GES DISC Datalist Improves Earth Science Data Discoverability

IN21B-0039

A. Li[1], W. Teng[1,2], M. Hegde[1,2], M. Petrenko[1,2], S. Shen[1,3], C. Shie[1,4], Z. Liu[1,3], T. Hearty[1,5], K. Bryant[1,2] B. Vollmer[1], D. Meyer[1]

Abstract

At the AGU 2016 Fall Meeting, the Goddard Earth Sciences Data and Information Services Center (GES DISC) unveiled a novel way to access data, named "datalist." A datalist is a collection of predefined data variables from one or more archived data sets, curated by a subject matter expert (SME). For that initial unveiling, the GES DISC science support team curated a predefined *Hurricane* datalist, and received very positive feedback from the user community. Datalist uses the same architecture that our new Web site uses, and has the same look and feel as that of the other data sets on our Web site. Datalist provides one-stop shopping for data, metadata, citation, documentation, visualization, and other available services. Since AGU 2016, we have been developing new datalists corresponding to the Big Earth Data Initiative (BEDI) Societal Benefit Areas and A-Train data. We have prototyped four datalists: *Hurricane, Wind Energy, Greenhouse Gas,* and *A-Train*. We have also started working with our external advisory group members to create their favorite datalists, and with other DAACs to explore the possibility of federated (cross-DAAC) datalists.
