captures critical field metadata.
C. Team archiving manager facilitates creating directory structure, organization and naming of data, and ReadMe templates on internal shared team drive. Stages data for later public repository.
   i. Structured by field site/subsites, then data type.
D. PI completes field site level ReadMe capturing essential metatdata and post-processing steps
E. TDAP prep: Use USGS metadata wizard to format and validate metadata of collection to repository standards.
F. USGS and internal review of data collection.
G. Data collection publicly released on TADP under Project Team. Link to relevant publications.

Use Case: SSERVI GEODES

Leads: Nicholas Schmerr (Univ. Maryland) and Jacob Richardson, (Univ. Maryland/ NASA GSFC)
- Multiple field sites: San Francisco Volcanic Field, Lava Beds National Monument.
- Field data types: Seismic, Gravity, GPR, Magnetic, LiDAR, dGPS, hXRF, IR, photos, field notes.
- Post-processed data products in a variety of formats.

TADP Desire: Search and discovery of data hosted at separate institutional data repository.
- Preference to deposit data more locally at the University of Maryland DRUM repository.
- Working to enable discovery of DRUM data within TADP to enable greater discovery among the public and with other field analog data.