

Units WG Notes

Pinned:

<https://github.com/EDlorg/Units-WG> (notes on QUDT submissions, R code for metadata incorporation)

<https://github.com/EDlorg/qudt-public-repo/tree/main> (our fork of QUDT)

<https://github.com/EDlorg/unit-registry> (vestigial)

https://ble-lter.github.io/QUDT_Browser/ Tim's QUDT browser

Regular meeting: second Wednesday, 11 Eastern

Connection info:

<https://ucsb.zoom.us/j/85166868304?pwd=a0RUUW1hQ3VGM0k3ZER0ejlkczs2UT09>

Membership:

Mary Martin <mary.martin@unh.edu>,

Marina Frants <mfrants@ucsd.edu>,

Christine Laney <claney@battelleecology.org>,

Stevan Earl <srearl@asu.edu>,

John H. Porter <jhp7e@virginia.edu>,

Sarah Claire Elmendorf <sarah.elmendorf@colorado.edu>,

Hillary Krumbholz <hkrumbholz@ucsb.edu>,

Suzanne Remillard <suzanne.remillard@oregonstate.edu>,

Mark Servilla <mark.servilla@gmail.com>,

Matt Jones <jones@nceas.ucsb.edu>,

Mya Yazbek <mya.yazbek@eastern.edu>,

Margaret O'Brien <margaret.obrien@ucsb.edu>,

Past members:

Chris Turner <chris@axiomdatascience.com>, (deliveries fail)

salinkatherine.muller2010@gmail.com,

kem325@cornell.edu,

kylezollo@gmail.com,

"Kristin Vanderbilt (krvander@fiu.edu)" <krvander@fiu.edu>,

Elie Arnaud <elie.arnaud@mnhn.fr>,

Anna LaBry <rmlabry@gmail.com>,

An T Nguyen <enthusiast@utexas.edu>, (deliveries fail)

2026-09-09

Present:

Regrets:

Agenda

Stevan - Go over results of the recent AI attempts

Margaret - Report on chat with Steve Ray and status of our QUDT fork/ pull request

All - Paper writing assignments

Notes

Margaret kept notes on using the QUDT software to merge our units into the fork. Link is here so she doesn't lose it. Probably too much detail for this group, but not for the QUDT folks.

Link: [Working with Maven](#)

2026-08-19

Present: Stevan and Margaret

Agenda

Come up with a plan for testing AI models. For the paper, we would like to say something about how well AI worked for generating these units. But our process so far has been makeshift.

Notes

Status to date of AI work

had to think way back as to how we arrived at the 84 units we have now, in our QUDT contribution. This was a partly manual process

Process: in Fall-winter 2025

Margaret had used claude to generate ttl for 180 new units (all the unmatched).

Inputs were a verbal prompt (based on Mary's), and 2 files, one describing the parent-child relationships of units, QKs and DVs, and one with a "set of rules". Files

(input and output) are all in git, under

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/claude_AugSep_mob

Mob ended up generating 2 batches of units, mostly because she would run out of claude-cred trying to do it all at once..

Claude generated ttl for 180 units. We divided these among WG members for manual vetting. Moderate success. See notes column here: for specific problems noted:

 QUDT contribution log.xlsx (see tab QC log, AI output)

Example Problems:

1. At first try, claude didn't seem able to understand how to create QUDT files (e.g., after directing it to the git repo and documentation, it didn't follow rules or copy the file structure). The parent-child graph and a list of rules helped somewhat. E.g., Claude still didn't follow rules: e.g, wrongly formatted DVs.
2. In general, an auto-generated process could mean we might be creating new QUDT-ttl for something that was a typo. We caught this during vetting stage (see notes, 2025-08-13):
 - Attribute is: Average_solar_radiation_rate
 - micromoles per square meter - second (umol m⁻² s⁻¹)
 - The unit is written as "micromolePerMeterPerSecondSquared"
 - Umol m⁻² s⁻¹ should have been micromolePerMeterSquaredPerSecond.This typo appears in a dozen or so packages from a sunsetted LTER site.
3. For many units, vetting was difficult (or not possible) without going back to the original dataset. After doing a few this way, others were postponed. This is not claude's fault – QK, especially, depends on other dataset metadata.

By February 2026, (4-5 months later) the group had enough. We decided to go with what was finished to date (about 80 units) for our next QUDT contribution.

TO DO: Describe what this group of 80 represents, in the unit corpus, ie, frequency of use. How many units were manually edited? What fields? (based on log in spreadsheet, at least half)

Conclusion:

It doesn't make sense to redo the 180 units Margaret did, since we know there are many that were already rejected from our QUDT contribution.

Next steps:

Goal is to properly evaluate the use of AI for this process.

We have a list of 84 units in ttl - AI-created, manually edited, and we have good confidence in them

Use this ttl file as our benchmark. Stevan will create an agent to generate these units again in several models (chatgpt, claude). We can compare those outputs to our benchmark. Added later: AI should also create accompanying QK and DV files, as needed.

2026-08-12

Present: John, Margaret, Marina, Stevan
Regrets

Agenda

1. Units contribution for QUDT
 - a. This git dir:
 - b. Margaret's notes (some stuff pertinent to the paper)  AI notes, mob
 - i. To resolve, this group:
 1. Confirm no unknown QKs? (CountPerMass, CountPerMassRate) - YES. no unknowns.
 2. Help finding some "informative refs for new QKs (see git file)
 - a. MoleSpecificBiogeochemicalRate
 - b. Others
 3. Frequency of use (of these 86) in datasets (John)
 - ii. To resolve with Steve Ray
 1. Use of skos:broader, qudt:organizedUnder. When do we need both?
 2. Is there a viz of QKs? (try protege first)
 - c. Discussion of AI use
 - i. QUDT contents vs broader scientific context
 1. See some of above with QK
 - a. Needs more context from datasets???
 2. It was relatively easy to create some QKs manually
 - a. But Claude had a very hard time doing that
 - i. Perhaps because I told it to match existing
 - ii. Rather than telling it to look at QKs for similar units
 - iii. Rather than telling it to be more liberal in how it found it out
 3. Many units have similar base elements (e.g., meter, second) but different multipliers (kilometers/hour, meters/second) are those "commensurate"? Dimension vectors? DMs don't map to QKs

directly because DMs don't include ratios where elements cancel out (moles/moles/second vs gram/gram/second – all have DM time only).

4. Discussion of hierarchy of quantityKinds looking at MassSpecificBiogeochemicalRate

2. Paper outline:  Outline1

Action items

Margaret:

DONE Add the file of “yes” units to git (claude_inputs, should be about 100). Derived from this tab:  QUDT contribution log

Lists of units here:

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/claude_AugSep_mob/inputs

Claude needed the trailing comma, because it concatenated all lines in a text file when uploaded, and couldn't separate them. So the comma became the delimiter.

Chat with Steve Ray, show and tell

Validate our fork, pull request if possible.

Stevan:

second AI attempt to generate units, less random than margaret's.

All:

continue edits to paper outline.

Next time:

Stevan - Go over results

Paper - come up with writing assignments

Margaret - Summary of talk with Steve Ray, Maybe a QUDT pull request

2026-07-08

cancelled.

2026-06-10

Present: Margaret, Marina, Stevan

Regrets: John (out hiking)

Agenda

1. Units contribution for QUDT
2. Paper outline: [Outline1](#)
 - a. John has made some suggestions in the outline particularly in the GOALS section. If we can get the goals clarified, so that they can guide the direction of the paper, that would be ideal.

Notes

Per John: hone the goal. What is the story that the world needs.

Ontologies need to improve, can't be static

Don't need to be encyclopedic (id what needs to be added)

How to generate

How much to use AI

Use Stevan's preflight check script:

https://github.com/EDlorg/Units-WG/blob/main/.github/skills/qudt-ttl-generation/scripts/ttl_preflight_check.sh

2026-05-13

Present: Margaret, Stevan, John, Mary

Regrets: Marina

Agenda

Finish added the conversion factors (via gpt) to the ttl files

Paper

Continue with outline

John's QUDT analysis of units found only in complex units (not as fundamental units)

Notes

Stevan's VS code prompts from February meeting for adding multipliers to AI-generated QUDT.
DONE. see mob_ttl branch.

Stevan ran mob's file thru a QUDT checker (Stevan will push this checker to git). See 4 errors below (all D1 errors)

```
Checking ../../../../AI_create_ttl/new_units_ttl_files/newQUDT_Units_mob.ttl
ERROR: unresolved TODO markers found
166:### TO DO, check def against others with
<http://qudt.org/vocab/quantitykind/ActivityConcentration>
259:## TO DO: check further. this one might be redundant, not that we've standardized
ordering of components (to per M2 per YR)
439:## TO DO may be redundant after reordering. check further.
537:# TO DO - check qk, see notes.
570:# TO DO - check qk, see notes
653:# TO DO - not sure of qk, see notes.
670:# TO DO - see gm-per-m-day
692:# TO DO - compare qks to other similar units. see notes
709:# TO DO - compare qks to other similar units. see notes
814:# TO DO
832:# TO DO
849:# TO DO
901:# TO DO
935:# TO DO, See :PicoMOL-PER-MIN-MilliGM
952:# TO DO - check other units with similar dimensions
ERROR: D1 used on non-dimensionless vector
1145: qudt:hasDimensionVector qkdv:A0E0L0I0M0H0T-1D1 ;
1179: qudt:hasDimensionVector qkdv:A0E0L0I0M0H0T-1D1 ;
1224: qudt:hasDimensionVector qkdv:A0E0L0I0M0H0T-1D1 ;
1239: qudt:hasDimensionVector qkdv:A0E0L0I0M0H0T-1D1 ;
```

John's analysis of complex units that have components that are not present:
We noticed this because we are adding KiloM3 (surprised it wasn't already there.)

Why this might be a problem:

1. Some units invoke units that don't exist. In the example below, this rule is included for unit:PERMILLE_PSI unit:MILLE is not present. So the rule can't be checked (or should fail)

```
qudt:hasFactorUnit [
  a qudt:FactorUnit ;
  qudt:exponent -1 ;
  qudt:hasUnit unit:MILLE ;
] ;
```

Note PERMILLE has replaced PER-MILLE

2. Impedes understanding of a unit, if you can't look up components. Helpful for figuring out conversion factors of a new complex unit

Share this as an issue at <https://github.com/qudt/qudt-public-repo>

NEEDS TO BE RERUN EXCLUDING REPLACED /DEPRECATED UNITS

[See Tab labeled "John's analysis, unit elements missing"](#)

In case we need it, this sparql returns the non-deprecated units (only the first prefix is necessary here):

```
PREFIX qudt: <http://qudt.org/schema/qudt/>
PREFIX unit: <http://qudt.org/vocab/unit/>
PREFIX qkdv: <http://qudt.org/vocab/dimensionvector/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX quantitykind: <http://qudt.org/vocab/quantitykind/>

SELECT ?unit

WHERE {
  ?unit a qudt:Unit .
  MINUS { ?unit qudt:deprecated true }
}

ORDER BY ?label
```

These 29 unit elements are in complex units, but not defined as QUDT units:

A2
C2
C3
C4
CentiM5
EV2
FT6

GAL_UK2
GAL_US2
GigaEV2
HZ2
HZ3
HZ4
IN6
J2
J3
K4
KiloGM2
KiloM3
L2
M0dot5
MicroMOL2
MicroS2
PA2
PlanckMass2
SEC3
T2
V2
W0dot5

To do: find some complex units with these elements, and see if any will be problematic, to use as examples.

2026-04-08

Present: Margaret, Stevan, Marina, John, Mary
Regrets:

Agenda

- Report back on QUDT contribution (Margaret).
- Work on paper outline:
<https://docs.google.com/document/d/1wU7W2OMXNUU5dFcLnTrbg5cRm0sOGMf0mBG7UWvtU9I/edit?usp=sharing>

Notes

Worked on outline, above.

If we want to invoke our use of AI in this paper, we need to refine the process a bit. Generally all we did so far was bumble around.

To do: Margaret and Stevan - outline the ideal process, try and create a prompt that will work better.

Number of units: mob uploaded a list (file). Claude had time limits, and would stall. Tell it to work on only a few at a time.

QKs - look for similar units in qudt (same DV, similar magnitudes)

Maybe answer the question: how much environmental science knowledge does Claude need to choose (if existing), or create a QK if not

2026-03-11

Present: John, Stevan, Margaret

Regrets:

Agenda

Report back on QUDT contribution (Margaret).

Review writing projects.

- Paper (John is leading)

- Blog post?

- Notes, status archived?

Draft outline of paper to work on:

<https://docs.google.com/document/d/1wU7W2OMXNUU5dFcLnTrbg5cRm0sOGMf0mBG7UWvtU9I/edit?usp=sharing>

Notes:

Margaret has not gotten far on assembling the entire contribution. Hoping by next month. John presented his outline (link above). Margaret and Stevan made significant edits.

Paper's Goal:

1. Our goal is to detail the mechanisms and considerations for adding new units to the QUDT unit ontology, and the extent to which the processes and decisions can be automated and augmented with AI resources.
2. Describe a workflow that can be used by a repository (EDI).

2026-02-11

Present: Margaret, John, Mary, Stevan, Marina

Regrets:

Agenda

progress vetting/creating the new units.

Log is here:

<https://docs.google.com/spreadsheets/d/12rVYLgtvBo48mmayZLya46Uja7W66syITajrCoSIXZM/edit?gid=547171838#gid=547171838>

New TTL definitions go here (add a file with yours, include your name or initials in the filename):

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/new_units_ttl_files

Notes

We now have TTL for about 100 units to contribute to QUDT. that's enough for our contribution. If there are any common (20 or more uses) that we have missed, those will pop to the surface when people start using these tools.

Next steps:

~~Get multipliers in. Margaret didnt put these into the claude prompt, so it's not in about 80% of the ttl units.~~

~~Process:~~

~~Margaret will double check John's and Mary's files.~~

~~John, Mary, Stevan, Marina will divide up Margaret's file and add offsets and multipliers.~~

~~John's process documented:~~

- ~~1. Find a similar unit in QUDT (same dimension vector - see notes on that below)~~
 - ~~a. Target unit: MicroMOL-PER-L-DAY~~
 - ~~i. Similar unit: MicroMOL-PER-L-HR~~
 - ~~ii. So take the existing multiplier and divide by 24~~
 - ~~b. OR~~
 - ~~i. Similar unit: GM-PER-L-HR~~
 - ~~ii. The multiplier for the new unit is the same, since the relationship between GM and KiloGM is the same as microMOL to MOL~~

Another method:

Use Claude:

Stevan's prompt (copied from chat), for VS code. The AI behind this is GPT-5.1-Codex-Max

The ttl file is a document describing new units that we are proposing to submit to the QUDT organization. Please review the mission, instructions, and mechanics of QUDT units: <https://qudt.org/>
The units listed in this file are missing the conversion multipliers to SI, which should be in the following format for each entry:
qudt:conversionMultiplier 0.01157407407407407407407407407407 ;
qudt:conversionMultiplierSN 1.157407407407407407407407407407E-2 ;
Assess the current units and edit the file to include the accurate, complete conversion factor for each proposed new unit.

See git repo for updated ttl files, with conversion factors.

Stevan to do:
Add conversion factors to files as needed.

Margaret to do:
Combine ttl files (remove problematic ones), add to qudt branch, validate.
Create DVs and QKs as needed.

Contact Steve Ray, that we have ~100 units.
Questions:

2026-01-14

Present: Margaret, John, Marina
Regrets: Stevan

Agenda

progress vetting/creating the new units.
Here:

<https://docs.google.com/spreadsheets/d/12rVYLgtvBo48mmayZLYa46Uja7W66sYlTajrCoSIXZM/edit?gid=547171838#gid=547171838>

Defined units here:

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/new_units_ttl_files

summarize potential issues down the road (eg, changing code, coordinating with QUDT.org...).

Notes

How to create your ttl file

1. Start a base file. Use the header from https://github.com/EDlorg/Units-WG/blob/main/AI_create_ttl/new_units_ttl_files/newQUDT_Units_mob.ttl
2. Copy your list of units into your new file (eg, names only, or a ttl entry from somewhere else - like the claudie output or a similar unit in production QUDT.)
3. Add it to https://github.com/EDlorg/Units-WG/blob/main/AI_create_ttl/new_units_ttl_files. Make sure to identify it as yours
4. Edit the ttl entry as needed
5. Validate your file with an online ttl tool like <http://ttl.summerofcode.be>
6. Check work into git

Future scenario

1. For EDI: are their update processes? what do they look like? ie, how to communicate with Jon (or other programmer working on ezEML)
2. Outline a possible process for submitting new units ad hoc (eg, a form that generates ttl, another WG campaign)
3. Do we recommend another dataset extraction from EDI, to be rematched? This would generate a new list of potential new units
4. Communication with QUDT:
 - a. Deprecated units:
 - i. if these were previously matched, what does their URI look like now? Has it changed (ie, does it now include a version number?)
 - ii. Or does QUDT have a deprecated-in-favor-of field, and our URI stays the same (preferred).
 - iii. How will we know?
 - b. Will EDI communicate with QUDT? how?
 - c. Is Margaret overthinking all this? Will nothing happen for 10 years, and then something totally new will be there

Outline for Article

1. on process for getting annotating units into metadata - analyze what fraction of datasets how have QUDT units
2. Adding new units to QUDT (WG activities described)
3. Need to involve Jon Iles for EZEML

Action Items

Marina: start a ttl file for your 20 units

John: add some unclaimed units to your list (thanks!)

