

[Schema.org Hack-A-Thon](#) [1]

Submitted by erinmr on Sun, 2014-06-29 14:26 Wednesday, July 9, 2014 - 13:30 to 15:00

Event: [Summer Meeting 2014](#) [2]

Session Type: [Breakout](#) [3]

Expertise Level: [Beginner](#) [4]

Collaboration Area: [Semantic Web](#) [5]

Abstract/Agenda:

Do you want your datasets to show up on the first screen of search engine results? Of course you do! This hands-on session is for people who would like to learn more about the extension to the [schema.org](#) [6] vocabulary for datasets and data catalogs. We will begin with the basics, introducing [schema.org](#) [6] and then highlighting the Dataset and DataCatalog classes. Many data providers generate dataset landing pages on-the-fly by pulling information from a relational database perhaps enhanced with content from a semantic knowledge base. The original structured data is often obfuscated if the dataset landing pages are formatted in HTML, but recent vocabularies provided by [schema.org](#) [6] can help restore the structure through shared markup vocabulary. Proper use of the [schema.org](#) [6] vocabulary encoded in microdata, RDFa, or JSON-LD formats exposed structured and parsable information to content published out as HTML, and the marked up pages are recognized by search engines.

Several ESIP partners have successfully implemented [schema.org](#) [6] Dataset for their dataset landing pages and will guide participants through the process, including a review of tools available from Google Webmaster Tools for validation and testing of search results. Participants in this session are encouraged to bring a sample of their dataset landing pages (e.g. the HTML source code for a representative dataset landing page) and receive guidance on how to add [schema.org](#) [6] Dataset markup.

Related information and examples: http://logd.tw.rpi.edu/schemaorg_dataset_extension [7]

Co-conveners: Doug Fils and Adam Shepherd

Notes:

Doug Fils

Adam Shepard

Introductions

Overview

Code example (jsbin)

Discussion

People want a box, discovery occurs when people know where to look. Current practices don't do a good job of allowing data discovery

Motivation: expression by community to being able to access data

Connections are lost in the code, portability/sustainability is decreased, becomes dark knowledge

What can we do about it?

APIs: difficult to use and not a good way to drive discovery

Linked open data. Doesn't necessarily drive discovery.

Current search is not enough

Current tools, why the hate?

Schema.org: placing machine readable information into xml

The resolution,: make the web representation of your data smarter

What is involved?: Dataset is the only "science data" focused vocabulary at this time.

-extend your own vocabulary

-no guarantee that others will use it

-there is a formal extension pattern (see FAQ)

shema.org is not a majic bullet, there are other ways. It's not focused on geospatial

It is in active development, new developments in the works [shema.org/actions](https://schema.org/actions)

There are tools(<http://goo.gl/Ro5kDv> [8]), any23 for apache

Live demo: paleo-search.appspot.com/#gsc.tab=0

Crafting a scientific knowledge graph

virtuoso, kaylee's (from google, sp?)

Bring forth the affordances:

<http://jsbin.com/jenesi/5/edit>

need to be used to be sustainable.....

Live examples:

Needs to be a data tab on google.com, but to do that data sets need to be marked up with shcema.org

If we mark up our dataset with information about tools, uses for the data, then new tools will be built that leverage that information

datasets.schema-labs.appspot.com

jsbin.com/tocidu/2/

use data catalog is you have multiple dataset in a group

www.google.com/webmasters/tools/richsnippets [9]

few examples, see notes for addresses....

tips for implementation....

Need to put notes in ESIP system...

How Adam and Doug implemented....ask for help if you need it.

Q: Is schema.org code viewable in pages

A: yes

goo.gl/72a7Kt

goo.gl/Ro5kDv

Session Leads:

Name: [Douglas Fils](#) [10]

Organization(s): [Consortium for Ocean Leadership](#) [11]

Email: dfils@oceanleadership.org [12]

Name: [Adam Shepherd](#) [13]

Organization(s): [Woods Hole Oceanographic Institution](#) [14]

Email: ashepherd@whoi.edu [15]

Notes takers:

Name: [Kevin Dobbs](#) [16]

Organization(s): [University of Kansas](#) [17]

Creative Common License: Creative Commons Attribution 3.0 License

Teaser: This hands-on session is for people who would like to learn more about the extension to the schema.org vocabulary.

Accepted:

Source URL: <https://commons.esipfed.org/node/2557>

Links

[1] <https://commons.esipfed.org/node/2557>

[2] <https://commons.esipfed.org/2014SummerMeeting>

[3] <https://commons.esipfed.org/session-type/breakout>

[4] <https://commons.esipfed.org/taxonomy/term/260>

[5] <https://commons.esipfed.org/collaboration-area/semantic-web>

[6] <http://schema.org/>

[7] http://logd.tw.rpi.edu/schemaorg_dataset_extension

[8] [http://goo.gl/Ro5kDv](https://goo.gl/Ro5kDv)

[9] <http://www.google.com/webmasters/tools/richsnippets>

[10] <https://commons.esipfed.org/node/1291>

[11] <https://commons.esipfed.org/taxonomy/term/597>

[12] <mailto:dfils@oceanleadership.org>

[13] <https://commons.esipfed.org/node/745>

[14] <https://commons.esipfed.org/taxonomy/term/391>

[15] <mailto:ashepherd@whoi.edu>

[16] <https://commons.esipfed.org/node/1983>

[17] <https://commons.esipfed.org/taxonomy/term/641>