

## ARMS: A Developing Metadata Standard for Describing Astrobiology Research Products

Richard M. Keller<sup>1</sup>, David F. Blake<sup>1</sup>, Thomas F. Bristow<sup>1</sup>, George W. Cooper<sup>1</sup>, Christopher E. Dateo<sup>1</sup>, David J. Des Marais<sup>1</sup>, Angela M. Detweiler<sup>2,1</sup>, Linda L. Jahnke<sup>1</sup>, Michael D. Kubo<sup>1</sup>, Barbara Lafuente<sup>3,1</sup>, Mary N. Parenteau<sup>1</sup>, Leslie E. Prufert-Bebout<sup>1</sup>, and Nate Stone<sup>4</sup>

<sup>1</sup>NASA Ames Research Center, Moffett Field, CA USA rich.keller@nasa.gov

<sup>2</sup>Bay Area Environmental Research Institute, Moffett Field, CA USA

<sup>3</sup>SETI Institute, Mountain View, CA USA <sup>4</sup>Open Data Repository, Gray, ME USA

**Introduction:** Internet-based search has provided us with a means of discovering and accessing an incredible array of global knowledge resources. But when it comes to finding astrobiology-specific resources, such as datasets or publications, text-based web search is not very precise or effective. A search for “astrobiology dataset” may yield hundreds of irrelevant web sites that do not describe datasets, while at the same time omit relevant sites that exclude the terms ‘astrobiology’ or ‘dataset’. What is missing is a way to annotate resources with more specific information that facilitates its discovery: astrobiology-specific metadata.

**Describing Astrobiology ‘Resources’:** The Astrobiology Resource Metadata Standard (ARMS) is a new metadata standard under development at NASA Ames Research Center, in conjunction with the Astrobiology Habitable Environments Database project (Detweiler et al., 2019). ARMS is intended to uniformly describe astrobiology ‘resources’, i.e. virtually any product of astrobiology research – including datasets, physical samples, software (modeling codes and scripts), publications, websites, images, video, presentations, etc. The current draft of ARMS defines 16 different metadata properties used to describe a given resource. A number of these properties are fairly generic, and cover aspects such as resource identification, personnel, funding, and publications. The true power in ARMS for search and discovery will come from four astrobiology-specific pieces of metadata used to annotate a resource:

- *Field location:* The field site placename or geographic coordinates associated with the resource (for field campaigns or missions);

- *Research theme:* The broad research area most relevant to the resource (as identified in the 2015 NASA Astrobiology Strategy Document (Hays et al., 2015));
- *Astrobiology disciplines:* The set of disciplines most relevant to creation or use of this resource;
- *Astrobiology keywords:* The set of topical keywords that best characterizes the resource. (A structured dictionary of ~800 keywords was developed by Ames scientists for this purpose.)

Table 1 illustrates how the 16 proposed ARMS metadata fields would be used to describe the Pavilion Lake Research Project’s (Lim et al., 2009) physical sample dataset. (Note that the exact formatting rules for the metadata fields shown in the table are still under development.)

**Status:** A draft of the ARMS standard is being prepared for circulation to the astrobiology community. The release of this draft is the first step in a community feedback cycle that will result in the release of an ARMS v1.0 standard.

**Summary and Future:** We have described a developing new metadata standard for astrobiology resources. Ultimately, we envision that a technology solution will enable precise, internet-wide search and discovery of distributed, annotated astrobiology resources. ARMS takes one step along this path.

### References:

- Detweiler A.M. et al. (2019) *AbSciCon*, submitted.  
 Hays L. et al. (2015) *Planetary Science Vision 2050 Workshop*, LPI Contribution No. 1989, id.8141.  
 Lim D. S. et al. (2009) *Fundamental and Applied Limnology* 173(4), 329-351.

**Table 1: Sample ARMS Metadata for Pavilion Lake Research Project dataset**

<b>Resource Name</b>	PLRP integrated dataset	<b>Initial Release Date</b>	7/23/18
<b>Resource Description</b>	Data from 2005-14 field excursions...	<b>Last Revision Date</b>	n/a
<b>Resource Type</b>	dataset	<b>Funding Sources</b>	ASTEP   CSA   NSERC   NGS
<b>Resource Locator</b>	https://odr.io/plrp	<b>Missions / Projects</b>	Pavilion Lake Research Project
<b>Lead PIs</b>	Lim	<b>Field Location</b>	Pavilion Lake, BC Canada
<b>Institutions</b>	NASA Ames	<b>Research Theme</b>	Habitability and biosignatures
<b>Points of Contact</b>	Keller RM   Lim DS	<b>Astrobiology Keywords</b>	paleoclimate   stromatolite
<b>References</b>	10.1127/1863-9135/2009/0173-0329	<b>Astrobiology Disciplines</b>	limnology