Mary/Stevan: continue QC step

Margaret: edit your ttl file
John: work on article outline

2025-12-10

Present: John, Stevan, Mary, Margaret, Marina
Regrets:

Agenda:

Go over QC results. Redistribute units, if they are not evenly distributed via their DVs. create a plan for ttl files, eg, where to add these to git. John may want to lead this, since he has already created one ttl file

QK=unknown vs action=postpone?

Notes

1. Discussed a few specific units.
 - a. Some of their elements should be reordered, to put time last (e.g. $\text{mg} * \text{d}^{-1} * \text{m}^{-2} > \text{mg} * \text{m}^{-2} * \text{d}^{-1}$). Often when that happens, the unit is already in QUDT. See action items, below
2. QK=unknown vs action=postpone?
 - a. Most thought the use of “unknown” was not a good idea if any reasonable QK was available. Units should be postponed if they are confusing
 - b. There is a lot of work to be done on QKs in general. Not something this group will deal with. Someday, people will want to search datasets by QK (eg, “photon flux”, “primary production”), and then that work will be important. Maybe a good project for AI - that is, analyzing dataset metadata to find QK terms and match likely units.
3. Discussed how to edit ttl entries.
 - a. Everyone is happiest keeping individual files with just the units they are working on (fewer git collisions). Margaret will concat these at the end. Left it up to individuals how they want to separate out their units (extract from claudet-ttl, generate de novo, or copy from a similar unit in qudt). For now, work on the units you have QC'd.
 - b. Check your file into this dir (be sure to identify it as yours):
https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/new_units_ttl_files
 - c. Margaret asks that you validate your ttl file as you work! E.g., with a tool such as <http://ttl.summerofcode.be>. You will need to include the header from the claudet file (defining the prefixes). Validation will help identify copy-paste errors.
4. Discussed possible deprecated units - QUDT appears to be removing some units from the production list. Do any of our datasets reference these? If so, what do we do?

Action items

Mary, Stevan, Marina: continue with unit validation

John: edit substitutions, mark units in [QUDT contribution log](#) if our action on them will change (e.g., no longer a new unit)

Margaret: start your turtle file.

2025-11-12

Present: Stevan, Margaret, John, Mary, Marina

Regrets:, Christine

Agenda:

go over QC results to date.

Notes:

About half the DV (and assoc units) are done.

Many units will have to be revisited, thru their datasets. Those will be marked as “postpone” or “discard” (probably some postpones will become discards).

John created some files that might be useful for finding related units in QUDT (an alternative to using sparq): https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/JP_Crossref_files

They are:

- RelatedQKs.cvs

- RelatedUnits.cvs (list of our new units, plus similar units (same DV) already in QUDT.

- RelatedQKs_summary.csv (count of now many time a QK is used_

- RelatedUnits_summary.csv

John will put these into git, in the AI_create_ttl directory.

Action Items:

Stevan, Mary, Marina: continue with unit QC, starting from the assigned dimension vectors here:

<https://docs.google.com/spreadsheets/d/12rVYLgtvBo48mmayZLYa46Uja7W66sylTajrCoSIXZM/edit?gid=1662521236#gid=1662521236>

John, Margaret:

Claude created some invalid DVs. revisit these and their units. See assignments in same spreadsheet above.

John will add substitutions and see about automating res-sorting of unit elements to try out all possible orderings to see what matches.

2025-10-08

Present: S. Earl, M. Frants, C. Laney, M. Martin, M. O'Brien, J. Porter

Regrets:

Agenda:

Go over QC process and divide up the remaining units (est 120 left)

Using these notes:

<https://docs.google.com/document/d/1wOVhs1c2KxVp1OVI5WGGzBM1bU2em653VvbJEYxwvQM/edit?tab=t.0#heading=h.e0aqouh89sbf>

2025-09-10

Present: John, Stevan, Margaret, Mary, Marina

Regrets: Christine

Agenda:

1. Go over the units that claude generated.
 - a. https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/claude_AugSep_mob/outputs

Tools

1. To see related units already in QUDT, use the browser here:
https://ble-lter.github.io/QUDT_Browser/
2. Link to sparql code to see units that use a specific dimension vectors: link to the header "SPARQL" in my notes file here: [AI notes, mob](#)

What other sparql queries do we need?

To do

The most likely scenario is that there are existing quantity kinds that we can use. We need to see the existing units that might be related to ours (via the dimension vector).

John will write some R code to produce quantityKinds that have the same DVs we are using. Possibly get lists of other units sharing the quantityKind (drawn from QUDT quantityKind).

This will go into a new dir in the git repo:

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl/claude_AugSep_mob/reference

Margaret: validate the 200 new units with the rest of QUDT (with maven). Add more fields for

1. UCUM codes. It's not in margaret's relationship diagram. There can be several, will one suffice?
2. ID field.

Git file should not have a date on it!

After that: group evaluates the maven-validated units using John's list of related QKs.

2025-08-13

Present: Mary, Margaret, Stevan, Marina

Regrets: Christine, John

Agenda

1. Generating the remaining units with AI. Margaret has had limited success.
2. Our plan for handling erroneous metadata. Here is one, that (after looking more closely at the dataset) was clearly a typo. See notes

Notes

Action items (AI):

Mary - run Margaret's prompt on claude, send her the files

Margaret: validate Mary's files (top braid and maven, with steve ray)

Margaret: get a short-term claude acct if you need it

Putting off dealing with erroneous metadata (as below example from CWT). finish process for generating new units first.

AI notes:

[AI notes. mob](#)

Erroneous metadata notes:

Not just un-mappable (like a 'martha_stewart') but had us trying to create a new unit in qudt for something that is a typo:

MicroMOL-PER-M-SEC2

pseudounit was micromolepermeterpersecondsquared

In knb-lter-cwt.1039.13

Attribute is: Average_solar_radiation_rate

Average solar radiation rate; micromoles per square meter - second ($\mu\text{mol m}^{-2} \text{s}^{-1}$)

The unit is written as "micromolePerMeterPerSecondSquared"

$\mu\text{mol m}^{-2} \text{s}^{-1}$ should be micromolePerMeterSquaredPerSecond instead.

This error occurs in additional CWT datasets:

knb-lter-cwt.1039.13

knb-lter-cwt.1040.19

knb-lter-cwt.1094.13

knb-lter-cwt.2007.13

knb-lter-cwt.1214.13

knb-lter-cwt.2007.13

knb-lter-cwt.4038.19

knb-lter-cwt.4100.19

knb-lter-cwt.4101.119

knb-lter-cwt.4102.18

knb-lter-cwt.4103.19

knb-lter-cwt.4104.14

Knb-lter-cwt.4106.14

We need to be on the lookout for the units we are creating now, and maybe EDI has some ideas for QC? Examining units in bulk is exposing some weird stuff. If it were an active LTER site, the obvious solution would be to inform the data mgr. But this site is no longer active.

2025-07-09

Present: John, Marina, Stevan, Margaret

Regrets: Mary

Agenda

1. Update from Steve Ray
2. ezEML with units
3. Experiences generating units with AI

Notes

Via email from Steve Ray -

they decided that statistical concepts are out of scope for QUDT. So they won't accept those units. Quoting his msg:

2. ConductivityVariance_NEON, if I understand correctly, describes the statistical concept of variance for a dataset. Statistical concepts are currently out of the scope of QUDT, and would at best be considered a qudt:Quantity layered on top of some dimensioned concept (as you did for Conductivity). In other words, Conductivity is a perfectly valid QuantityKind, but the variance of a dataset of conductivity values isn't.

That seems fine to me – probably stats concepts should come from a stats ontology.

ezEML:

Per Mary: The ezEML qudt functionality worked great for me. I think Jon had some outstanding questions in his last email. It would be great if everything got resolved to give him the green-light on moving that forward.

ACTION: Suggest that Jon announce its availability to a broader group of testers, and give access to ezeml-d for further testing by other contributors.

Experiences with AI -

Mary's results are here, using claude.ai:

https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl

Margaret tried claude.ai and chatgtp (both are free versions, so complexity and time limits)

Summary:

- I couldn't reproduce Mary's results in claude with the same query text.
- Neither system produced correct files. It might be fine, however - to go through results and clean them up manually. I got a little further with chatgtp. My notes: [AI notes, mob](#)
 1. I could upload the flowchart and a pseudo-ERD to try to teach it the cardinality.
 2. The output format matched QUDT's ttl files (claude's did not)

Manual cleaning:

I worked on Mary's output using TopBraid. Meh. Steve Ray sent instructions for using Maven, but I haven't installed that yet.

2025-06-11

Present: Stevan Earl, John Porter, Mary Martin, Margaret O'Brien, Marina Frants, Jon Ide.

Regrets:

Agenda:

Go over possible implementation(s) in ezEML, with Jon Ide.

Notes:

Background: New feature in ezEML: trusted users can publish directly from ezEML (presumably instead of going through a curator). Mary mentioned that she would always have to download so that she could add QUDT annotations. Jon thinks it would be easy to incorporate some of our work into ezEML.

Requirements:

EDI needs to run the service, not depend on ours. Jon has translated John's R code to python. Needs testing still.

The lookup table lives on the EDI server (periodic downloads, e.g., after community additions).

Operation:

Check box for user to say they want to add QUDT annotations.

Way to check the annotations

TBD:

- Viewing annotations: maybe in one list, rather than attribute by attribute.
- Attribute IDs: these are optional in EML, but required if an element has an annotation. Code should look to see if an ID is already present and not replace those.
- EML version - ezEML can read 2.1, and currently exports an edited doc as 2.1. To add annotations, it will have to upgrade a doc to 2.2. It might be best to upgrade all edited docs to 2.2
- What if doc is already annotated? Don't replace.
- user interacts with the annotation – how? ability to reject an annotation, but not to edit (not expecting users to have that level of expertise)
 - Issues will arise here.
 - Way in the future, there could be external annotation vetting, e.g., to involve external experts.

TBD down the road, maybe

- ?? a way for people to pick a unit off the QUDT list (someday). Maybe not necessary, if they can type in a term and then EDI can match it to our lookup table.
- Quantity kinds? Units are many:many with QK. QKs could be a way to group attributes, and search datasets. QKs link the unit to the measurement. For environmental science, are incomplete in QUDT. probably another working group would need to address this.

Other Notes

(margaret's) for Mary's generated ttl files in git:

(https://github.com/EDlorg/Units-WG/tree/main/AI_create_ttl).

Manual examination.

<https://docs.google.com/document/d/1wOVhs1c2KxVp1OVI5WGGzBM1bU2em653VvbjEYxvwQM/edit?tab=t.0>

Summary:

Some duplicates and missing DVs and QKs.

1. [claude.ai](#) isn't following the flowchart, Ideally, it should, and would probably have
 - a. created the DVs and QKs as required (no duplicates) - because it would have been told to look for a DV, and then decide among the QKs.(also might be a test of whether mob can write a flowchart or not)
 - b. Found QKs that matched the DV
2. Instruct [claude.ai](#) to export all conversion multipliers and offsets as decimals
3. Still not sure what is going on with those "property needs to have at least 1 values, but found 0" messages. The focus nodes are not familiar, and there are tokens in the paths. Question for Steve.

2025-05-14

Present: Mary, John, Stevan, Margaret

Regrets: Marina, Christine

Agenda: continue with unit creation discussion.

1. Recording units we're submitting in git
2. Also
 - a. Mary's latest chatgpt output (~20 units)
 - b. Her code (currently none exists, its some sort of web history)

Assemble some questions re using claude.ai for Steve ray

1. What character set? (check if their file declares that)
 - a. Eg, symbols in utf-8 or hex

Action items

Add checking for an existing QK and DV in claude.ai process (Mary)

Here: <https://github.com/qudt/qudt-public-repo>

Margaret: look at her pull req for hex in symbols

-> Unit symbols are NOT in hex. E.g., this: "uS/cm", "nmol/g"

Margaret: write to Steve Ray, invite him to June meeting for show and tell

Send link to ttl output in git

Git repos to clone:

<https://github.com/EDlorg/units-WG>

Where we put our work in progress

<https://github.com/EDlorg/qudt-public-repo>

Where we put stuff that is ready for QUDT.org

2025-04-09

Present: Margaret, Mary, Stevan, John

Regrets:

Agenda

Go thru process to create a unit

https://github.com/EDlorg/Units-WG/tree/main/Creating_units/Guides

Try Mary's idea of using chatgpt

<https://chatgpt.com/share/67f56884-a45c-800f-ac85-2110c46d1912>

Issues with 1) quantityKind and 2) Dimension Vector

I still have not come up with a way to record the units we've added (or if this is necessary).

Maybe something like this registry:

<https://github.com/EDlorg/ecocomDP-dataset-management/tree/main/queue>

Notes

Need to decide on cutoff for how rare things should be. At what point will new units be added by request?

Chatgpt looks promising! It has not gotten the hasDimensionVector field correct yet. Mary will try to refine the chatgpt output. She needs some example ttl to see that chat gtp is getting it right

These three units were added by us,

unit:MilliMOL-PER-M2-HR

unit:MicroGM-PER-L-DAY

unit:MicroGM-PER-GM-DAY

Below is the ttl example, from qudt file at

https://github.com/EDlorg/qudt-public-repo/blob/main/src/main/rdf/vocab/unit/VOCAB_QUDT-UNITITS-ALL.ttl

unit:MilliMOL-PER-M2-HR

a qudt:DerivedUnit, qudt:Unit ;

dcterms:description "Hourly molar flux equivalent to .001 moles per meter squared per hour, typically used to express an areal production (if plane is a surface of the earth or ocean).^^rdf:HTML ;

qudt:applicableSystem sou:CGS ;

qudt:applicableSystem sou:CGS-EMU ;

qudt:applicableSystem sou:CGS-GAUSS ;

qudt:applicableSystem sou:SI ;

qudt:conversionMultiplier 0.0000002778 ;

qudt:hasDimensionVector qkdv:A1E0L-2I0M0H0T-1D0 ;

qudt:hasQuantityKind quantitykind:MolarFluxDensity ;

qudt:plainTextDescription "Hourly molar flux equivalent to .001 moles per meter squared per hour, typically used to express an areal production (if plane is a surface of the earth or ocean)."

qudt:symbol "mmol/m<C2><B2>/hr" ;

qudt:ucumCode "mmol.m-2.h-1"^^qudt:UCUMcs ;

rdfs:isDefinedBy <http://qudt.org/2.1/vocab/unit> ;

rdfs:label "Millimole per Square Metre per Hour"@en .

unit:MicroGM-PER-GM-DAY

a qudt:Unit ;

dcterms:description "mass ratio as 0.000000001-fold of the SI base unit kilogram divided by 0.001-fold of SI base unit kilogram per day"^^rdf:HTML ;

qudt:applicableSystem sou:CGS ;

qudt:applicableSystem sou:CGS-EMU ;

qudt:applicableSystem sou:CGS-GAUSS ;

qudt:applicableSystem sou:SI ;

qudt:conversionMultiplier 0.0000000001157407407 ;

```

qudt:hasDimensionVector qkdv:A0E0L0I0M0H0T-1D0 ;
qudt:hasQuantityKind quantitykind:MassSpecificBiogeochemicalRate ;
qudt:plainTextDescription "mass ratio as 0.000000001-fold of the SI base unit kilogram divided
by 0.001-fold of the SI base unit kilogram per day (8640 seconds)" ;
qudt:qkdvDenominator qkdv:A0E0L0I0M1H0T1D0 ;
qudt:qkdvNumerator qkdv:A0E0L0I0M1H0T0D0 ;
qudt:symbol "ug/g/day" ;
qudt:ucumCode "ug.g-1.d-1"^^qudt:UCUMcs ;
qudt:ucumCode "ug/g/d"^^qudt:UCUMcs ;
rdfs:isDefinedBy <http://qudt.org/2.1/vocab/unit> ;
rdfs:label "Microgram per Gram per Day"@en .

