

June Watercooler - EML Best Practices Document: comments and requests



Published on *LTER Information Management* (<http://im.lternet.edu>)

Home > June Watercooler - EML Best Practices Document: comments and requests

June Watercooler - EML Best Practices Document: comments and requests

Mon, 06/07/2010 - 10:36am — mobrien

Working Group area:

http://intranet.lternet.edu/im/news/committees/working_groups/emlbestpractices

Introduction:

The EML Best Practices working group has been reassembled to update the document first created in 2004. The group will meet in Albuquerque at the LNO, and plans to have a draft ready for this year's annual IMC meeting at KBS. It is our goal to

1. bring the document up to date for the latest version of EML (2.1.0),
2. to include recommendations for some additional data types,
3. to consider EML-content which could be delivered from our emerging dictionaries (e.g., controlled vocab and units), and
4. to facilitate publication on the IM website.

Notes from Tuesday, June 8, 2010

IM_VTC_EML_Best_Prac_20100608

in attendance: duane, Nicole, Margaret, hap, jonathan, mark, yang, linda, john, sven, wade, thersa, corinna

- what needs to be changed?
- o authoring section - what if someone leaves
- f one of reasons we suggest listing IM office as contact
- f keeping contact info current is hard! on levels - need to be more specific and focus on quality
- f previous document had tiered trajectory as heavy influence
- f some aspects are arbitrary - little middle ground
- f we should back off on the levels some
- f focus on important drivers for machine usability
- f now tends to be binary - level 5 or not
- f like to see more on quality
- f keywords, units, attributes
- f did not do much on content standardization in previous document
- f think more about content, less about structure content standardization should this group take on?
- f yes - if not here, where?
- f can use “must” “should” to delineate degree of consensus
- f issue - how methods are entered
- f very few give specific method steps
- f we may decide its OK to aggregate
- f 3 different places for methods in EML
- f human readable
- f but can be displayed selectively with tools if marked up appropriately
- f guidelines for content are a good way to go... process look to see what sites do...
- f but what should we do where there is wide variance
- f can recognize variance and discuss to come up with recommendations
- f locations of research sites as metadata or as data
- f EML coords as bounding box or down to sample level?
- f currently recommend bounding box for areas with samples
- f probably “should” not “must” if disagreement
- GIS data - issues of FGDC vs EML metadata
- f many sites not using EML for GIS data
- f locations as data vs metadata
- f can see how original recommendations map to future uses
- f ESRI metadata and EML
- f describing tables in EML
- f is level 5 EML for GIS data really useful beyond discovery-level
- f could use in PASTA
- f can use in KEPLER for GIS data
- f attributes in DBF file
- f metadata is part of shapefile - so you get nested metadata - FGDC inside EML
- f want to get FGDC translator working better - then can pull out FGDC data

into EML easily
 f can include attributes, ranges, domains, codes etc.
 f FGDC lacks units etc. that are needed for EML data integration
 f methods are important
 f experimental design
 f sampling unit or unit of analysis
 f want to include this as specifically as possible
 f different locations for methods
 f different scopes
 f ties into derived products -may only want methods for some attributes
 f getting methods in any form is difficult!
 f aren't there standard methods? could those be incorporated? machine readable
 needed for PASTA
 f Data Manager Library tests for "pasta ready"
 f does the metadata match the data?
 f if it doesn't, can't load data
 f in reviewing best practices, at level 3 there is a lot to make a doc PASTA-ready
 f could also include other quality checks
 f but won't be restricting harvest - just reporting would be nice to get normalized
 forms of tables - not just denormalized forms for input back into a database
 f do I want to offer normalized data
 f or share denormalized
 f I'd suggest both
 f try to develop tools for doing that we can share examples were really helpful
 f had fun doing those last time!
 f all of old links are now broken when SVN was refactored attributes and
 chemical units - units of WHAT
 f right now attributes are text
 where can we get periodic table data in here
 attribute standardization?
 for right with EML structure now need to have element name in attribute
 description
 could use kludge of setting up standard units that include objects of the units
 - with basic units as link to them in unit registry would be needed for unit
 conversions in PASTA
 f but that is not planned to be automated
 f robust descriptions are enough EPA Storet codes might be a good way to go
 f links methods, units, objects together
 f lots of opinions on STORET

- next steps

1. o want you to list several representative EML IDs from your site
2. o look to see what metadata standards other EONs are using?
 - f good idea, but sometimes hard to do
 - f NEON may be using EML

f may be broader issue

- Virtual Updates [1]

- Copyright © 2012 Long Term Ecological Research Network, Albuquerque, NM -
This material is based upon work supported by the National Science Foundation
under Cooperative Agreement #DEB-0236154. Any opinions, findings, conclusions,
or recommendations expressed in the material are those of the author(s)
and do not necessarily reflect the views of the National Science Foundation.
Please contact us with questions, comments, or for technical assistance regarding
this web site.

Source URL: <http://im.lternet.edu/node/599>

Links:

[1] <http://im.lternet.edu/taxonomy/term/169>