ASSESSING PLANETARY DATA ACCESS AND USE. R. M. Borden[1] and B. W. Bishop[1], [1]School of Information Sciences, University of Tennessee, Knoxville, TN 37996 (rborden4@vols.utk.edu).

Introduction: Planetary scientists often rely on and reuse data from past and ongoing missions for their current research. This data comes from instruments on orbital spacecraft, landers, and samples brought back from lunar missions. There are several places to find this data. Data from any NASA mission will be in one of the Planetary Data System's (PDS) Nodes. The PDS is the formal archive for NASA mission data, as well as derived products from NASA-funded planetary research [1]. There are six science discipline Nodes, including Atmospheres, Geosciences, Cartography and Imaging Sciences, Planetary Plasma Interactions, Ring-Moon Systems, and Small Bodies. Each of these nodes may have one or more different search interfaces to help scientists and other data users find the data they are looking for within the node or in other nodes.

With multiple places to find data and ways to find it, this project was developed to explore which places planetary scientists most often use when finding data for their research, and their experiences accessing and using it. When planetary scientists look for data to use for their work, where, how, and why do they go to those resources? The aim of this project was to explore these questions through a survey of current users of planetary data.

The FAIR Data Principles of findability, accessibility, interoperability, and reusability have been gaining momentum as a way to assess data and data repositories in recent years with the original publication now cited over 1,000 times [2]. The utility, versatility, and charm of the FAIR acronym help explain its popularity and application in a variety of fields including biology, life science, plant science, environmental science, and other data-intensive sciences [3-6]. The original goal of these data principles was to promote machine-actionable data and metadata because of human limitations in data processing and also to allow for humans to spend more time doing original research while automating as much as possible [2]. Still, understanding human information-seeking behavior makes data access and use better for humans as well as machines.

After being submitted to the PDS, data should not just sit around forever. The importance of storing the data in a publicly accessible archive like the PDS is precisely so it can be available for reuse by scientists for future research. The results of studies like this one can be used to find the best way to make this data more easily accessible and usable by all potential future users.

Methods: A survey was developed to assess planetary scientists' experiences finding, accessing, and using data. While NASA's Planetary Data System (PDS) is a common place to look for planetary data, the language of the survey was left intentionally neutral to allow for scientists who might be using data from other agencies or archives. The survey was based on a set of interview questions that have been used to interview scientists dealing with Earth-focused data [7]. These questions have been modified to be relevant to planetary data uses. The survey questions are formatted to match up with the FAIR data principles of findability, accessibility, interoperability, and reusability.

The following questions were asked in the survey:

Job-related demographics

1. *What is your current job title?*
2. *How many years have you been working in your current job?*
3. *How many years have you been working with planetary spatial data?*
4. *What is your highest level of education?*

Findability

5. *What tools or websites did you use to help you find the data? (PDS Data Search, a specific PDS node search tool (which one?), ESA's Planetary Science Archive search tool, USGS PILOT, Map-a-Planet 2, other)*
6. *Did the data have a persistent identifier (i.e., a long-lasting unique reference to an objects location) (e.g., DOI; PURL)?*
7. *Did the data have metadata?*
8. *Did you use the metadata to help you find the data?*

Accessibility

9. *How did you access the data? (Download, process before downloading (i.e., USGS Map Projection on the Web service), view only without downloading, other)*
10. *If downloaded, what format were the data in?*
11. *Was the data free?*
12. *Did the data have use constraints (e.g., limitations of use)?*
13. *Was the metadata accessible?*

Interoperability

14. *Was the data in a useable format at time of download/access?*

15. *How was the data encoded?*
16. *Was the data using encoding common to other data used in your research (i.e., same format)?*
17. *Was the data machine-actionable (e.g., able to be processed without humans)?*

Reusability

18. *Did the data geographic scale or resolutions impact use of the data?*
19. *Did the coordinate systems used impact use of the data?*
20. *Did the metadata provide sufficient information for data use?*
21. *Were there any other issues with the data that impacted use of the data?*

Closing

22. *Please provide any other feedback about planetary spatial data fitness for use.*

These questions were entered into the survey software QuestionPro. The survey was publicized with announcements in the Lunar and Planetary Institute's Planetary News e-mail newsletter, the Planetary Science Institute's Planetary Exploration Newsletter, and on the Young Scientists for Planetary Exploration Facebook group.

Preliminary Findings: There have been 39 responses to the survey at the time this abstract is being written. About half of the participants in the survey have found their data via the PDS search interface or a search tool hosted by one of the PDS discipline nodes, such as the Geosciences Node Orbital Data Explorer. Another quarter of the responses indicated use of the PILOT or Map-a-Planet 2 tools hosted by the USGS Astrogeology Science Center or the ESA Planetary Science Archive search tool. The remaining responses indicated they used some other way of finding the data. These other ways included things such as individual instrument or mission websites, Google/Google Mars, and JMARS.

Another question asks data users about what they do with data once they find it. Of the responses so far, ~80% of users download data as they find it, and 10% each either process before download or use it online without downloading. However, 25% of respondents said that the data was not in a usable format at the time of download. Several other disciplines have data portals that automate transposing data at the time of download into common usable formats and this is functionality that could be built into these systems. 70% of respondents said the geographic scale or image resolution impacted use, and 51% said the same about the coordinate system(s) used. 13% said the metadata did not provide sufficient information for data use.

There were a number of text responses to the question asking about issues with the data impacting its use, which can be used to make recommendations for future PDS format/interface update considerations.

Some questions could clearly have been worded better, such as the one asking about "encoding" of the data, which confused many respondents who said they didn't know what the question was asking.

Future Work: Once the survey is closed, all results will be downloaded and analyzed. The full results of this survey will be published in a peer-reviewed journal. This survey is a first attempt to understand the ways in which potential users of planetary data discover and evaluate the data they wish to use. Future surveys could focus on specific PDS Nodes or search tools, or on certain aspects of the user experience.

References: [1] PDS Roadmap Study for 2017 – 2026. [2] (1) Wilkinson, M. et al. (2016). *Scientific Data*, 3, 160018. [3] (2) Diepenbroek, M. et al. (2017). *Journal of Biotechnology, 261*(Supplement C), 177-186. [4] (3) Rodríguez-Iglesias, A. et al. (2016). *Frontiers in Plant Science, 7*, 641. [5] (4) Wilkinson, M. D. et al. (2017). *PeerJ Computer Science, 3*, e110. [6] (5) Wolstencroft, K. et al. (2017). *Nucleic Acids Research, 45*(D1), D404-D407. [7] Bishop, B. W. & Hank, C. F. (2018). *International Journal of Digital Curation, 13*(1).