```

unit:MicroGM-PER-L-DAY

```

a qudt:Unit ;
dcterm:description "A rate of change of mass of a measurand equivalent to 10-9 kilogram
(the SI unit of mass) per litre volume of matrix over a period of 1 day."@en ;
qudt:applicableSystem sou:CGS ;
qudt:applicableSystem sou:CGS-EMU ;
qudt:applicableSystem sou:CGS-GAUSS ;
qudt:applicableSystem sou:SI ;
qudt:conversionMultiplier 0.00000000001157407407407407 ;
qudt:hasDimensionVector qkdv:A0E0L-3I0M1H0T-1D0 ;
qudt:hasQuantityKind quantitykind:MassConcentrationRateOfChange ;
qudt:symbol "<C2><B5>g/(L<C3><A2><E2><80><B9><E2><80><A6>day)" ;
qudt:ucumCode "ug.L-1.d-1"^^qudt:UCUMcs ;
rdfs:isDefinedBy <http://qudt.org/2.1/vocab/unit> ;
rdfs:label "Micrograms per litre per hour"@en .

```

2025-03-12

Present: John, Margaret, Mary, Marina, Stevan
 Regrets: Christine (on vacation)

Agenda

Prioritize units to create (using John's latest match results)

~~Go thru process to create a unit~~

In John's spreadsheet (below, "Unmatched units"), classify each unit as
 Y = will be added to QUDT

S = adjust substitution rules to transform this unit

X = discard for now. Includes some units that may be added after we get some advice from QUDT managers.

We got through about half the spreadsheet, and will continue, hoping to finish by the April meeting.

Materials:

Unmatched units:

<https://docs.google.com/spreadsheets/d/1bi8ZA8JfRIiVYEGTCtq0iM3u3Et9QJKWWBXSIE5JpPE/edit?gid=0#gid=0>

John's notes:

Thanks to improvements in QUDT (now version 3.0) we are up to 92% overall (up from 91%), with NEON up to 97% and EDI up to 95% (DataONE still lags at 89%), without any new substitutions - just due to additions to QUDT. I've limited the list to units used 5 or more times, but sorted it to go from most uses to least uses.

The qudtUri column is always NA (otherwise it would have matched QUDT) and the qudtUnit column is the guess for what the QUDT unit might be after the string substitutions.

Some can probably be addressed by improving the string substitutions (e.g., nanomolepergram which came out as NanoMOL-e-PER-GM (getting rid of the -e- would fix it), and g*m2. Or by eliminating some entities (weight-PERCENT).

Notes and how tos

https://github.com/EDlorg/Units-WG/tree/main/Creating_units/Guides

I still have not come up with a way to record the units we've added (or if this is necessary).

Maybe something like this registry:

<https://github.com/EDlorg/ecocomDP-dataset-management/tree/main/queue>

2025-02-12

Present: Margaret, Mary, Stevan, John
Regrets

Agenda

Announcements:

1. John et al paper accepted by Sci Data.
 - a. Need to revise citation for <https://ceur-ws.org/Vol-2969/paper10-s4biodiv.pdf>
 - b. Remove: For example, although KiloGM-PER-KiloM2 is a QUDT unit, KiloM2 was not (yet) a unit.
2. Our new units are in the latest QUDT release.

New units next steps

define additional units (with QKs, DVs where needed).

Updated list of units to be defined.

- Rerun John's match code
 - Need list of QUDT units that need to be added to transformations e.g, issue with feet_us
- Organize
 - Git issues - in our fork. Also would record our work.
 - One issue/unit? Decide on a git practice, write some guidance.
 - No. keep fork-issues for specific items about units we're working on.
 - It is helpful to keep some notes while working on a group of units. These could be stored in git also, but not the fork.
 - Decision: use our working group's git for this.
- How-to docs - Feedback please?
 - Margaret's drafts: narrative:
<https://docs.google.com/document/d/1pjc-BRHGZkS2JVnhNfQ-fYL7h-XH6xHWal778PIRRzk/edit?tab=t.0>
 - Flowchart:
https://docs.google.com/drawings/d/1-MrSh3Kov0ik88KDWBjCpqS5YbA79K-z_e8zoSmUQ6U/edit

Next Steps:

- Prep for march:
- Margaret - organize git for notes and how tos (done by March 1)
- John - rerun match code (done by March 1)
- All - review list of to-do-units, read how tos

2025-01-21

Present: Margaret, Steve Ray

Agenda: Ad hoc mtg, to go over our units submission.

Instructions (from his email):

Repo has a src folder when we make changes, followed by a build process that generates the target folder contents. Please take a look at our updated instructions at <https://github.com/qudt/qudt-public-repo/wiki/Installing-QUDT-for-Contributors> and at <https://github.com/qudt/qudt-public-repo/wiki/GoodGitPractices>. The first link implies that you need to install Maven locally, but the second link explains how you can avoid that by using the workflows available on GitHub.

For now, we will use the GitHub workflow. Will try it out for our first pull request.

<https://github.com/qudt/qudt-public-repo/pull/1098>

Other notes from Steve:

QUDT is considering domain profiles, e.g. filters to subset the entire list. Possibly they will use QKs for that. EDI could make use of that, but it means lots of QK definitions.

Showed him Tim's QUDT browser. This pulls from the website - good, since that is updated with each release (monthly). Latest snapshot has interim changes (loc?)

Things to follow up on w/Steve

How to find the validation report (happens during push)

How to find the results of the Github action > "Format sources and commit to branch" (I just get a yes/no)

Draft, how-to-new-unit:

Doc: [Create Unit notes](#)

flowchart: [Creat_Unit_workflow](#)

2025-01-15

Present: M. O'Brien, S. Earl, J. Porter, M. Frants, M. Martin

Regrets:

Agenda:

Announcements?

MS turned in Jan 5, per John
Jan. 14 - MS sent off to reviewers for their feedback

Quantity kinds

Margaret's progress:

List of unmatched units:

<https://docs.google.com/spreadsheets/d/11PRwbHTLXSgby1lz591gWfyKLaR4QOijB4rNsK8svuM/edit?gid=208332157#gid=208332157>

I used this spreadsheet to examine datasets and determine best fit QK (col G).

TO DO: go over these, especially those in color

Then, looked at our list of current units to add, to pick out high-priority QKs.

QKs needed for our current list of new units, with defs:

<https://docs.google.com/spreadsheets/d/1b1pGa8Dd1IX-ill4tMWtq39vUyGJqFaaZK3dIEh3Ov0/edit?gid=0#gid=0>

Topics:

1. Something seems wrong, because interestingly, there are units we plan to add that don't show up on the unmatched list (searching on unitLC col, which I think is pseudounit).

What is going on?

unit:MegaGM-PER-HA-YR

unit:MilliGM-PER-GM-HR

unit:MilliGM-PER-KiloGM-Day

unit:NanoMOL-PER-GM-HR

unit:MicroGM-PER-L-DAY

unit:MilliM-PER-M2

unit:MilliMOL-PER-M2-HR

unit:NanoMOL-PER-M2-SEC

John: PROBABLY CAME FROM DATAONE, WHICH DID NOT HAVE OTHER ATTRIBUTE INFO.

2. I don't understand the NEON error units. Christine, can you elaborate? Variance is square of entire unit, so confirm my changes.

- a. unit:DEG_C2

- b. unit:MicroS-PER-CM2

- c. unit:MicroMOL2-PER-M2-SEC2

- d. Here is Steve Ray's annotations on those:

- i. https://docs.google.com/spreadsheets/d/1_MHWm5mN7i0jTsar8Uuipr5nQVAJuTCh/edit?gid=1917304109#gid=1917304109

3. one of our new units was KiloM2, which actually exists as a unit (not a problem, except that I think we said it didn't exist in the paper!)

Decisions:

EDI should (eventually) use a hierarchy of measurements from ENVO, not the QKs from QUDT. ENVO is where a hierarchy of specific processes (like N-cycle) belongs. Advantages: simplifies EDI's tasks (only one ontology to search), and our task, because we don't have to pick a perfect match for QK, just find the best applicable, and add a new one when needed. Keep new QKs general. Adding new ones is good-citizen behavior, pick "unknown" when you don't have access to the dataset.

Margaret tasks:

1. ~~In spreadsheet~~
<https://docs.google.com/spreadsheets/d/1b1pGa8Dd1IX-ill4tMWtq39vUyGJqFaaZK3dIEh3Ov0/edit?gid=0#gid=0>
 - a. ~~Make sure none of the proposed QKs (col F) are already present, that we have just missed due to naming idiosyncrasies. Look at existing QK definitions: Example is PhotosyntheticPhotonFluxDensity, which is almost MoleFluxDensity, but not quite, so we will add MoleFluxDensity for things like MilliMOL-PER-M2-HR.~~
2. ~~Add/define new QKs to the QK file.~~
3. ~~Assign QKs in Units file~~
4. ~~validate~~
5. ~~Create pull request~~
6. ~~Call Steve.~~

Steve Ray questions

Is qudt:applicableUnit populated with code? Or do we put that in

2024-12-11

Present: Margaret, John, Marina, Mary, Stevan

Regrets: Christine

Agenda

~~Change in meeting time? (per Stevan)~~

Paper edits

Quantity kinds

Reviewer comments:

Please provide a written response to ALL requests in your 'response to reviewers' document to avoid potential delays.

Editor comments (none present if this section is blank)

* Please perform two standard checks, which we request for all datasets making use of input data from third parties:

1) Please cite all your sources in the reference list - preferably at the dataset level, and/or any relevant publications if more relevant (the data providers often provide guidance on how they wish to be acknowledged). For resources that do not have formal metadata, please quote the URLs in the text when they are mentioned in ()s.

2) Confirm all the sources are openly available - i.e. your re-use or re-distribution is compliant with the terms and conditions/licenses for data sharing (again, the data provider should provide guidance on this).

* Please remove the section numbering.

* Please move the data availability section to up to follow the discussion.

* Please add data citations for the datasets on repositories to the reference list (see https://www.nature.com/sdata/publish/submission-guidelines#data_citations). Please add the reference numbers to wherever the datasets are mentioned in the text - the main position should be the first part of the Data Records in a sentence describing where the data have been deposited.

Reviewer #1 (Remarks to the Author):

In their study "Using a Units Ontology to Annotate Pre-Existing Metadata", Porter et al. present an approach to link diverse representations of units to standardized units in the QUDT ontology using string transformations. By their approach they were able to map 91% of the unit instances from a large corpus of ecological metadata. This mapping enables better dataset comparison and annotation by providing consistent unit definitions and related information. In addition, they provide a resulting tool as a web service and R functions for annotation of Ecological Metadata Language (EML) documents.

Overall, the review is very well written and provides a good overview about the mapping procedure.

I have a few minor remarks as well as two major revision requests with respect to the developed web service:

Minor remarks.

Lines 146-147: How exactly was this done? Did you consider using machine learning approaches for this purpose?

Lines 173-175: Here, I am missing detailed descriptions how the web service was developed.

Lines 176-180: The manual checking section could benefit from more details on the process and the feedback from the data managers.

Line 213: Spelling mistake: Replace “wiil” by “will”

Lines 224-225: Are they already mentioned somewhere? In a table? How will this be achieved in the future?

Lines 260-261: This sentence could be re-phrased for clarity.

Major remarks:

I am missing more details regarding the setup of the web service and instructions on how it can be embedded by users. Additionally, it would be helpful to include further information about the service (such as the link) in the manuscript. The web service could also be significantly improved by adding more user-friendly features, such as a direct link to the results in the QUDT database, instead of a link that needs to be manually copied.

In addition, the link <https://vocab.lternet.edu/webservices/unitsws.php> is not working.

Figure 2: As reader, I would expect here a graph that is easier to understand at first glance.

Reviewer #2 (Remarks to the Author):

This work is highly relevant in providing methods for converting uncurated and situation-dependent unit representations into machine-readable units to enable automated processing of large data corpora. Machine readable units are essential for making data FAIR. Together with FAIR descriptions based on observational frameworks this will help to make research much more reusable.

The context of the problem is well introduced, and relevant unit representation frameworks presented. However, I recommend including recent efforts related to the SI Digital Framework and discussing its relevance in the context of the other frameworks. The methods applied are comprehensively and reproducibly documented, and the data outputs for each process step are provided in the EDI repository. Manual inspections by 24 data managers were conducted to verify the quality of the output. The work also contributed to the identification of missing unit references in QUDT.

A REST-based web service to provide semantic annotations based on the lookup table is mentioned in the paper, but the link to the service is missing. The I-ADOPT Framework is referenced in the paper, but its name is misspelled. Fig. 4 showcases how additional

annotations could be included in the EML metadata of the measured attribute and link to entities, units but I recommend to also add the quantity (property).

Corrections:

line 16 - a single complex unit (m/m² is a complex unit)

line 213 - typo: These will

2024-11-13

Present: Mary Martin, Colin Smith, John Porter, Jon Ide, Mark Servilla, Greg Mauer, Marina Frants, Corinna Gries, Stevan Earl

Regrets:, Christine Laney

Agenda

EDI discussion with Mark and others

Summary info:

<https://github.com/EDlorg/Units-WG>

<https://github.com/EDlorg/qudt-public-repo>

Paper (submitted in August, Mark and Paul have a draft)

Topics:

1. If needed, an intro to QUDT
 - a. What
 - b. Who
 - c. Why
 - d. Browse: https://ble-lter.github.io/QUDT_Browser/
2. WG's main products:
 - a. look up table, entity #6 in <https://portal.edirepository.org/nis/mapbrowse?scope=edi&identifier=1715>
 - b. New units for QUDT: <https://github.com/EDlorg/qudt-public-repo> (our fork) - anticipated done by end of 2024
 - c. John's code: https://github.com/EDlorg/Units-WG/tree/main/RCode_JP
3. Incorporation into EDI, e.g., their plans for
 - a. EML annotations
 - b. ezEML
4. Units WG 2.0 plans (2025 ->), e.g.
 - a. Additional unit submissions
 - i. ezEML
 - b. Lookup table maintenance (eg., after QUDT additions)
 - c.

Units vetting progress and questions:

Follow-up: Created list of camel-case preferred units and UCUM codes at:

https://github.com/EDIdorg/Units-WG/blob/main/RCode_JP/DataFiles4R/unitsPreferredNamesWithUCUM.xlsx

Notes

Recording and chat text is here (sibling to this doc):

https://drive.google.com/drive/folders/1wJl0y1ewN10M5W8Q-9l6jM_uKX7BYRUy

Feedback from EDI participants:

Questions:

Where is the lookup table in git? Is there an api method to return it?

https://github.com/EDIdorg/Units-WG/blob/main/RCode_JP/DataFiles4R/unitsWithQUDTInfo.csv

No api method.

Also in the dataset

Why would you want to use QUDT in the first place

- LTER IMs asked for something to harmonize units
 - Qudt had the most complete list, appears to have stable support
- You can do more sophisticated [things] with measurements if you incorporate QUDT, using it's quantitykinds and dimensions
 - Improved search (if you have adequate QKs to group datasets) - you can say "find areal primary production measurement", and connect to dataset via their unit
 - Dimensions allow conversion

Important to remember that standard strings help, either the camelCase spelling we use, or the QUDT-style (maybe UCUM?)

John: user experience (e.g., in ezEML)

Browse, using QK

Some sort of autocomplete, smarter than just starting at the start of the string

Many IMs (or assistants) might not know how to reduce a unit (if that's all they get) to a QK so they would probably use an autocomplete.

Is there a preference on where the annotation should appear? I.e., attribute level, additionalMetadata list, annotation section
Not at present

Demo from Colin:

Before we leave today, I'd like to share a little work I've done to wrap J. Porter's web service for annotating existing EML files.

[Here it is incase we don't get to it:](#)

<https://spinneret.readthedocs.io/en/latest/user/quickstart.html#adding-qudt-annotations-to-an-eml-file>

This is very much a WIP and will likely change moving forward.

Two areas of exploration

EZeml: Jon - ways to incorporate QUDT need to be explored. John has ideas for shortening a browse list

Shadow metadata: Colin - can we add QUDT to schema.org metadata

PASTA - not likely to incorporate annotations into L1 anytime soon.

EDI will discuss this topic during their developers meetings. Anticipating more meetings with them, particularly Colin (schema.org) and Jon (ezeml)

Unit vetting progress

Collecting unit questions here:

https://docs.google.com/document/d/1UuORxt7_2aJhukHQBSez-NOdChtLU7qwsP5Y4rExFI/e/dit?tab=t.0

2024-10-09

Canceled. Agenda moved to November.

2024-09-11

Present: John Porter, Margaret O'Brien, Stevan Earl, Marina Frants

Regrets: (Mary, Christine)

Agenda

Response from John's presentation to LTER IMC

Oct-Dec topics- what's next (2025)?