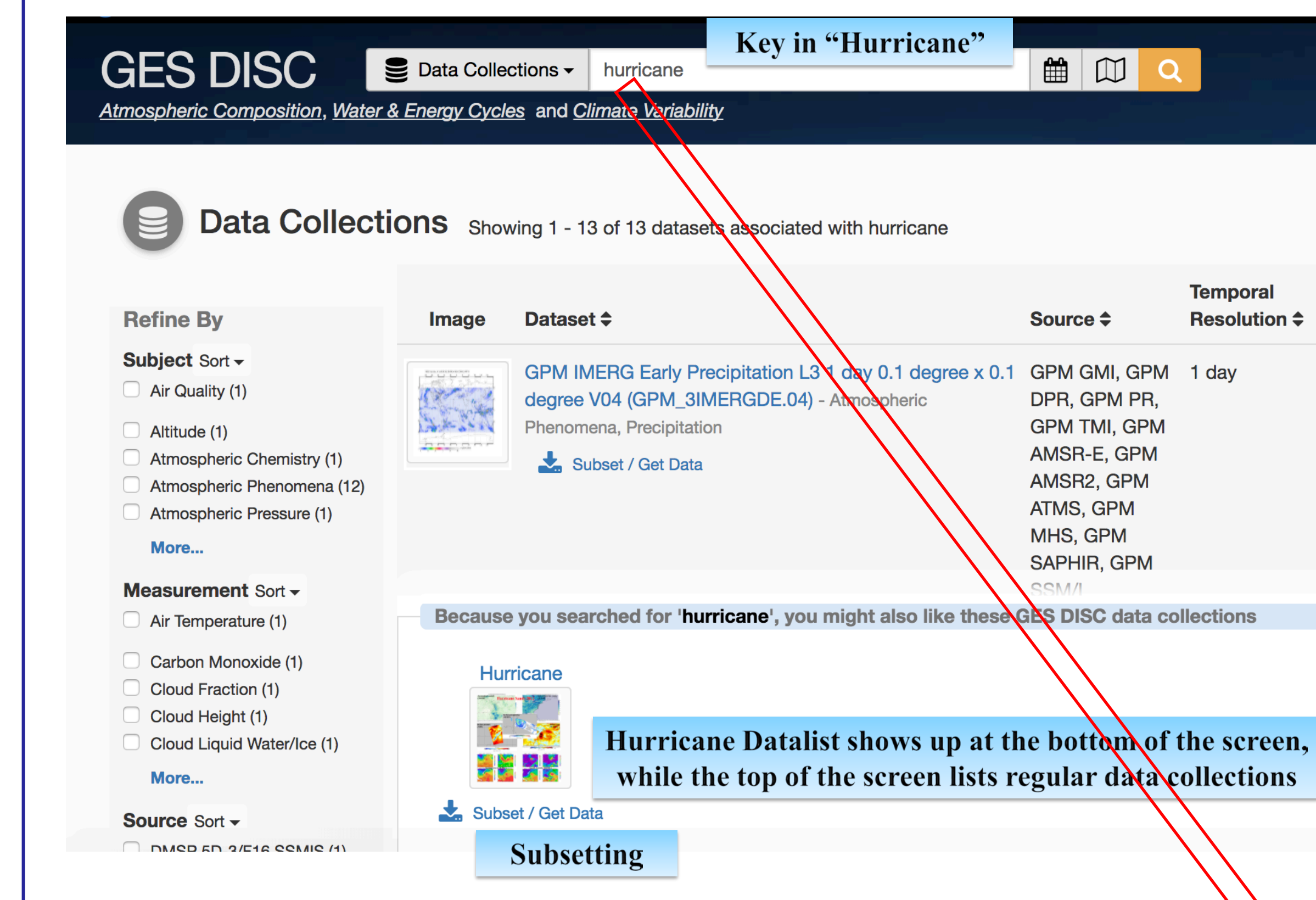
In order to make datalists operational or federated, it is critical to have a common supporting metadata model. To this end, we have mapped our datalist metadata model to the then unpublished UMM (Universal Metadata Model)-Var (Variable) (June version) and found that UMM-Var, together with UMM-C (Collection), and, possibly, UMM-S (Service), will likely satisfy our basic requirements. Data set shortname and version are specified in UMM-C; variable name, long name, units, and dimensions are specified in UMM-Var. UMM-Var also facilitates the use of ScienceKeywords, to allow tagging at the variable level, and of Characteristics, for optional variable characteristics. Measurements is useful for grouping of variables, and Set is promising for defining datalists. The UMM-Service model could be used to specify available services for datalist variables. Together, UMM-Var, UMM-C, and UMM-S would seem to form the basis for federated datalists. Conversely, the development and deployment of datalists would contribute to the evolution of UMM.
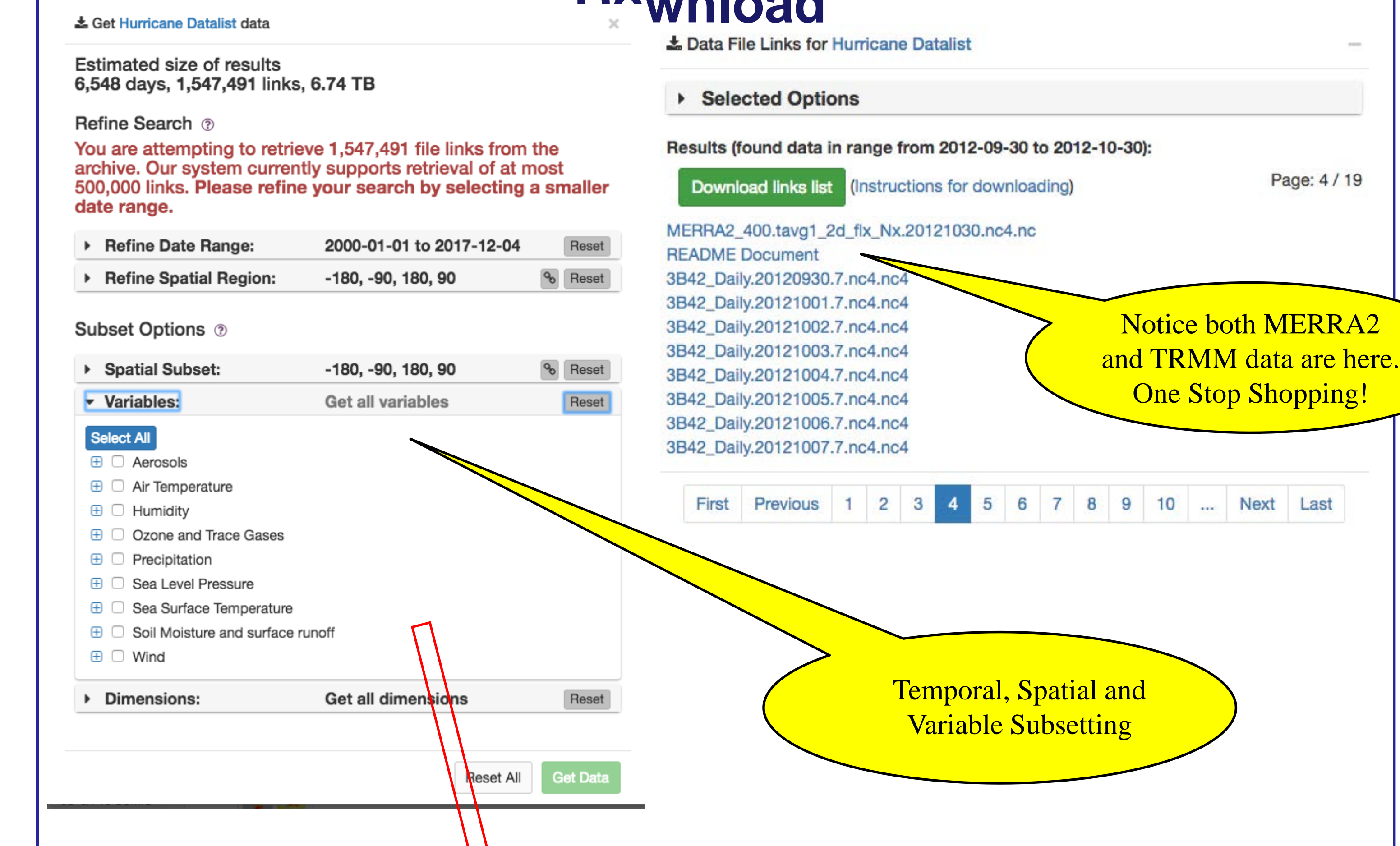
Authors

