

bit.ly/2QvxEPM

Schema.org - Developing a Plan to Govern science-on-schema.org

<https://2020esipwintermeeting.sched.com/event/VaXr/schemaorg-developing-a-plan-to-govern-science-on-schemaorg>

ESIP WINTER MEETING - THURSDAY LUNCH: ESIP Testbed Proposal Writing Find Ruth, Adam, Chuck, Lewis, Steve or Doug

Tuesday, January 7 • 4:00pm - 5:30pm

AGENDA

[Proposal for Governing science-on-schema.org as the ESIP Schema.org Cluster](#)

[Investigate schema.org release strategy & metadata structure for extensions](#)

[Geoschemas.org](#)

[How do we serve content? SHACL shapes? Extensions to JSON-LD?](#)

MEETING NOTES

[Attendees](#)

[NOTES](#)

ABSTRACT

AGENDA

1. Proposal for Governing science-on-schema.org as the ESIP Schema.org Cluster

<https://github.com/ESIPFed/science-on-schema.org/issues/30>

How do changes to guidance documents and the repo happen?

[Decision to let interested people become moderators](#)

Ruth: Github may be barrier to contribution for non-technical people
- Mark, Dan Fuka agree, but Mark thinks non-techies need to “get over it” in the use of GitHub, but we need to make an extra effort to get non-techie, subject-knowledge expert input.
- RUTH: Key is the need to address non-techie requests.

Matt Jones: need both Git and another environment to capture requests.

- Dave V. says something like StackOverflow

- Chuck: says YAMZ

- Email lists also used by W3C DCAT, and CF group

We can use the ESIP Cluster resources as avenues outside of Github to use.

- Ruth: never close an issue without a resolution
- Kai: use the '@' mechanism in Git to keep people engaged
- Mark P: lazy consensus as acceptance, i.e. no dissent after a period is agreement. But we must be sure that we solicit the views of people who are not following the details of the cluster. Be inclusive while still making reasonable progress.
 - Matt: agrees, but let issues that become stale fall through the cracks (Mark agrees). Dan noted that if they are real issues they will reemerge.
- Ethan: CF for NetCDF is also using Github and it's going well for non-techies
 - Ethan: Matplotlib has a nice [one-pager](#) (on governance)
 - Python library process (pip/pep [8000](#) series)
- TODO will update governance markdown with rules about when issues go stale etc.

2. Investigate schema.org release strategy & metadata structure for extensions

<https://github.com/ESIPFed/science-on-schema.org/issues/29>

3. Review ESIP Winter Meeting tagged Issues

<https://github.com/ESIPFed/science-on-schema.org/milestone/2>

- <https://github.com/ESIPFed/science-on-schema.org/issues/4>
 - Matt: can we use: (Citation Ontology)
<metadata resource> cito:documents <dataset resource>
 - Steve: make additionalType an array of types it conforms to
 - Is Wikidata a good place to have URIs about things?
 - Ruth: maybe they aren't a good place?
 - Lewis: COR might be a good place over DBpedia and Wikidata
 - Controlled vocabulary for format types: <https://cn.dataone.org/cn/v2/formats>
Some ontologies like [Semanticscience Integrated Ontology](#)

Have many predicates defined for "is related to" including:

[is related to](#)

- + [has attribute](#)
- + [is attribute of](#)
- + [is spatiotemporally related to](#)
- + [is generically related with](#)
- + [refers to](#)
- + [is mutually related to](#)
- + [is comparable to](#)
- [is referred to by](#)
- + [is represented by](#)
- + [is referenced by](#)
- + [is described by](#)

- [in relation from....](#)

DECISION:

1. Linking Metadata docs to Datasets:
use <Dataset> `schema:subjectOf` <Metadata>
 2. `dcterms:conformsTo` inside <Metadata> to link to encodingFormat
 3. Vocabulary for File Encodings:
 1. W3C Data Exchange WG - Profiles Ontology (use dcterms:conformsTo)<https://www.w3.org/TR/dx-prof/>
 2. ESIP COR seeded with DataONE file formats
 3. Lewis: SWEET has the following the data formats - <http://sweetontology.net/reprDataFormat> these are of course limited to the Earth and environment knowledge space.
 4. Lewis: The Dataone controlled vocab is much more 'complete' but does not provide IRI's.
 5. Lewis: Master list of all registered Media Types can be found at IANA - <http://www.iana.org/assignments/media-types/media-types.xhtml>
 6. Write an ADR: Needs to be written for data and repository managers. Liaisons submit to relevant social media places (email lists, etc.) Includes comment by dates.
- b. **Recommendation on specifying dateModified for Datasets**
<https://github.com/ESIPFed/science-on-schema.org/issues/67>

data? Doug: is sitemap lastmodified enough? Since to view dateModified in the harvested

Matt: No because inside the harvested data graph there might be resources that also need to be re-harvested (e.g. downloading an updated metadata file)

I understand.. But datemodified is valid for any creative work.. So are you just saying this is a recommended field on subjectOf type nodes?

DECISION: add recommendation to the guidance docs

- c. **What copyright or software licenses should be used**
<https://github.com/ESIPFed/science-on-schema.org/issues/60>

DECISION: update the guidance documents to recommend appropriate ontology (e.g., CC licenses) and if nothing else exists use spec URL. License vocab <https://spdx.org/licenses/> (text or URL)

- d. Philemon: Spatial coverage of GeoCoordinates always resolves to Antarctica by Google Dataset Search?
- e. Dan Fuka: How to describe privacy or restricted access in schema.org
- f.
- g. <https://github.com/ESIPFed/science-on-schema.org/issues/13>
- h. <https://github.com/ESIPFed/science-on-schema.org/issues/54>

4. Geoschemas.org

How do we serve content? SHACL shapes? Extensions to JSON-LD?

MEETING NOTES

Attendees

1. Adam Shepherd
2. Alistair Ritchie

   Names - Alphabetically	   Names - Alphabetically
   ESIP Staff - Organizer, Presenter	   ESIP Staff - Organizer, Presenter
 kai - Organizer	 kai - Organizer
 Andrew Macleod	 Andrew Macleod
 Caller 01	 Caller 01
 Daniel R. Fuka - Me	 Daniel R. Fuka - Me
 David Vieglais	 David Vieglais
 Debbie Roth	 Debbie Roth
 Dirk Fleischer	 Lucy Profeta
 Lucy Profeta	 Lulin Song
 Lulin Song	 Mark Parsons
 Mark Parsons	 Micah Wengren - Web
 Micah Wengren - Web	 mike

Doug Fils

Micah Wengren

Mike?

Mark Parsons

Lulin Song

Lucy Profeta

Debby Roth

David Vieglais

Daniel Fuka

Andrew Macleod

Kai

Kirk Fleischer

Gary Berg-Cross

Matt Jones

NOTES

ABSTRACT

This session will walkthrough the ESIP Github repository at <https://github.com/ESIPFed/science-on-schema.org>

See: <https://github.com/ESIPFed/science-on-schema.org/milestone/2>

Discussion:

- * How do we govern as a cluster?
- * Monitoring updates to schema.org?
- * Strategies for proposing changes to core schema.org?
- * Extensions at geoschemas.org How to Prepare for this Session: Review the contents of <https://github.com/ESIPFed/science-on-schema.org>