Mechanism for accepting new units in future

quantityKind (examining measurements)
EDI integration
Annotation QC

Notes

Report from John

Some issues with R and character sets. For odd characters, wrong output from XML
R-package: <u... It might be related to UTF 8, but it happens on EML that is in ascii too.
Stevan says its the BOM in windows
(<https://stackoverflow.com/questions/39593637/dealing-with-byte-order-mark-bom-in-r>).

Units submission (mob report)

- Progress on vetting, with topbraid composer. See git repo for commits.
- Most errors in our units are missing/wrong dimension vectors and/or quantityKinds.
 - Dimension vectors are relatively straightforward
 - QKs less so: you need to know about the entire measurement. I have not added any new QKs at this time, using “Unknown”

This git repo: <https://github.com/EDIdorg/qudt-public-repo>

Per John: We may also want to look at the site listing:

<https://docs.google.com/spreadsheets/d/17UMklPudPptd3RYvQIC4Di-o40vs8m88/edit?usp=sharing&ouid=105733120713818056674&rtpof=true&sd=true> for others to consider

Look at this, and determine if there is value to pulling more out of this. For the units submission (2024). Otherwise, they would go into a 2025+ submission

UnitsWG 2.0 projects

1. quantityKind and measurements
 - a. What is it valuable for?
 - i. Subsetting units? Might help if using a pick-list in a web form
 - ii. Annotating measurements – A future project.
 - iii. Advanced stuff – QK's must match in order to perform unit conversions. Dimension vector isn't enough
 1. L-1 example - complex ratios reduce to 1/m, (optical oceanog, light attenuation, k_d) and one of ours to be submitted (MilliM-PER-M2, measurement unknown)
 - b. More complex knowledge modeling coupled with scientific expertise will be needed for QK work.
2. New-units vetting

- a. Anticipating intermittent work, easiest if done in batches, e.g., potential new units reviewed every 6 months
- b. Aspects TBD
 - i. form for accepting new contributions
 - ii. can QC be done with form-validation, or does the committee need topbraid composer?
 - iii. Integration of new QUDT units into our lookup table
 - iv. Integration of lookup table into [whatever EDI decides to do with it]
- 3. Entraining non-LTER data managers? (eg, other EDI or D1 contributors)
 - a. Outreach to expand involvement, and
 - b. Broaden scientific expertise (especially if we take on quantityKind and measurements)

Action items

~~Write to Mark, invite to October meeting (calendar invite):~~ done.

2024-08-14

Present: Margaret, Marina, Mary, Christine, Stevan

Regrets: John (vacation)

Agenda

Paper

See your email for text

To be submitted 8/20 (when John returns)

Comments?

Dataset:

<https://portal-s.edirepository.org/nis/mapbrowse?scope=edi&identifier=1715&revision=1>

Margaret's TO DOs before it goes to production server

- ~~1. Add John to access-edit list (didn't find a user for john individually, so used VCR)~~
- ~~2. Remove ref to webservices in the R-entity; those are php/html and not included.~~
- ~~3. Reorder entities~~
- ~~4. Rm prefixes from entity names~~

Notes

MS comments here:

manuscript

Line num	comment
----------	---------

6	In affiliations, Scripps should be listed as in La Jolla rather than San Diego
Header pg	“John Porter” listed twice in authors
throughout	check to be sure that new text uses a single space between sentences in keeping with the format overall
20	“Ad hoc” is italicized in the text, so do that in abstract, to be consistent
100	do not think that we need F and M initials in this context: Keil, J.M. and Schindler, S.
110	I think it should say “seems” rather than “seemed.” (Since it’s still true, presumably).
128	I’d put a comma after “program”, I think it makes the sentence clearer.
130	Dataset ref (I added it to Zotero) Porter, J., M. O'Brien, M. Frants, S. Earl, M. Martin, and C. Laney. 2024. Mapping EDI, NEON and DataONE units to the QUDT ontology, 2022 ver 2. Environmental Data Initiative. https://doi.org/10.6073/pasta/03de5f726541999caeed8ef2a55e8c3c (Accessed 2024-08-14).
136	Fig 1 caption mentions the dataset and github. The dataset ref should replace that
151	Delete stray comma in middle of line (after “pseudounit string”)
155	Probably best to avoid URLs in text. Either leave it out, or use the general QUDT ref (#12) (IMO, the URIs are fine to leave in, even tho these are URLs.)
169-170	Dataset ref goes here. No git urls. No ref to qudt ttl file. Mention of QUDT version number is enough.
172-173	Delete ‘be’ ie, should be “..could potentially be matched...”
183-184	Should be “were developed”; also either “inspection...was performed” or “inspections...were performed”
251/252	Reword? Two tools that draw upon the lookup tables to facilitate adding QUDT units to EML metadata were created as annotations Suggested: Two tools were created that draw upon the lookup tables to facilitate adding QUDT units to EML metadata as annotations.
226	Remove space between “QUDT” and “.” at end of sentence.
294	Incomplete sentence (delete “We therefore compiled a list of.”)

381	Please add to the Acknowledgement section: "The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the National Science Foundation through the NEON Program." This is in the google doc as well.
191	DataOne (45.8%) and from NEON...to... DataOne (45.8%), and NEON
200	Other tables & figs have the label as part of the caption (below the object). If that is the pattern the journal wants, add "Figure 2" to front of line 200
230	Text is "Tab. 1.". Is this a reference to table 1? If so, it belongs on line 229. Presuming table 1 caption is below table.
255	Should that be Fig. 3, not 4?
294	Sentence frag, see para below: We therefore compiled a list of. QUDT
299	Use "percentage" rather than "percent"?
310	"There remain" can be dropped - just use "8% of the units may never..."
310	Reword "In many cases, these rarely used," - a bit confusing. Perhaps just concatenate the two sentences "In many cases, these rarely used, unmatched units are nonsensical. For example, what unit is meant by "4a" or "atlantis/alvin" or "martha stewart" or -9999?" to "Many of the unmatched units are nonsensical, such as "4a", ..., and -9999."
312	dangling quote: these terms" actually
313-316	Per an email conversation, we were going to drop the highlighted text (comment 3) and combine the rest of the paragraph (lines 310-313) with the first paragraph of the discussion. (for what this looks like, see text below these tables of comments)
319	Ref should be to the QUDT docs, #12
334	Ref is (S. Ray, pers. comm.)
347	Add period to sentence, after "test the models". No ref needed.
349	Ref should be to Figure 4b ("not Fig 3, bottom")
351	Extra space between "a" and "collection"
362	Add "and" before last element in series: "...iAdopt, and encourage development

	of extensions...”
381	To list of grant numbers, please add after the NSF grants “, NASA grant NNX14 AR62 A and BOEM grant MC15 AC00006”

Submission doc

Contributing authors	Add orcid (see google doc, top table) Martin: https://orcid.org/0000-0003-1791-5584
Margaret's entry	Funders: Please add 2 more funders (in the dataset) Funder: NASA Funder ID: https://ror.org/027ka1x80 Number: NNX14 AR62 A Title: Santa Barbara Bight Biodiversity Observation Network Funder: Bureau of Ocean Energy Management (BOEM) Funder ID: https://ror.org/03tzscr25 Number: MC15 AC00006 Title: Santa Barbara Bight Biodiversity Observation Network Remove 1637632 and 2224726, add 1831937

Re first para of discussion (lines 313- 316, above. Sentence frag also rm from line 294)

Despite the generally successful mapping of existing units in the three-organization corpus to QUDT, we encountered some gaps in the QUDT coverage. For example, although KiloGM-PER-KiloM2 is a QUDT unit, KiloM2 was not (yet) a unit. QUDT has a well-defined methodology for proposing additional units, and we are developing submissions to help eliminate the gaps we identified. We estimate that the proposed additions to QUDT will add an additional 1-2% of coverage. There remain ~8% of units that may never be matched to QUDT. In many cases, these rarely used, unmatched units are nonsensical. For example, what unit is meant by “4a” or “atlantis/alvin” or “martha stewart” or -9999? In many cases, we believe these terms actually belong in other parts of the metadata (e.g., methods, missing values, lists of codes and locations).

PERMISSIONS

Margaret,

I understand that you would like to include a personal communication in your upcoming publication, in the context of qualifiers for units. You shared your intended quote as "These are generally for legacy units from QUDT's origin, and this practice is not currently recommended because such a context-specific specialization edges toward the realm of a qudt:Quantity rather than a dedicated unit (S. Ray pers. comm.)"

You have my permission to quote me on this topic. Best wishes on the publication!

Steve Ray
CEO, QUDT.org

2024-07-10

Regular meeting canceled, replaced by out-of-band mgt last week.

2024-07-03

SPECIAL MEETING TIME 1 PM EDT, noon CDT, 11a MDT, 10a CDT

Present: Margaret, Stevan, John, Marina

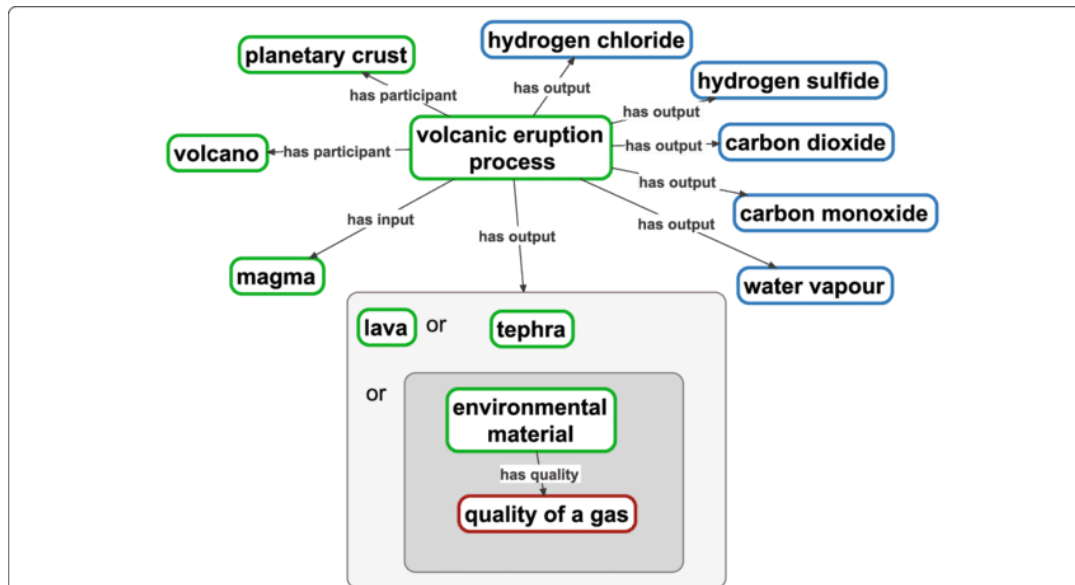
Agenda:

Work on paper at:

https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit?usp=sharing

Discussion:

Last figure - hard to read? Better an RDF triplet? See Stevan fig. has example that includes ofEntity and hasContextObject as well as hasUnit



Remove existing figure 5 and move Fig 5 to discussion for new RDF-style diagram including other elements

Actions:

John will reconcile terms based on new list at start of Methods

Margaret will draft new figure 4 in the Discussion - RDF Triples (be sure to BOLD the unit stuff :)
)

Margaret and John will work on assembling DATASET to go with paper.

Wed 7/10: updated draft.

Decide on Tuesday if we'll have our regular meeting (would have been 7/10).

Marina and Stevan out: week of wed July 17.

2024-06-27

Present: Margaret, Steve Ray.

Out of band meeting to fix my top quadrant issues. Naturally, I could not reproduce the start-up errors for him (a good thing, right?), so we went on to discuss some other things. These notes are distilled from that discussion.

Christine's spreadsheet, with Steve's input:

[link]

Some specific notes:

Units that are variance measurements: there is no way to indicate arithmetic, so these will need to be named to fit the pattern, e.g, CM3-PER-CM3-2 should be CM6-PER-CM6

These will also need a new quantityType, to indicate that is a VarianceOf (variance of some measurement)

EQ-PER-L is a concentration of charge, and the unit is MOL-PER-L. (see more about quantities and quantity kind below)

From Steve to Christine, via email:

I sifted through your spreadsheet and put some feedback in column J.

Also cc'ing Margaret, who reconnected this week.

Regarding rows 94 and 95, try pasting the following query at

<https://www.qudt.org/fuseki/#/dataset/qudt/query>

```
SELECT *  
WHERE {  
  ?qk skos:broader* qudtqk:DimensionlessRatio .  
}
```

The result shows all our current specializations of dimensionless ratio quantity kinds. You could invent some new ones if you want.

Quantity and QuantityKind

Quantities and the entities involved are “at the edge” of QUDT (e.g., chemicals, substrates, etc). there is a tendency for people to want to wrap everything about a measurement into the unit, and that is incorrect. The quantity is basically the measurement class, and we don't have a good ontology for those (the measurement itself). Quantities are in the science-domain, and are context for the unit. There are frameworks these could be built on, but that hasn't gotten very far.

What we can do now is to make sure that the quantityKinds in QUDT are correct, and will match to both the units we have, and the datasets they came from. We have not looked at the datasets (measurements) beyond the units yet, but we may have to.

The part of the quantityKind that has to be correct is its dimension vector. It's main use is conversion. (no match > no conversion)

Notes on some of the quantityKinds we are interested in.

“Mass-concentration” is not a thing. (even though in chemistry it is thought of as a ‘type of concentration’ even in QUDT def! Fix that, please.). so from a unit POV, do not call mg/l a concentration. It's a density.

mob comment: I still find this irritating. First, because I don't know how to steer people to the right quantityKind. Maybe it's from the quantity side of things. (or maybe most people never have to see that part. somehow.)

Second, the fact that for our use, most of the time something like mg/l is referring to a mixture (not the density of a substance), and that there are different things in the numerator and denominator. Those differences go into the Quantity (not QuantityKind, and not unit!).

Example of quantities, numerator and denominator

Steve gave me an example of another rules-based ontology (proposed standard) he is building for the building industry. It does have some examples of quantities, and the attached rules: E.g., a cooling system must have a medium, e.g., hasMedium water or something like that (screenshot here)

Link: <https://open223.info/>, <https://explore.open223.info/brick/Quantity.html>

Scales:

Generally, these are dimensionless. Because you can't do conversions with them, unless you know lots of context. That context could be in the unit description or references, but it is not encoded.

pH - has a unit. dimensionless

Reynolds number - has a unit, also dimensionless

Eventually, we will get to water density (seawater, sigma-theta), which is a deltaQuantity - e.g., the actual density (ρ) of water is 1.024123, and sigma-theta is $(\rho-1)*1000$ so what is reported is 24.123. In datasets it should be called a "density anomaly"

Steve likes that we plan to require hasNormativeReference, (or at least hasInformativeReference) in all new units.

General QUDT mgt

Board currently is 4 people

Don't remove, deprecate.

If deprecated, replace.

Note that the first units to go in were from NASA, so that's why you see lots of stuff about rocketry that is probably over-specified for a general units ontology. (one wonders... was this in response to the mars crash? That would be cool.)

2024-06-12

Present: John, Margaret, Marina, Mary, Stevan

Regrets:

Agenda:

Paper:

https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit

Notes:

Worked mostly on figure 1.

John will incorporate our comments to date, for recirc next week.

Goal: submit by July 1.

(Margaret has had no luck with topbraid, and is setting up a chat with Steve Ray).

2024-05-09

Present: Margaret, Colin Smith, Corinna Gries, Jon Ide

Out of band meeting with EDI

Margaret updated EDI on WG activities. Brief notes here.

Showed them (or emailed as follow up):

1. Paper draft, workflow fig 1:
https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit
2. QUDT sample content (ours):
<https://github.com/EDIdorg/Units-WG/blob/main/QUDTcontributions2023.ttl>
(unit:MicroGM-PER-L)
3. Lookup table:
https://github.com/EDIdorg/Units-WG/blob/main/RCode_JP/DataFiles4R/unitsWithQUDTInfo.csv
4. John's web service demo:
 - a. HTML forms: <https://vocab.ltnet.net/unitsws.html>
 - b. Raw web service: <https://vocab.ltnet.net/unitsws.html>
5. Example of EML with annotation added via John's code (a screen search for "annotation" shows 3 instances):
<https://portal.edirepository.org/nis/metadataviewer?packageid=knbn-lter-hbr.20.12&contentType=application/xml>

EDI plans (still in development):

Location of annotations in EML metadata: they will be able to support annotations in any of the possible locations. But when EDI codes annotations, they probably will prefer to put all annotations into a group, under the <annotation> element (eml:eml/annotations, right above <additionalMetadata>).

The most complicated topic will probably be how ezEML handles QUDT integration. Questions include:

Would we still use the std/custom designations?

