

Metadata Schema for UAB Research Data Catalog

<https://digitalcommons.library.uab.edu/datasets/>

Version 1.0

7/15/2025

Claire Warner

Notes

1. Fields below are populated in the batch upload spreadsheet by the data processing code in GitHub unless noted.
2. * Indicates a field is made searchable in Digital Commons

Digital Commons Field	Spreadsheet Column Header	Description	Formatting Notes
Title*	title	Title of dataset.	No HTML formatting needed.
Author ORCID*	orcid	Names (First Middle Last) followed by hyperlinked ORCID. Repeated for every author with an ORCID	HTML formatting: <p>Author Name 10000-0123-4567-8910</p><p>Author Name 20000-0123-4567-8910</p>

Publication Date*	publication_date	Date of publication.	YYYY-MM-DD
Abstract*	abstract	Abstract or description of dataset.	HTML formatting: <p>Abstract here.</p>
Keywords*	keywords	Keywords related to dataset.	No HTML formatting needed. Include as a comma-separated list.
Disciplines*	disciplines	One or more disciplines from the Digital Commons Three-Tiered Taxonomy of Academic Disciplines (https://digitalcommons.elsevier.com/en_US/dc-disciplines-taxonomy). This is a built-in field specific to Digital Commons, and can be filled in during manual curation, as it won't be populated by information from the repository metadata.	No HTML formatting. Separate tiers with colons, and disciplines with semicolons. Ex: Life Sciences: Genetics and Genomics: Computational Biology; Life Sciences: Immunology and Infectious Disease
Related Items	related_data	Links to any DOIs listed as related items. Relation Type is filled in based on the relation types in the DataCite metadata schema (https://datacite-schema.readthedocs.io/en/4.5/appendices/appendix-1/relationType/) in the processing code.	HTML formatting: <p>Relation Type 1: DOI 1 prefix + suffix</p><p>Relation Type 2: DOI 2 prefix + suffix</p>
Source Publication	source_publication	Field for the primary publication the dataset is based off of. Currently unused as all related publications including the source publication are put in Related Datasets, but left as an option for future use.	N/A

Repository*	external_rep	Name of the repository where the dataset is stored.	HTML formatting: <p>Repository Name</p>
Distribution License	distribution_license	License under which the data is shared. This uses a built-in Digital Commons field which accepts a URL for the license (controlled value). Talk to your Digital Commons rep. for a complete list of the allowed values.	No HTML formatting. Add as a URL using “http” not “https”. Ex: http://creativecommons.org/licenses/by/4.0/
Access Instructions	access_link	States whether the data is available openly and what license it is shared under.	HTML formatting. Ex: <p>This data is available under the CC-BY 4.0 License</p> or <p>Access to this data is restricted.</p>
Funder*	fundref	Includes the following if available: funder name, funder DOI, grant URI, grant number. Repeated for multiple funding sources.	HTML formatting ex: <p>Funder: National Science Foundation Funder DOI: 10.13039/100000001 </p>Collaborative Research: The Chemical Ecology of Shallow-water Marine Macroalgae and Invertebrates on the Antarctic Peninsula 1341339</p>
Link to Dataset	source_fulltext_url	DOI URL of the dataset	No HTML formatting.
Author*	author#_fname author#_mname	Author # First Name Author # Middle name or initial	No HTML formatting.

	author#_lname author#_institution	Author # Last Name Author # Institutional affiliation (takes only one affiliation) These 4 columns of the upload spreadsheet feed into the default Digital Commons author field (along with a few other fields such as email, which we leave blank). The columns are generated programmatically for as many authors as are present on the dataset. (# = 1, 2, 3, and so on).	
--	--------------------------------------	--	--