[1] NASA Goddard Space Flight Center
[2] ADNET Systems Inc.
[3] George Mason University
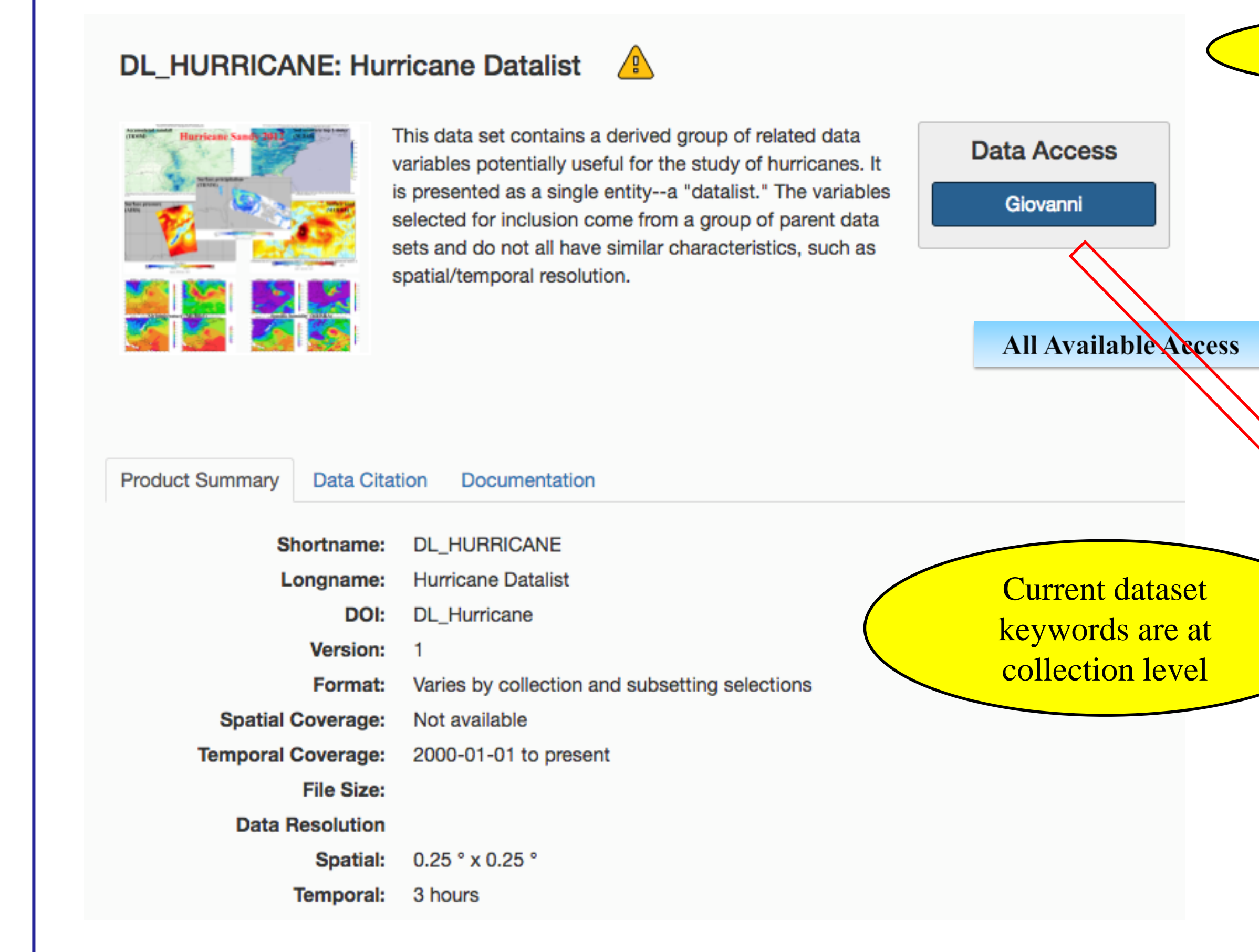[4] University of Maryland, Baltimore County
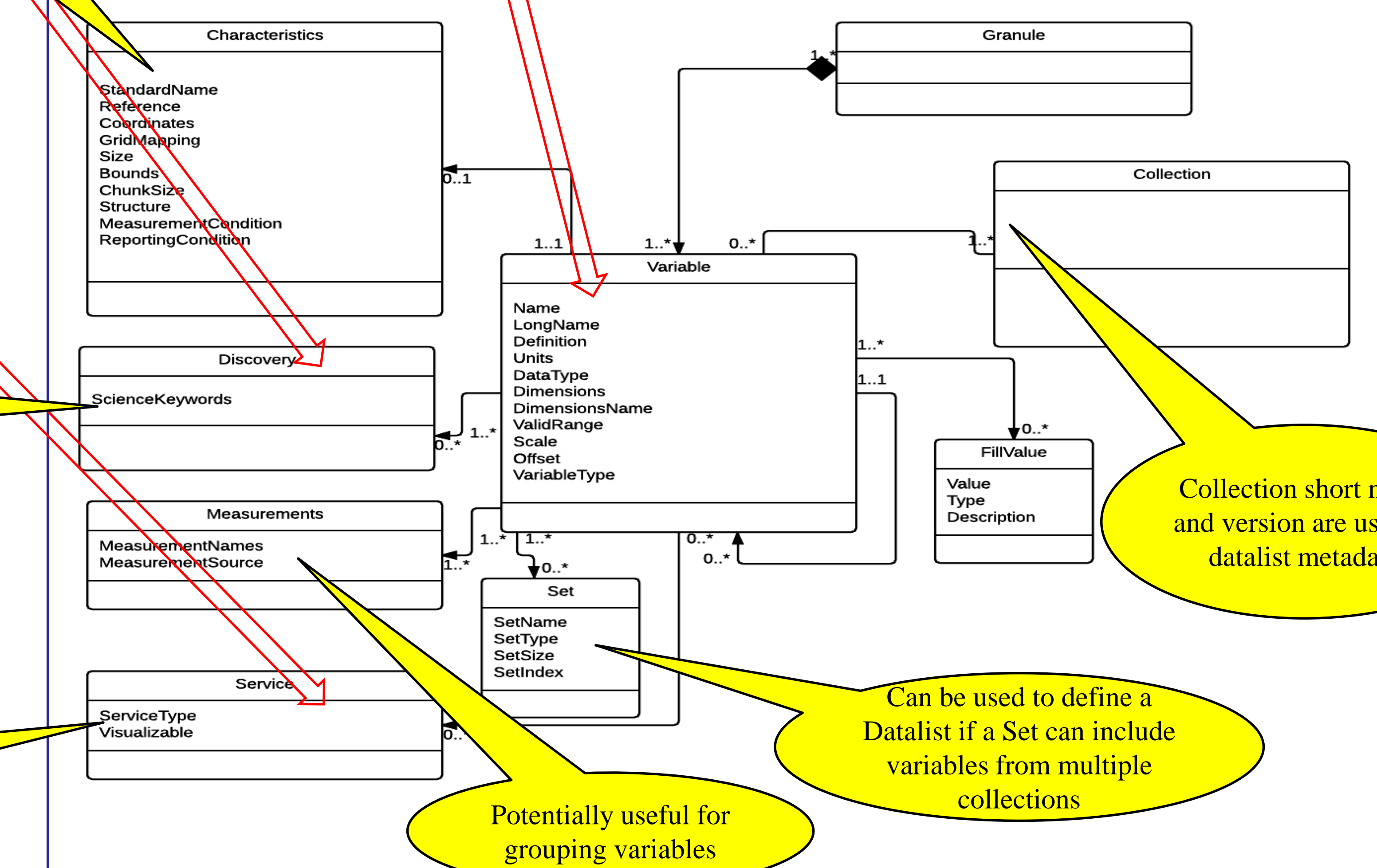[5] SGT

Search and Result

Key in "Hurricane"

Hurricane Datalist shows up at the bottom of the screen, while the top of the screen lists regular data collections

Datalist Landing Page

Subsetting and Data Download

Notice both MERRA2 and TRMM data are here. One Stop Shopping!

Temporal, Spatial and Variable Subsetting

UMM-Var and its Relationship to other models as of June 30, 2017

Potentially useful

Current dataset keywords are at collection level

Collection short name and version are used in datalist metadata

Can be used to define a Datalist if a Set can include variables from multiple collections

Currently datalist services are at collection level

Potentially useful for grouping variables

https://disc.gsfc.nasa.gov