Would submitters have to do units twice?

Could parts of QUDT be used in custom unit fields?

What about data packages being updated (where they did units the old way, would these be migrated? how?).

What if the user typed in their favorite unit representation, could then ezEML offered the QUDT equivalent - (based on John's lookup table). If no exact match, offer close matches? (what criteria).

2024-05-08

Present: John, Margaret, Mary, Marina, Christine, Stevan

Regrets:

Agenda:

Paper:

https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit

Notes

Submit to Scientific Data (Nature journal)

- 90% of submissions appear to be descriptions of data (archived datasets). Only ~ 5% (171) are articles about data mgt. (another 100 are about uses of data)

Next Steps:

- ~~1. Margaret will work on Flowchart figure 1~~
2. John will clean up text and look at putting numbers into a table
3. Look at histogram figures - either make them more informative or eliminate them and use a table instead
4. Clean up terminology for unit instances, unique units, pseudounits and QUDT units, ie, consistent language, needs to be clear that
 - a. Ad hoc units, Instances, Unique units (all are "raw"; ad hoc units and instances are the same thing)
 - b. "Raw" data = from 3 sources
 - c. Lookup table is product of 2 major steps
 - i. creation of the pseudounit intermediate, mapping)

2024-04-10

canceled

2024-03-13

Present: Margaret, John, Christine, Marina, Stevan

Regrets: Mary

Agenda

1. Reports?
2. paper.

Notes

No commits since last month to the QUDT fork. Margaret has some problems with TBC. so no units-checking today.

Paper link:

https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit

Progress: John has updated the text considerably.

Everyone: edit freely

Deadlines: by next meeting (April)

Major editing done

Choose journal

journal	Page charge	audience	Impact factor
Ecological informatics	3690 (w/o UC 15% discount)	ecology	5.1 (since 2021) - https://www.sciencedirect.com/journal/ecological-informatics/about/insights
Scientific Data	2290		10 (5 year) - https://www.nature.com/nature-portfolio/about/journal-metrics
Gigascience	2512		9.2 https://oair.org/j.php?i

			ssn=2047-217X
Journal of Big Data	2k (1700 euro) (open access)		8.1 (2 year) - https://journalofbigdata.springeropen.com/
Journal of information and management	> \$3700	Very general - would need to tailor intro differently	9.9 - https://www.sciencedirect.com/journal/information-and-management
Journal of information science	None (but no OA)	Library sci	3.1 (5 year) - https://journals.sagepub.com/home/jis
Data Sci Journal	650 UK pounds		2 - https://www.scijournal.org/impact-factor-of-data-sci-j.shtml

Action items:

Paper edits

Look for projects that can pay page charges for us (EDI, NEON, local LTER site), invoicing mechanism

2024-02-21

Present: Margaret, John, Stevan, Mary

Regrets:

Agenda

- Progress using TopBraid Composer
- Tips for using text editors instead.
- Paper?

Notes

Highlights!

- We spent most of our time perusing TB composer and fixing errors.
- We cannot export the SHACL validator output.
- We still need a sustainable method for validating. But it may be enough that the working group knows what is allowed, and acts as a filter for any new units (e.g., which fields are

allowed, cardinality) and that we leave validation up to QUDT (Steve). Maybe by the time we have created 1-00-200 new units that will be true.

Process we will use for corrections

Everyone:

Clone our Github fork and Edit QUDT contributions there:

https://github.com/EDlorg/qudt-public-repo/blob/main/vocab/unit/VOCAB_QUDT-UNITS-ALL-v2.1.ttl

Steve Ray's preferred method is that we edit the main doc, and then do a pull request from our fork.

Our WG's units start approx line 35600 (find on page "EDI " to see that comment. Do not edit above this line (please - those are the already-vetted QUDT units).

For you all to edit, Margaret or Stevan (admins) need to add you to the team. We need git account names for Sarah, Christine and Marina.

To decide which units to work on, use the most recent screenshot of errors in this dir:

https://drive.google.com/drive/folders/1iZsPqTM2_jA4S4HiB-43IHZ29ngI7RyB

Path: EDI/UnitWG/QUDT_contributions_in_progress/SHACL_error_reports/

Validation:

Margaret will run validator periodically and add a new screenshot. This method will work for the short term, but we need something better.

Potential new fields to add:

normativeReference

informativeReference (e.g., for a wikipedia or other entry describing the unit)

2024-02-05

Present: Steve Ray, Margaret, Marina, Mary

Ad hoc meeting with Steve Ray, to figure out TB Composer.

No digital notes, as we were too busy. Some paper notes

2024-01-24

Present:

Regrets:

Agenda

go over progress with TopBraid Composer (TBC, TBCFE)

Notes

2024-01-10

Present: Stevan, Marina, John, Margaret

Regrets: Sarah

Agenda

Catch up.

Action items

1. Install the software (6.0 or 6.1, as appropriate)
 - a. Get familiar with it (group mtg)
 - b. Set up a short mgt with [Steve Ray](#) - to make sure we're doing it right, efficiently.
Specifics tbc
2. Comments to John's draft
3. Databits article (general sci audience)

2023-12-13

Present: Steve Ray (QUDT), Margaret, Steven, Christine, Sarah, Li Kui, John

Regrets: Marina

Li is here representing imexec, to get a sense of what the working groups are doing, and report back.

Agenda

1. Guest - Steve Ray, from QUDT
2. Feedback on our submission
3. Paper
 - a. https://docs.google.com/document/d/1KXe5QAADXDXQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit?usp=share_link
 - b. Overall layout/themes
 - c. Present our EML annotation as a use case

- d. Writing responsibilities?, e.g.,

Notes

1. Guest - Steve Ray, from QUDT
 - a. introductions
 - b. Our assembled questions
 - i. Quantity
 - ii. QuantityKind
 - c. SHACL language used as a way to validate QUDT
 - i. TopBraid Composer (free version) version 6.1 helpful
 1. Uses rdf:type shortcut (letter a)
 2. EDG "Edge" is the new commercial tool from TopQuadrant
 3. Hidden download link on TopQuadrant web site for TopBraid
<https://archive.topquadrant.com/topbraid-composer-install/>
 - d. Note, if you alter the URL for a unit by adding .ttl it will give you the Turtle file
 - e. You can use Pull request to create new branch where you can append to existing units file
 - i. May be easier to work with a full version of the file
 - f. Wolfram Alpha model for units also good, but it is not FAIR (expensive)
 - g. Discussion of common names with _qualifier
 - h. Discussions of grams carbon etc.
 - i. quantityKind can't have a value
 - i. Discussion of quantityKind in reference to Fluxes
 - i. Many different kinds of fluxes
 - ii. Can also use altLabel (only one label per language, but no limit in altLabel)
 - j. Applicable units - in Advanced User Guide
 - i. Units point to most general quantityKind (where all children are applicable)
 - ii. Applicable quantityKind goes up quantityKind tree until it finds a quantityKind that has at least one unit. quantityKind further down the list won't be listed as applicable
 - k. Discussion of numerator and denominator dimension vectors
 - i. Still under development
 - ii. Note, if dimension vector ends in D1 should not have any other dimensions
 - iii. dimensionless/time has dimension T-1 and D0 not D1
2. Feedback on our pull request (if available)
 - a. Presuming we will not have heard back from QUDT, given Thanksgiving week and AGU week.
3. Paper

- a. https://docs.google.com/document/d/1KXe5QAADXDQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit?usp=share_link
- b. Overall layout/themes
 - i. Present our EML annotation as a use case
- c. Writing responsibilities?, e.g.,
 - i.

Action items

Steve will send Margaret the link to topbraid composer (free, 6.1). We will set up a session with him to work thru our issues. 6.1 will probably be only for mac, so others will have to get the free 6.0 download from link below

Chat notes:

John Porter (he/him) to Everyone (Dec 13, 2023, 1:22 PM)

<https://github.com/EDlorg/Units-WG/blob/main/QUDTcontributions2023.ttl>

Christine Laney to Everyone (Dec 13, 2023, 1:26 PM)

Somehow I missed that CentiM3-PER-CentiM3 already existed ... can take it out of our ttl file

You to Everyone (Dec 13, 2023, 1:39 PM)

<https://archive.topquadrant.com/topbraid-composer-install/>

Stevan Earl to Everyone (Dec 13, 2023, 1:45 PM)

I have to run for school pick up. Huge thanks to Steve, and I look forward to hearing about the rest of the session!

Sarah Elmendorf [she/her] to Everyone (Dec 13, 2023, 1:56 PM)

I have to run too thanks for the informative meeting!

John Porter (he/him) to Everyone (Dec 13, 2023, 2:41 PM)

https://www.vcrlter.virginia.edu/data/test_unitsws.html

John Porter (he/him) to Everyone (Dec 13, 2023, 2:57 PM)

<https://archive.topquadrant.com/topbraid-composer-install/>

https://github.com/EDlorg/Units-WG/blob/main/RCode_JP/EML_unit_annotate_functions.r

2023-11-08

Present: Stevan, Margaret, Mary, John, Christine, Sarah

Regrets: Marina

Agenda

1. Review progress on goals from last meeting:
 - a. A set of units for a pull-request: ~20-30 units, with comments (for qudt experts) as needed
 - b. Manuscript to circulate among our group
 - i. From John:
https://docs.google.com/document/d/1KXe5QAADXDQ9b0Y4mPag4_nbY2qEW_SCPbcil2tRSw/edit?usp=share_link

Notes

Problems we are encountering:

With our contributions

1. How best to assign quantityKind (can't do it accurately without understanding the entire measurement.
 - a. Margaret's soln: the most general, e.g., rate, flux
 - b. John: (quantities that matched this
2. We are not assigning dimension vectors consistently.
 - a. How to use "D" - can it refer to only part of the unit?
 - i. Dimensionless in the numerator E.g., a ratio/time: A0E0L0I0M0H0T-1D1, or
 - ii. D0 only if ALL dimensions reduce to nothing.
 - iii. Documentation:
 1. Schema:
 2. <https://github.com/qudt/qudt-public-repo/wiki/DimensionVector-Vocabulary-Submission-Guidelines>

With our process

1. Rare units - what will the cutoff be, e.g.,
 - a. number of uses.
 - i. our current list covers units appearing >20 times. Suggestion that 10 uses is a good cutoff
 - ii. We could use number of datasets also
 1. E.g., >1

These problems are added as issues in our git repository:

[links]

- Consistent dimension vectors <https://github.com/EDlorg/Units-WG/issues/4>
- How to validate <https://github.com/EDlorg/Units-WG/issues/6>
- quantityKind questions <https://github.com/EDlorg/Units-WG/issues/5>
- Order of operations (ie, process for adding new dimension vectors, quantityKinds) <https://github.com/EDlorg/Units-WG/issues/7>

Action items

- All - get your units into the ttl file by Nov 16 (next Thursday)
- ~~Margaret: add the rest of the git issues~~
- ~~Margaret: Friday, Nov 17, submit these to QUDT.~~
fork QUDT, add our units, create a pull request.
In the comment, refer to our list of issues (problems we had during unit creation), and that we need their input to resolve these. Request that we be part of the review discussion

2023-10-11

Present: Stevan, Margaret, John, Sarah, Christine

Regrets: Marina, Mary

Agenda:

Critique/evaluate/cross-check new units in the Git ttl file:

<https://github.com/EDlorg/Units-WG/blob/main/QUDTcontributions2023.ttl>

Discuss comments on existing units. Collected in the “Add Unit Mappings to QUDT” (this was a list of units for the LTER IMs to use in their mapping checks)

https://docs.google.com/spreadsheets/d/1m5oe9fKvG3CmqJzL_FcFTZQjTGm5-w2NjW-nitAjWIk/edit#gid=1411219426

Notes

FYI Margaret is collecting a group of 8 units here (in the temp directory):

<https://docs.google.com/spreadsheets/d/1jLbg3TQ79mIDl9ftyQg7RqmjeT-lkiTgnJF5G-33t30/edit#gid=207956447>

Example datasets that have QUDT units annotated:

Vcr.395.1, vcr.396.1, cap.712.1

Guidelines for our new units

1. Make sure we are all creating symbols in utf8. Determining character set might be difficult if you are editing in a browser.

- a. This might help: <https://string-functions.com/encodeddecode.aspx>
2. QuantityKind
 - a. First create the dimension vector. Then look up that vector in QUDT.ttl, and add all the quantityKinds that apply (to the unit! Not to the vector). Be careful though, because there is not a 1:1 relationship between QK and dimension vector.
Example
 - i. Our candidate is "Milligram per Gram per Hour"@en-us (probably the quantitykind is specific update rate); Its vector is = A0E0L0I0M0H0T-1D0, which is in 24 quantityKinds, e.g., Frequency.
 - ii. The best way to find out the quantitykinds of interest to us is to look thru the datasets that used this unit, and analyze their measurements.
3. ConversionMultiplier.
 - a. To convert TO the new unit (e.g., for unit = KM, conversion = 1000
4. If you only include one spelling in a label, Default to english spelling (metre, not meter)
5. If you think we find mistakes in QUDT
 - a. Make a note here:
https://docs.google.com/spreadsheets/d/1m5oe9fKvG3CmqJzL_FcFTZQjTGm5-w2NjW-nitAjWik/edit#gid=1411219426
 - b. If you copy that element (stevan's example was for gm-per-hectare), add a comment to your ttl entry.
6. Make comments in the ttl file for questions, comments on each other's units
7. System: we seem to be safe with SI, but others apply so add as you think best...

Our questions for QUDT experts:

1. What is the difference between dterms:description and qudt:plainTextDescription
2. Is there a 1:1 relationship between QK and dimension vector? NO! E.g., these two units have the same dimensions: "Milligram per Gram per Hour"@en-us and
3. Any ideas for including vectors for more complex units? E.g., "milligram per gram per hour" actually has mass in both the numerator and denominator, and these are meaningful. e.g., something like A0E0L0I0M[1-1]H0T-1D0
 - a. related : most likely, the entire measurement is something like "milligrams [nitrogen] per gram [dry weight organism] per hour". We've been pretty successful getting people to use only unit parts that can be dimensionalized, which now means those other meaningful parts are not captured.
 - b. See instructions for numerator and denominator.

Goals for November:

2. A set of units for a pull-request: ~20-30 units, with comments (for qudt experts) as needed

3. Manuscript to circulate among our group

2023-09-06

Present: Stevan, Margaret, Marina, John, Sarah

Regrets: Mary

Agenda

~~Start building out the important units here~~ SEE BELOW FOR FINAL DECISION ON PROCESS

- https://docs.google.com/spreadsheets/d/1m5oe9fKvG3CmqJzL_FcFTZQjTGm5-w2NjW-nitAjWlk/edit#gid=0
- (if they have filled out “Proposed_New_Unit_to_QUDT”, col c)

To a spreadsheet resembling:

- <https://docs.google.com/spreadsheets/d/1FxGkYNnCyFYKTjV-BQ3TBI-Elwz8Z3BB9-z2cXCxFe4/edit#gid=0>

QUDT Guidelines:

<https://github.com/qudt/qudt-public-repo/wiki/Unit-Vocabulary-Submission-Guidelines>

John’s code update

- For the adventurous, attached is R code for creating either 1) an edited EML document containing annotations as part of the <attribute> element, or 2) an XML snippet consisting of an <annotations> element containing all the individual <annotation> elements. The latter only works if you already have “id=” attributes as part of your <attribute> tag, because it relies on that id to reference.
- When processing an entire EML document you have the option of incrementing the package revision number (probably won’t work for NEON Christine!) and having id= attributes automatically added to the <attribute> tags if they weren’t there already.
- I’m still working on the code that automatically adds id= attributes to the <attribute> tags and creates an EML document with all the annotations at the end, but I thought you’d enjoy playing around with this.

Notes

QUDT **required** elements:

- rdf:type
- rdfs:label
- qudt:abbreviation
- qudt:conversionMultiplier
- qudt:conversionOffset
- qudt:hasDimensionVector

- qudt:hasQuantityKind
- qudt:plainTextDescription
- rdfs:isDefinedBy qudt:ucumCode

QUDT **optional** elements:

- qudt:code
- qudt:longDescription
- qudt:symbol
- qudt:uneceCommonCode
- qudt:exactMatch
- expression (LatexString)
- prov:wasInfluencedBy

We want to assemble a sample of our units for a first submission to QUDT. planning on 40-50 (10 units each).

Assemble these in a trtl file, in Git.

Use a spreadsheet for housekeeping, probably the same one that the LTER IMs were working in.

Process

Working in this folder:

<https://drive.google.com/drive/folders/1binrTv4y3GEs1ypduIJo-DosPoM-D9cQ>

1. Claim a unit in the log:

<https://docs.google.com/spreadsheets/d/12rVYLgtvBo48mmayZLya46Uja7W66syITajrCoSIXZM/edit#gid=0>

2. Clone the git repo

3. Create a ttl ("turtle") entry in this file

<https://github.com/EDlorg/Units-WG/blob/main/QUDTcontributions2023.ttl>

Recommendation: Find a similar existing unit in QUDT, copy its entry, edit as necessary.

<https://qudt.org/2.1/vocab/unit>

Tools:

QUDT guidelines:

<https://github.com/qudt/qudt-public-repo/wiki/Unit-Vocabulary-Submission-Guidelines>

QUDT units

<https://qudt.org/2.1/vocab/unit>

Validators:

UCUM: <https://ucum.nlm.nih.gov/ucum-lhc/demo.html>

TTL:

TTL syntax:

lines end in ;

entry ends in .

Next Meeting: critique each other's units. Goal is to have 50 or so units done, to start the conversation with QUDT.

2023-08-09

Present: Margaret O'Brien, John Porter, Mary Martin, [Marina Frants](#), Stevan Earl

Regrets: Christine Laney, Chris Turner

Agenda

1. go over the feedback we received at the ESIP/IMC meeting;
 - a. Can we further engage the LTER IMC (e.g., the spreadsheets?).
 - b. Some notes have been collected here:
<https://drive.google.com/drive/folders/1TTw0q-8EJt76MtxHe7ZN1j34E4jC5Xvt>
2. plan next steps
 - a. QUDT contributions
 - b. Report back to EDI (e.g., incorporation into EZeml)

Notes

John has already written more web services (and ids for parent element)

Still to be decided: what style of annotation do we want to use? All our examples have been an annotation as a child of the attribute element. Programmers prefer simple lists instead. We should plan to be able to do all three.

Discussion of nominalDay etc.

Qudt-day maps to EML-nominalDay (number of days elapsed)

Qudt-month

Qudt-year

In EML standard unit list, these three: nominalDay, nominalMonth, nominalYear all refer to elapsed time. Example measurements include things like

“Age”: use nominalYear

“Day-since-start-of-experiment”: use nominalDay

“Length-of-fiscal-year”: use nominalMonth

How to break up QUDT list to get it to work with EZEML - pick list too long

Break by dimensions? Using dimension vector

Core Metabase also will require editing of “table”

They could put it into annotation table?

Unit dictionary and unit type stable

New units for qudt:

Focus on what people have indicated as important at:

https://docs.google.com/spreadsheets/d/1m5oe9fKvG3CmqJzL_FcFTZQjTGm5-w2NjW-nitAjWIk/edit#gid=0

(if they have filled out “Proposed_New_Unit_to_QUDT”, col c)

Construct a spreadsheet of this style:

<https://docs.google.com/spreadsheets/d/1FxGkYNNcyFYKTjV-BQ3TBI-Elwz8Z3BB9-z2cXCxFe4/edit#gid=0>

Action items

1. Start a conversation with QUDT.
 - a. New units (about 250)
 - b. A few issues (microG... should be microGM... is there a way to make these consistent?)
2. John will add “nominal” elements to substitution list
3. John will continue work on draft manuscript
4. Tell IMs the Sept 1 deadline for adding to
https://docs.google.com/spreadsheets/d/1m5oe9fKvG3CmqJzL_FcFTZQjTGm5-w2NjW-nitAjWIk/edit#gid=0
 - a. Deadline so we can work on high priority units
 - b. Still working on mechanism for adding units in the future
5. John will continue work on R functions for adding annotations to EML documents

2023-07-10

Present:

Regrets:

Agenda: Finalize ESIP poster.

<https://docs.google.com/presentation/d/1nXtU0oVitFJc50WtEFbvC2w1utvF1DJAc01d2cEODIQ/edit#slide=id.p3>

zoom link for 7/7 (1100 EDT)

<https://ucsb.zoom.us/j/85166868304?pwd=a0RUUW1hQ3VGM0k3ZER0ejlkc2s2UT09>

2023-07-06

Present: Stevan, Margaret, Marina, Mary

Regrets: John, Christine (conflict with another EDI WG)

Agenda

Ad hoc mtg, prep for IMC/ESIP

Resolve:

1. Poster was designed for the IMC. Perhaps it does not belong up during the ESIP meeting. Thoughts/Resolve.

Summary:

As we suspected, this poster is not quite right for an ESIP audience. We want something that is more general, and describes the work we have already done. This time did not work for John and Christine (apologies; mob did not look at the EDI calendar...)

Solution for collecting feedback from the IMC:

we can use a table in a google doc, or a spreadsheet (organized like the table from the ppt slides).

Solution for the ESIP poster:

feature a summary of our past work, especially John's mapping, etc.

IMPORTANT: John, don't print the poster as is! We should talk first.

Upcoming tasks:

Friday (Margaret, in France, UTC +01) will do some edits. Others are more than welcome to do this as well.

This group will meet Friday AM (US time zones), same bat-channel.

John, I rescheduled the units WG meeting for 1100 EDT on Friday 7/7 -- if we don't hear from you, we will assume this time will work.

zoom link for 7/7 (1100 EDT)

<https://ucsb.zoom.us/j/85166868304?pwd=a0RUUW1hQ3VGM0k3ZER0ejlkc2s2UT09>

poster link (for active editing)

<https://docs.google.com/presentation/d/1nXtU0oVitFJc50WtEFbvC2w1utvF1DJAc01d2cEODIQ/edit#slide=id.p3>

ppt link for IMC slides (context):

https://docs.google.com/presentation/d/1AB1uWsNx_IBeL0C5MBkNhToEUt-Nz0xl5rdpnnocbX8/edit#slide=id.p

Next steps for this WG:

assembling the list of units to be submitted to QUDT. Stevan (rightly) pointed out that we should decide on the criteria for accepting a unit before we start vetting the list. Perhaps we can hold an ad hoc meeting during ESIP for this. If not, that will be the topic of our next regular zoom call.

2023-06-14

Present: John, Margaret, Mary

Regrets: Christine, Stevan, Marina

Agenda

1. Potential workshop at LTER IMC annual meeting (July 17).
 - a. Text from Greg Mauer under notes.
2. Discuss what to do next with John's mapping of found_unit:QUDT_unit
 - a. Link:
<https://docs.google.com/spreadsheets/d/1nht8kkukEMPzUVnAAKV32vJabKMmW3Bm/edit#gid=568393210>
 - b. What does the lookup file look like (see the one in github - possible revisions)
(https://github.com/EDIdorg/Units-WG/blob/main/RCode_JP/DataFiles4R/unitsWithQUDTInfo.csv)
 - c. Web service - outline functionality
 - d. Communication (to managers of existing datasets)
 - e. Pie-in-the-sky – tell EDI that we could 'enhance' a good fraction of their datasets.
What ideas do they have in that area?
3. Discuss results of possible QUDT spreadsheet action item, determine next steps
 - a. SPREADSHEET LINK:
<https://docs.google.com/spreadsheets/d/1FxGkYNnCyFYKTjV-BQ3TBI-Elwz8Z3BB9-z2cXCxFe4/edit#gid=0>

Notes

IMC

Email from Greg Mauer:

Dear Units working group, I have something for you all to consider in your next working group meeting. The IMC Annual Meeting is coming up soon (July 17th) and we are looking for

constructive ideas for workshops. Could your working group benefit by running some kind of workshop at the annual meeting?

We imagine workshops will last around an hour, and it would be your choice on how to run them. We are hoping that they will be constructive and help move working group priorities forward, so the use cases you might plan a workshop around could be:

Does your working group need new participants to help with specific tasks or features?
Generate buy in and try to recruit new members with a demo and a request for skilled help (programmers, writers, etc)

Do you want to scale-up your group's work across LTER? Expand your WG activities by training IMs and getting them started on a task.

Do you need input? Gather IMC community feedback to help your working group make decisions and create a polished final product.

If any of these sound useful to your working group, we encourage you to consider a workshop at the annual meeting. IMexec is happy to help you plan and execute your workshop too. Please reach out to me, or anyone else on IMexec, if you are interested in this idea.

Possible Ways to use WG code QUDT Units

input	How	output	issues
EML Doc	R (DIY, local)	EML doc with units annotated and attribute IDs added if not already there	Requires local configuration
Directory of EML Docs	R (DIY, local)	New directory of EML docs with units annotated and attribute IDs added if not already there	Requires local configuration
QUDT ID	R (DIY, local)	Annotation snippet	Requires local configuration
EML doc	WS	EML doc with units annotated and attribute IDs added (if not already there)	Option: update revision number Option: annotation inside <attribute> or as an <attributes> list
Single unit string	WS	QUDT URI	
Single unit string	WS	Annotation snippet	

List of units	WS	List of QUDT URIs	
List of units	WS	Annotation snippets (forest)	
attribute ids and associated units as a csv	WS	<annotation> element containing all the units that match with QUDT	Won't work without attribute ids
QUDT ID	WS	<annotation> snippet	Possible option: output an stmml/unitList

Draft Workshop Agenda

1. Introduction <15 minutes
 - a. Progress
 - i. Mapped 6xxx units to QUDT (numbers, next steps)
 - ii. How to use QUDT in EML docs (annotation demo)
 - iii. Web service design
2. Web service demo
3. Discussion
 - a. General comments
 - b. Print poster of table above - and provide DOTS to put on it to "vote"
 - i. Include extra lines for other ideas
 - ii. Include explanation at top
 - c. Ideas for ezEML, EML assembly line

IMC session notes

4. Have people review the units used at their site and matches to QUDT
 - a. Probably not at this meeting - probably better done online, or on your own
 - b. Unless the web service is ready for preview, then they could test out the web service with their own units.
5. Feedback we need:
 - a. What will make it easy for you to incorporate QUDT into your site EML
 - i. You want to create EML the way you make now, and it gets Modified to include annotation EML
 1. We could give you R Code to ingest EML, produce modified version that includes annotations (DIY)

2. Set up a web service where you can upload an EML document and receive back a modified version with annotations added
- ii. You want to create your own annotations
 1. Help EZEMML and EML Assemblyline incorporate tools for adding annotations correctly
 2. Web service - input a raw unit or list of raw units and get back either the matching QUDT URI or an annotation
 3. Web service - input a raw unit and get back an annotation
 4. Web service that takes in a LIST of raw units and IDs and produces an <annotations> element with all the annotations included in it.

Action items:

~~Margaret will check with Jones regarding the syntax for annotation~~

-> It was there the whole time: <http://qudt.org/schema/qudt/hasUnit>

Margaret - poster of table, for vote-by-sticker

Margaret - ppt for introduction

Draft:

https://docs.google.com/presentation/d/1AB1uWsNx_IBeL0C5MBkNhToEUt-Nz0xI5rdpnnocbX8/edit#slide=id.g20903dde4ab_4_0

John will work on web service for returning annotations based on input of raw unit

Single outputs only (unit str > qudt URI, unit str > annotation)

2023-05-10

Present: Margaret, John, Marina, Stevan

Regrets: Mary

Agenda

Review QC results

Next steps

Notes

We are in awe of Mary's ability to track things down - found unit annotations that were good, but that the unit was inappropriate to the underlying data. Not something we need to address - but impressive!

Example copied/condensed from her email:

<https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-luq.119.1545979>

attName: Tag

attDescr: Tag # that is attached to tree

Unit: metersPerSecondSquared

Type: real

<https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-mcm.301.4>

attName: PON (microg L-1)

attName: Particulate organic nitrogen concentration found at lake depth

Unit: metersPerSecondSquared

Type: real

<https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-luq.190.17>

attName: MAIN STEM

attName: The main stem integer tag number of an individual tree.

Unit: metersPerSecondSquared

Type: real

“Unknown” should not be UNITLESS

NA is too ambiguous. Is N/A the same?

perMill - PPTH is definition in QUDT but may be issue with isotopic scientists - Margaret will check - may need to ask QUDT to add synonym

Uninterpretable (but possibly the submitter intended to fill this in later:

Unknown

NA

N/A

Dd

Ddd

No units

Margaret will post “possible additions spreadsheet” on Google Drive and parcel out the task of checking whether equivalent units in QUDT already exist.

Action Items

1. Look in QUDT to see if any fractional units are included, e.g.
 - a. .01 meter squared
 - b. Goal is to know whether we should propose any of these
2. John will work on fixing known problems (N “no” in spreadsheet <https://docs.google.com/spreadsheets/d/1nht8kkukEMPzUVnAAKV32vJabKMmW3Bm/edit#gid=568393210>) [unitsWithQUDTCheck.xlsx](#)
3. QUDT new additions
 - a. spreadsheet of UniquePossibleAdditonsToQUDT: - post in google drive
 - i. **Everyone**: Confirm that there is not already a unit in QUDT with an alternate spelling (ie, that does not match the spelling that we came up with)
 1. Assign spreadsheet parts to WG members
 2. One place to start: https://ble-lter.github.io/QUDT_Browser/
 3. Instructions: LINK HERE
4. Address questionables
 - a. When does per-mil = parts per thousand?
 - i. Isotopic fractions
 1. Mob checked this with her spouse, who works in this field. He says that permil is equivalent to ppt, and many people use them interchangeably. (also pointed out that some others may disagree, but that is unavoidable).
 - ii. Seawater salt concentrations
 - iii. [Some other context]

Future ToDos

1. Start BP items to address
 - a. Unit level
 - i. How to create a customUnit that cannot be mapped to QUDT
 - ii. When to use dimensionless (or unitless)
 - b. Attribute level (save for a follow up working group)
 - i. How to choose a unit correctly (?)
 - ii. Examples of what not to do (anonymized)
 1. the datasets that Mary found (above)
 2. Edi.417.1, see additionalMetadata/unit for “micromolar”
2. Start draft of 2 newsletter posts: EDI and LTER databits

- a. Cover basics of how QUDT will work with EML metadata
 - b. Cover basics of how “context matters”; context = entire attribute (context out of scope for this WG)
 - c. Discuss/List uninterpretable units (context = the WG could not map these at all, and most likely no one would be able to reuse this data without follow up or investigation)
- 3. QUDT - new additions
 - a. Contact Kai Blumberg

2023-04-12

Present: John Porter, Mary Martin, Margaret O'Brien Marina Frants, Christine Laney
Regrets: Stevan

Agenda

Review automated unit assignments, determine next steps.

Action items

Margaret: ~~deposit John's spreadsheet. Done;~~ started a folder called “Products”
~~Write up the QC protocol and distribute~~

Units QC

Everyone: Quality control on John's spreadsheet called “unitsWithQUDTCheck”

Location:

<https://docs.google.com/spreadsheets/d/1nht8kkukEMPzUVnAAKV32vJabKMmW3Bm/edit#gid=568393210>

Why: the spreadsheet contains assigned QUDT units for every unit found in an EDI/NEON/DataONE dataset (column C). This was semi-automated, so the list should be checked to ensure these are correct.

We are confirming that the automated mapping to the QUDT unit is correct. Column C is the original unit (found in the EML), and there will be misspellings, white space, etc. But if you are pretty sure they all meant - e.g. - "millimeter", then they should all map to QUDT “MilliM”
The spreadsheet contains the QUDT link (column F), in case you need to check more of the entry that is present in these columns.

Protocol:

The 5 LTER-related people are each assigned ~300 units to confirm correct:

John, Margaret, Marina, Mary, Stevan

The spreadsheet is sorted by unit, alphabetically. That means unit variants are grouped together (e.g., "acre", "acres"; "centimeter", "centimeters") which makes each person's assignment much shorter than 300, because you can probably vet a whole group together.

Column A: name of the person assigned that unit

Column B: mark each unit, as to whether it is OK:

allowable answers:

Y - assignment is correct

N - assignment is incorrect

Q - assignment is questionable, or more information is needed

Column J (last column) - optional, if you need to add a note

Notes:

The spreadsheet is generated by a script, so don't remove any lines! Keep in mind that column C (the unit as it appeared in EML metadata) may have been misspelled - that's fine. Our goal will be a list we can give back to each contributor (LTER site) with the match between their unit and QUDT.

We will discuss the results at the May meeting. We may need a way to go back to the original dataset.

Notes

Progress since last meeting:

John - tons. (below is the summary sent to imexec rep, Chris)

The Units working group has been making substantial progress. First, as discussed at the EDI session (webinar?), we have a reasonable draft list of units which we can link to specific QUDT URIs. This includes approximately 82% of units that can be resolved (some like "dimensionless" are not resolvable) making up 1,125 unique raw unit descriptions, and we have now added all of the 1,753 UCUM codes recognized by QUDT, regardless of whether they match anything currently in our corpus of EML documents. You can see the list at: <https://www.vcltr.virginia.edu/data/unitsWithQUDTInfo.csv>

Second, we've got draft R code where you can give it the ID info for any dataset in EDI and it will return a copy of the EML for that dataset with unit annotations automatically added. Optionally, if the <attribute> elements in EML contains a IDs (e.g., <attribute id="Ent1.Table1.Attr1">) it will create a file containing an <annotations> element that could be inserted into an otherwise unmodified EML document. I've attached copies of sample output documents from that code. If anyone would like to try it out, let me know and we can share the draft code, as long as it is understood that it is still DRAFT code!

The things we are still working on are 1) improving the unitsWithQUDTInfo.csv file to include more units, perhaps dealing with some of the special cases such as "stemsPerMeterSquared" which could be resolved as "NUM-PER-M2", and 2) refining the draft R code for automatic annotation of documents. There are some issues there we are still working on (the valueURI is pretty clear, but we are still discussing the propertyURI and which ontology to use in defining it). We are also trying to figure out how best to make this capability available to information managers that generate their EML in different ways.

Margaret

1. Dimension vectors, draft how to:
https://docs.google.com/document/d/1P8QLI24P97rmAgT_zc_6BFSrUXYbFHxUJu19IJ9jdTw/edit
2. Working on understanding quantity, quantityKind, wondering if that we can use those to subset a long list

2023-03-08

Present: Margaret O'Brien, Stevan Earl, Mary Martin, John Porter

Regrets: Christine Laney

Agenda:

1. Follow up on webinar from feb 21
 - a. Breakout notes:
https://docs.google.com/document/d/1FDKWkJCNLXX9xiQBxUav67HBY6pUMUH_oXbr4xXrfWc/edit
 - b. Does this feedback bring up any specifics we should address or add?
2. Quick review of where we are on the below ("products"), what are the next steps.
 - a. Mapping between raw unit, qudt unit (John)
 - i. What can others in the WG do?
 - ii. Recommended format?
 - b. QUDT submission started (margaret)
 - i. Identifying more units to add

- ii. Required fields
- iii. Domain expertise

Notes

Next steps:

John wants to play with creating annotations from the raw-qudt mapping table. We will decide later where the annotation will go (in the EML, either at attribute, or annotation section). He will use the example from the 2/21 slides.

Label = rdfs:label

URI = qudt uri

For the property, use the example for now (semanticscience.org), but we may want to change this.

He also wants to generate an entire EML doc with annotations added for units.

Margaret wants to tackle some of the unmapped units. There are many, but a lot are garbage, and sorting the chaff is tedious. John and Stevan think it would be better to generate annotations for what we can now, and then let the legitimate units bubble up. Our job would be to tell the community what was required to submit a new unit to QUDT and mediate that process.

John will work on developing the tables linking raw units with QUDT URIs, Labels and UCUM codes (from QUDT) and some draft interfaces for testing them. Issues: annotations inside <attribute> tags or in <annotations> or <additionalMetadata> tags at the end of the document need to be included. Ideally have both QUDT and UCUM annotations.

Possible interfaces 1) put in raw unit, get back annotations for inside <attribute> tag, 2) give URL to EML document, service will extract raw units, match them up and return <annotation> tags suitable selected location (inside <attribute> or end of document). 3) could be done via web service, or R that accesses needed tables over the web.

Better: margaret should concentrate on descriptions of units for QUDT additions.

<https://github.com/qudt/qudt-public-repo/wiki/Unit-Vocabulary-Submission-Guidelines>

Our spreadsheet, collecting our submissions:

https://docs.google.com/spreadsheets/d/1zEhVQXMtxEWoKmaDEGYryaiqmgyCUHyxHIQ_SE3tsjU/edit#gid=0

Mary will add some units

Margaret will spend her time on dimension vectors. Provide a summary for next time.

Example dimension vectors:

<https://qudt.org/2.1/vocab/dimensionvector>

How to:

<https://github.com/qudt/qudt-public-repo/wiki/User-Guide-for-QUDT#3-introducing-dimension-vectors-for-si-units>

<https://qudt.org/schema/qudt/QuantityKindDimensionVector#:~:text=A%20Quantity%20Kind%20Dimension%20Vector.of%20a%20system%20of%20units.>

Still to do:

1. The BP stuff. E.g.,
 - a. at the measurement outlining how to describe a “day” vs a date (like day since start of year).
 - b. At the unit level: how to submit a new unit to QUDT
2. Mary: for the standardUnits, applications like EAL could be adding the QUDT unit right now! We should agree on where in the EML we want the annotation to go.
3. people will be making EML all different ways. There should be some coordination.

2023-02-15

Webinar prep - Margaret, John

Slides (EDI unitsWG folder):

https://docs.google.com/presentation/d/1xZxVzeDDiWW4n56kjiI4VfBNb0PIKY38IHpZvpUc9aM/edit#slide=id.g20903dde4ab_2_23

Questions to anticipate,

1. How do I map my units to QUDT?
 - a. We're doing that for you! (see slide 17, John's excel)
 - b. Or try it yourself, manually, here: https://ble-lter.github.io/QUDT_Browser/
2. What goes into my EML unit field?
 - a. Something from that list (john's spreadsheet)
 - b. Something readable
 - c. Minimum: some other string, as long as its mapped to QUDT
3. What does an annotation look like?
 - a. See slide XX (example)
 - b. For the annotation, all you really need is the spreadsheet, and a bit of code to shove it into your EML. (you can continue to manage units and create custom units the way you are doing it now)
4. What if my unit is not in QUDT?
 - a. We missed it. Lets talk (maybe it should be added to qudt)
5. Show me the QUDT entry for this unit ____, right now ____
 - a. nominalDay

- b. Fraction
 - c. Microeinsteins...
- 6. My PIs insist on including things like “carbon” in the units. What should I do?
 - a. Even if your customUnit is e.g., “mg Carbon per liter” you can still map it to the appropriate QUDT unit in an annotation, which is mg-per-liter. Ref intro slide on unit vs measurement
 - b. Keep in mind that conversions that include things like atomic weight are more complex.
- 7. How do we get new units into QUDT
 - a. They have a process for contributions, we will have many to add. This will happen soon - we need the new units to be there for software tools to work the way you need them to.
- 8. How will future communication with QUDT work?
 - a. Probably through an ongoing EDI working group.
- 9. How will the tool to create EML content (unit name, stmml for custom unit) from QUDT work?
 - a. We don’t know yet. We need input from people who are interested in that topic, to design it.
- 10. Will a tool distinguish between standard units and custom units? How?
 - a. Don’t know yet, but it could be done. The most important thing is that the annotation is there, that points to QUDT. that's the part of the measurement that can be used for data interpretation, integration and conversions.

2023-02-08

Present:

S. Earl

Marina Frants

C. Laney

M. Martin

M. O'Brien

J. Porter

Regrets:

Agenda:

1. Upcoming EDI webinar - 21 Feb 2pm ES
 - a. Presentation from LTER ASM:
https://docs.google.com/presentation/d/1wfWPqympdGmIAMzrKTPVTfdhiiCdAUFC23xMkgawvnc/edit#slide=id.g146de5db504_7_0
 - b. New slides needed:
 - i. To do
 - ii. Why? (do this at all)
 1. Make it so units have MEANING (eg, via resolvable uris)
 - a. Help facilitate unit conversions
 - b. multiple spellings can be tied to the same concept (URI)
 - i. Reconcile variant representations of unit names
 2. Enable searches based on (dimensions) Helping consistence searches based on type of data
 3. Easier evaluation for use in synthesis
 - a. List of datasets sharing sets of units
 - iii.
 - iv. Mappings completed so far
 1. Automated:
 - a. by john, mainly sequential search-replace of string parts
 - b. Units from all datasets in EDI, NEON, D1 (total=325k uses iof units)
 - c. (~7k entries. Currently able to map ~400) but these total 67% of unit uses
 - d. Still to do (using manual mapping):
 - i. Determine how many unmapped should be pursued - we expect another 100 units might be found
 - ii.
 - iii. Validate automated mappings
 2. Manual - several LTER sites, NEON,
 3. Why do the mapping
 - a. A DM interested in mapping to qudt will need the mapping. This will help
 4. Still to do: Confirm they are correct
 5. New slide: summary, TBD #, # able to map, # mapped
 - 6.
 - v. What units are “missing”?
 1. Mainly a manual process, somewhat subjective
 2. Can propose new units to QUDT
- c. Potential discussions to anticipate
 - i. Including QUDT units in EML
 1. Annotations

- 2. customUnit construction
 - a. API/Tool for automating
 - b. Create table of stmml for QUDT units for easy lookup
 - ii. Possible need for programmatic access to QUDT (e.g., during dataset construction with R)
 - 1. Easy to pull in JSON for QUDT release
 - d. Next Steps
 - i. WG
 - 1. Additionally refining mappings
 - 2. Propose additional units to QUDT
 - a. Needs to be refined by subgroup
 - 3. Design output from QUDT to create STMML
 - a. Focus on Integrate with
 - i. EZeml
 - ii. EAL
 - b. Create resources that can be used by sites
 - i. (e.g., tables of equivalents)
 - ii. Your next steps (as DMs)
 - 1. We will be contributing new units to QUDT. we will need your expertise to finalize that list
- 2. Site units
 - a. Best format
 - b. Next steps
 - i. John's work:
 - 1. Look for a file called UnitswQUDT.csv add it to the Google drive.

AutoSave Off UnitswQUDT.csv Search

File Home Insert Page Layout Formulas Data Review View Automate Help PDF Architect 7 Creator

Clipboard Font Alignment Number Styles Cells

B1 unit

	A	B	C	D	E	F	G	H	I
1	qudtUnit	unit	qudtUri	pseudounit	orgList	numOrgs	TotalUses		
2	dimensionless	dimensionless	NA	dimensionless	EDI NEON	3	72533		
3	NUM	number	http://qudt.org/vocab/unit/NUM	number	EDI NEON	3	32025		
4	PERCENT	percent	http://qudt.org/vocab/unit/PERCENT	percent	EDI NEON	3	26544		
5	DEG_C	celsius	http://qudt.org/vocab/unit/DEG_C	celsius	EDI NEON	3	21318		
6	M	meter	http://qudt.org/vocab/unit/M	meter	EDI NEON	3	11666		
7	MilliM	millimeter	http://qudt.org/vocab/unit/MilliM	millimeter	EDI NEON	3	9539		
8	DEG	degree	http://qudt.org/vocab/unit/DEG	degree	EDI NEON	3	8259		
9	CentiM	centimeter	http://qudt.org/vocab/unit/CentiM	centimeter	EDI NEON	3	6580		
10	none	none	NA	none	EDI NEON	3	5207		
11	GM	gram	http://qudt.org/vocab/unit/GM	gram	EDI NEON	3	4399		
12	V	volt	http://qudt.org/vocab/unit/V	volt	EDI NEON	3	1928		
13	HR	hour	http://qudt.org/vocab/unit/HR	hour	EDI NEON	3	1027		
14	KiloPA	kilopascal	http://qudt.org/vocab/unit/KiloPA	kilopascal	EDI NEON	3	978		
15	SEC	second	http://qudt.org/vocab/unit/SEC	second	EDI NEON	3	762		
16	MIN	minute	http://qudt.org/vocab/unit/MIN	minute	EDI NEON	3	653		
17	DAY	day	http://qudt.org/vocab/unit/DAY	day	EDI NEON	3	591		
18	MilliL	milliliter	http://qudt.org/vocab/unit/MilliL	milliliter	EDI NEON	3	405		
19	L	liter	http://qudt.org/vocab/unit/L	liter	EDI NEON	3	94		
20	NanoM	nanometer	http://qudt.org/vocab/unit/NanoM	nanometer	EDI NEON	3	82		
21	proportion	proportion	NA	proportion	EDI NEON	3	73		
22	HZ	hertz	http://qudt.org/vocab/unit/HZ	hertz	EDI NEON	3	10		
23	M-PER-SEC	meters per second	http://qudt.org/vocab/unit/M-PER-SEC	meterpersecond	EDI DataO	2	26423		
24	HA	hectare	http://qudt.org/vocab/unit/HA	hectare	EDI DataO	2	1900		
25	DeciBAR	decibar	http://qudt.org/vocab/unit/DeciBAR	decibar	EDI DataO	2	1745		

3. Accumulating units for a QUDT submission

- https://docs.google.com/spreadsheets/d/1zEhVQXMtxEWoKmaDEGYryaigmqyCUHyxHIQ_SE3tsjU/edit#gid=0
- Required fields still missing for most: offset, multiplier, dimension vector, quantities

4.

Notes:

Margaret finished the SBC list of units. And mcr.

Add to slides

Scope of units we are looking at - d1, neon, edi, lter sites

2023-01-11

Present: Margaret, Stevan, Corinna, Christine

Regrets: John,

Agenda

1. Reports from action items
2. John's spreadsheet
3. API (Stevan?)
4. Progress on our own units
5. Next steps ? UCUM?

Notes

Margaret is using https://ble-lter.github.io/QUDT_Browser/ to find sbc units

Also can use: https://www.qudt.org/doc/DOC_VOCAB-UNITS.html

Or this file, using browser-find:

<https://docs.google.com/spreadsheets/d/1vf4T1i69XrjuoKrLwmWVOZJr1m6kAREJSakAaruXuDM>

Action items:

Stevan, Margaret, Christine: Create scope-based tables like Corinna's for CAP, SBC, NEON.

Put the table here:

<https://drive.google.com/drive/folders/1hl6o1o6lmoC8WjMqhqqazPuKO2wyfPPv>

Output should be in the same format as Corinna

2022-12-14

Present: Margaret O'Brien, Christine Laney, Stevan Earl, Chris Turner, Marina Frants, John Porter,

Mary Martin, Corinna Gries

regrets:

Agenda

Reports from action items

Git repo

Use existing or new... Probably too confusing to put with old one. Yes on new GIT repository - but are we READY for it?

Discussion of some past activities using old units dictionary database in creating EML 2.2 list. Access to old units database limited by old PHP code. Database is stored in a database. Marina able to use information sent to modify database.

View units in EML assemblyline uses new list? YES both EML assemblyline and EZeml it doesn't use deprecated. You need to FILTER to get the non-deprecated list, but there is a column that indicates status of unit. In EML 2.1 can't use new spellings - need to upgrade to 2.2.

Steps -

1. create psueudo-unit column with "condensed" units - lower case, no spaces, slashes/per to same, some basic units we may be able to deal with plurals etc.
2. Sort on pseudo-units
3. Manually add QUDT and/or UCUM units for a selected number of units
 - a. Share on Google Sheets - put your initials on ones you add
 - b. Would be good to have list of QUDT units
 - i. Margaret has exports from QUDT
<https://docs.google.com/spreadsheets/d/1vf4T1i69XrjuoKrLwmWVOZJr1m6kAREJSakAaruXuDM/edit#gid=1072861445>
 1. Column C probably the closest
4. Try to figure out how to automate
 - a. Possibly convert pseudo unit to UCUM code that can link to QUDT
 - b. Need to investigate API to QUDT via R or Python
 - i. Margaret knows someone to talk to regarding the API (Kai Bloomberg)
 - ii. Need to look up the API document

John will do 1-2 and some of 3

Margaret to do #4

NTL plans to use QUDT codes as an annotation. QUDT unit and QUDT attribute name in ECSO (<https://bioportal.bioontology.org/ontologies/ECSO>) (Dataone/NCEAS ontology) - put units twice - once "raw" once as QUDT. Also attribute name as "raw" and link to ECSO <https://bioportal.bioontology.org/ontologies/ECSO>

Relationship between QUDT and STMML? Not worried about right now.

Notes

John reported on our progress to the IMC. Text (via email):

We've mostly been in "harvesting" mode lately – developing lists of units used at EDI, NEON and DataONE to see if we can crosswalk them with some of the various Units resources (e.g., QUDT). The assembled list (allowing case to matter) has roughly 8,500 terms but only 426 of

them are used by more than one organization. Most of the unique units come from DataOne (5,388) with EDI accounting for 2,651 units not used by other organizations.

Incidentally, one very interesting additional stat. Although only 426 (5%) units are shared by two or more organizations (i.e., EDI, NEON, DataONE), they account for 73% of all the units used in data packages. So those 426 sound like a great place to focus our efforts!

21 units were used by all three orgs

2022-11-09

Attending: John Porter, Margaret O'Brien, Stevan Earl, Mary Martin, Marina Frantz

Regrets:

Agenda

1. Review progress on the table of units from datasets in EDI and NEON (October action item).
2. Start a git repository

Notes

- 1) Move meeting to 1 hour earlier (agreed)
- 2) Take a look at the Google Drive repository
<https://drive.google.com/drive/folders/0B0Gg0cAZlQbieHdnTEhkNIYzRVk?resourcekey=0-KacyEVXdOs9CcVhyEXBYUQ&usp=sharing> to see if we should use it, or do something with GitHub
 - a) Need to make sure everyone has access
- 3) Old unit-registry in GIT, but that is out of date so created a new one **units-WG DONE**
 - a) Need to add location of new repository to unit-registry readme as the new place.
- 4) Contents of units-WG (new git repository)
 - a) Directory of existing units - as summary files, ie, same as
<https://drive.google.com/drive/folders/1DzMYp21XB3cOH8m1KK9XGFxGhIjhmSQa>
 - i) LTER
 - ii) NEON
 - iii) DataONE (still need to pull down from Matt Jones email below)
 - iv) ArcticDataCenter (ditto Matt Jones)
- 5) What about next step?
 - a) Understand the tables and how we should analyze them

- b) Ways of sorting so that similar units show up near each other in list
- c) Note squaredMeters and MetersSquared won't sort near each other
- d) Are there advanced lexical analyses?
 - i) Name normalizer for EDI name spellings used to regularize names
 - ii) Ask Jon Ide about tools/approach he used - Margaret will do
 - iii) Want to CLUSTER the existing units
- e) Some may sort together
- f) Some may sort together in terms of frequency of use
- g) Plan
 - i) Integrate existing sheets - John will work on
 - ii) Include summary columns for each system + overall totals
 - iii) EDI and DataOne duplicate - so do without DataOne?
 - iv) Hold off on analysis of DataONE for now, so focus on EDI and NEON
 - (1) Decode JSON files from DataONE into summarized CSV form (already done by Matt for ALL DataONE)
 - v) Do file eliminating case differences?
 - (1) Can split strings on camelcase - may help for lexical analysis
 - (2) Can do it both ways (with and without case)
 - vi) Look at integrated summaries to identify cut offs
- 6) Final products????
 - a) See what matches QUDT
 - b) See if definition matches what we mean
- 7) What about next, next step?
 - a) Eventually Need to regularize units
 - i) Perhaps link to number rather than QUDT
 - b) Crosswalk/map units with QUDT?
 - c) Don't want something that sits between QUDT and EDI - but how to deal with synonyms?
 - d)

The below msg was Sent by Matt Jones, 2022-10-25. The summary of D1 units has been added to our drive.

Margaret helped us with a similar type of attribute analysis over the last few years as we've worked on developing attribute vocabularies for annotation. It's pretty easy to get similar lists of units out of the DataONE network if you'd like to see where patterns in units usage are consistent (or not) across repositories. Here are a couple of SOLR queries that produce results that might be of interest to your analysis. If you wanted to use this, I could put a few minutes into distilling the JSON outputs below into CSV formats sorted and grouped by count as John has done, but I thought I would just throw out there that these index queries are pretty easy to do ad

hoc in DataONE, and see if there is any interest in this broader dataset before I put the time into that.

- All units for all metadata formats in DataONE (use the start/rows pager parameters to page through the results):
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+attributeUnit:*&fl=attributeUnit&wt=json&rows=1000&start=0
- All units for the EML family of metadata documents in DataONE
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+formatId:*eml*+AND+attributeUnit:*&fl=attributeUnit&wt=json&rows=1000&start=0
- All units for all metadata formats EXCEPT the EML family in DataONE
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+-formatId:*eml*+AND+attributeUnit:*&fl=attributeUnit&wt=json&rows=1000&start=0
- All units for just the metadata from EDI, Arctic Data Center, and LTER repositories
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+dataSource:*EDI+AND+attributeUnit:*&fl=attributeUnit,datasource&wt=json&rows=1000&start=0
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+dataSource:*ARCTIC+AND+attributeUnit:*&fl=attributeUnit,datasource&wt=json&rows=1000&start=0
 - https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+dataSource:*LTER+AND+attributeUnit:*&fl=attributeUnit,datasource&wt=json&rows=1000&start=0
 - You can get a list of repository identifiers to filter against here:
<https://cn.dataone.org/cn/v2/node>

Matt

I decided to do it anyways, the DataONE-wide unit list is attached, which I generated using this bash script:

curl

```
https://cn.dataone.org/cn/v2/query/solr/?q=formatType:METADATA+AND+attributeUnit:\*&fl=attributeUnit,datasource&wt=json&rows=1000000&start=0 | jq .response.docs[].attributeUnit[] |  
sort | R --no-save -e 'library(dplyr); df <- read.table(file(description="stdin"), header=FALSE,  
col.names=c("unit")) %>% group_by(unit) %>% summarise(n = n()) %>% arrange(-n);  
write.csv(df, "dataone-units.csv")'
```

You could add some other grouping variables like repository or metadata standard if so desired.

2022-10-12

Attending: Margaret O'Brien, Marina Frants, Chris Turner, Christine Laney, Mary Martin, John Porter

Regrets: Stevan Earl

Agenda

1. Assemble feedback from ASM.

Action Items

1. Compare EML unit content to existing dictionaries.
 - a. Questions to answer:
 - i. How consistent are the QUDT units for these? (definitions, speciation of speciation, other?)?
 - ii. What additional units would have to be defined (by us)
 - iii. Are patterns for QUDT id construction consistent and readable enough
 1. to be presented in HTML versions of metadata (for scientists)
 2. To find a unit to use (for people creating EML).
 - a. For NEON, a small group, for EDI a large untrained group
 - b. GOAL: Construct UCUM unit for everything in EDI and NEON
 - i. EML standardUnits
 - ii. customUnits in EML datasets
 - iii. Count of how often each are used (per scope - ltersite, neon, edi)
 - c. Process:
 - i. Acquire EML
 1. EDI: request corpus of current EML from Mark Servilla (Margaret)
 2. NEON: (Christine Laney) One EML doc per data product, in a zip file.
 - ii. Construct a table with columns (John can create cols 1, 2. Possibly manual after that - TBD)
 1. original name (stdUnit name or customUnit name)
 2. count
 3. UCUM
 4. Source
 5. QUDT id
 6. Other QUDT fields (guessing 10 fields?)

Notes

Today's discussion Summary (see further notes below):

1. There are certain advantages to using UCUM that we don't want to give up - ie, the ability to access dictionaries other than QUDT.
2. If we were to go straight to QUDT (per Tim's suggestion below) we would need to ensure that the QUDT id will be something that can be used easily by both scientists and dms, mainly its consistency and readability.
3. Cursory examination using Tim's browser turned up mistakes and inconsistencies in qudt. There could be others, and our conversation with them will include these. We will probably be expected to contribute both new units and corrections to existing ones.

Tim Whiteaker's QUDT Browser:

https://ble-lter.github.io/QUDT_Browser/

Be patient. It takes a minute to load with the slow wifi.

Text from Tim:

Use QUDT for the identifier

- QUDT and UCUM are pretty equivalent in terms of readability for a layperson in my opinion.
- QUDT is the thing we're eventually mapping to, so using its ID makes the mapping clearer than using UCUM.
- Since we're mapping to QUDT, using its ID is one way of helping the community get comfortable with its nomenclature.
- Speciation, which UCUM would allow but not do anything with, belongs in its own element.

For the conversation with QUDT (some of these questions may be answered on their website but I haven't looked)

- What's the long term plan for QUDT? How is this system funded? Purely by donations and volunteer time?
- Will QUDT always be available for public consumption, or is it going to evolve into a closed system like ISO where you have to pay to read the specification?
- How frequent are updates? It looks like the last update was in August 2.
- How long does it take to be discussed and approved/rejected.
- Is there a pattern for deprecating units, for example when a mistake was made? For example, if a unit was GM-PER-CentM2 and a future version correct that to GM-PER-CentiM2, would GM-PER-CentM2 resolve to something that says, "Hey here's this unit but actually its deprecated and you should use this one instead"?
- Are QUDT IDs persistent? I guess that's related to the above bullet.

Include just the QUDT ID in EML, or all the other QUDT stuff, or something in between?

- For readability without having to view an external resource, I think you include the ID, the human-friendly name, and optionally the unit type (e.g., MassPerArea, or maybe mass per area), abbreviation, and description. Just me griping: the [descriptions at QUDT](#) are awkward and may have inconsistent usage of commas and decimals, and appear to use British English spellings. Does EML allow that? I can't pull up anything with this WiFi.

For backwards compatibility, I am fine with the plan to call QUDT units custom units. So, IMs can use EML standard units, QUDT custom units, or their own custom units. If the next is a breaking change, then you could consider dropping stmml and just having <unit> and making it annotatable, like data table attributes.

Technical details of Services

- [units.edi.org/emlforunit/\[QUDT_ID\]](https://units.edi.org/emlforunit/[QUDT_ID])
- E.g., units.edi.org/emlforunit/GM-PER-CentiM2
- Returns EML snippet

EDI role

- Host service
- Host GitHub repo for data package best practices
- Approve pull request to updating best practices for units
- Host webinar on units

Best ways to communicate good practice with unit descriptions

- GitHub repo for data package best practices
- Host webinar on units

Outline input for our QUDT/UCUM decision

- Just do whatever Tim says, starting with giving him some Key Lime Pie

Working Group Notes

GM_Carbon-PE-R-M2-DAY	grams Carbon per square meter per day	g. m- 2. d- 1{ C}	A metric unit of volume over time indicating the amount generated across one square meter over a day. Used to express productivity of an ecosystem.	Mass PerAr eaTi me
GM_Nitrogen-PE-R-M2-DAY	grams Nitrogen per square meter per day	g. m- 2. d- 1{	A metric unit of volume over time indicating the amount of Nitrogen generated across one square meter over a day. Used to express productivity of an ecosystem.	Mass PerAr eaTi me

N}

Problems in Descriptions:

- “Volume” in descriptions
- No “Carbon” in [GM_Carbon-PER-M2-DAY](#) description
- Nitrogen is not generally used to “ express productivity of an ecosystem.”

4.

2022-09-14

Attending: Margaret, Corinna, John, Stevan, Marina, Chris

Regrets: Mary

Agenda:

final planning for ASM session

Suggestion: change time to 1 hr earlier

Notes:

Per John: The updated (6/2020) list of UCUM codes is on Github:

<https://github.com/ucum-org/ucum/tree/main/common-units> as text and spreadsheet.

The corresponding document to the one from 2016 is:

<https://github.com/ucum-org/ucum/blob/main/common-units/TableOfExampleUcumCodesForElectronicMessagingwithPreface.docx>

It is still not clear to me where the g/m2 is translated into g.m-2 but this is the most recent info from UCUM.

Added during intervening weeks:

CODATA working group: <https://codata.org/initiatives/task-groups/drum/>

Their inventory:

https://codata.org/wp-content/uploads/2022/08/DRUM_Units_Inventory_080122.pdf

Paper: <https://www.nature.com/articles/d41586-022-01233-w>

Stop squandering data: make units of measurement machine-readable

In the age of big data, it is time to ensure that units are routinely documented for easy, unambiguous exchange of information.

2022-08-24

Attending: John, Stevan, Margaret, Mary

Regrets: Elie

Agenda

Out of band meeting, for ASM planning. See assignments from 8/10

Notes

From John:

I've created a draft presentation using Google Slides for the Introduction to the Units session at the ASM in the shared IM folder under working groups/unit dictionary.

The direct link is:

<https://docs.google.com/presentation/d/1wfWPqympdGmIAMzrKTPVTfdhiiCdAUFC23xMkgawvnc/edit?usp=sharing>

However, I've left you two the hard parts! Slides 5 & 6 are title-only. But don't feel restricted to those, you are also welcome to rework the historical background and motivation. I look forward to seeing what you come up with!

What can UNIT DICTIONARY services do? Provide EML Snippets for units

- Allow integration with ezEML and other editors

- Needs a stable back end - API

Focus on SERVICES that sites can use or can be incorporated in site systems and ezEML etc.

Need EDI and site endorsement

Service that can be integrated with many others, e.g., ezeml

Service should be hosted by EDI.

Would be good to have corpus of all the units used in EML documents - with count of sites and data packages that use each unit

- Also helpful to have stmml definition

- Could be used to assess suitability of different resources (e.g., QUDT)

Could then throw it at QUDT or mapping tool (future)

Good news:

There is a lot of units work going on right now, so we have stuff to pick from.

ASM Session outline

Intro talk - 20 min

Outcome: IMC understands what WG has been doing, endorses where we want to go next

Best backend (wg's opinion is likely to be QUDT)

Discussion:

Services needed for units

- Minimum: service returns STMML for custom units

- Service contains units of interest to IMC

WG next steps (please endorse)

- Open communication with back end [QUDT, or equivalent]

- Assemble list of units for contribution to backend

- Identify interfaces to backend

2022-08-10

Attending: Mary, Margaret, Elie, John, Corinna

Regrets: Marina

Agenda

Dig into QUDT (and UOM, see below). Ensure that services exist (or can be written) that will meet our needs.

Potential activities with LTER IMs for ASM (90 min in person session, late September)

Question: are the OBOE units in UOM taken from the EML list? If yes, is it 2.2? (appears to be 140 units total)

https://github.com/mobb/units-of-measurement/blob/main/units_of_measurement/resources/mappings.csv

Notes

QUDT - to answer:

1. Does it contain units we need?
 - a. EML2.2
 - b. Others, per data pkg survey,
 - i. https://github.com/EDlorg/unit-registry/tree/master/edi_unit_analysis
 - c. See https://www.qudt.org/doc/DOC_VOCAB-UNITS.html
 - i. ? other querying tools?
 - ii. Fairly sophisticated
2. What is mechanism to add units:
<https://github.com/qudt/qudt-public-repo/wiki/Unit-Vocabulary-Submission-Guidelines>
3. What is mechanism to extract (for EML) [likely, we will write this]
 - a. As <stdUnit>, <customUnit>
 - b. As STMML for additionalMetadata
 - i. Need resource that can provide the information needed to populate STMML
 - c. Can we use UOM for this?
 - i. See <https://github.com/units-of-measurement/units-of-measurement>
 - ii. Mappings include ecoinfo
https://github.com/units-of-measurement/units-of-measurement/blob/main/units_of_measurement/resources/mappings.csv
 1. Could be used to map EML standard units to other standards (e.g., UI)
 2. But what about non-standard units
 - d. Annotations? Could include URI for standard unit in other systems
4. Homework

Critical to ASM workshop:

PPT assignments - draft

1. Pin down the why - why do we need a units resource? JOHN - draft
 - a. What are the challenges people have encountered with units? MARY
 - b. What more would people like to do with units that we can't do now? STEVAN
 - i. Draft:
<https://docs.google.com/presentation/d/1wfWPgymPdGmIAMzrKTPVTfdhiCdAUFC23xMkgawvnc/edit?usp=sharing>
2. Discuss resources (intellectual) (e.g., QUDT, units-of-measurement.org)
 - a. Comparisons of resources (matrix?) MARGARET
 - i. Attributes: look at existing table
(<https://docs.google.com/spreadsheets/d/1kt11SsthejOeRffPH39mUp-jRuPUAGJUoR2bMg970o/edit#gid=0>)
 - ii. Other attributes: future use - e.g., including a unit as an annotation
3. What are next steps that we should pursue? CORINNA
- 4.

Other potential (non PPT) activities

- a. Discussion/examples - smaller groups?
- b. Define requirements/needs

Ad hoc meeting: 2 weeks from today, 8/24 same time (noon eastern)

Note from Matt via slack:

Matt Jones 7:17 PM

@mob I just learned today of this units harmonizing service that Kai made:

<https://units-of-measurement.org> For example, here's the mapping showing OBO, NERC, QUDT, OBOE etc for velocity in m/s: <https://units-of-measurement.org/m.s-1>

7:17

thought it might be of interest for your units working group

2022-07-xx

canceled

2022-06-08

Attending: Margaret, Mary, Stevan, Kyle, Matt, Corinna, Marina

Regrets: John, Elie

Agenda

1. Consider changing time to 1 hr earlier, to better accommodate a new member from France. (complication: we may also have a new member from Alaska - in the opposite direction)
2. News
 - a. ASM workshop accepted (schedule TBD)
 - b. margaret attending units summit at IDW (virtual, week of june 20)
3. Review candidate lists:
<https://docs.google.com/spreadsheets/d/1kt11SsthejOeRffPH39mUp-jRuPUUpAGJUoR2bMg970o/edit#gid=0>
- 4.

Notes

Stevan: what is status of D1 work with units? D1 is using qudt quite a bit, but has not got a solution.

List of unit onts that D1 supports: ____ (TBD, ask Bryce, all are in bioportal)

D1 project (psyncie?) extracts annotations, creates searches on annots

Is there a mapping between EML and qudt? Yes, but old (Josh)

Unit is a required part of EML,

Still need a mechanism that does not interfere with current workflows for creating EML.

EDI would consider adding a unit annotation, if a mapping existed

Mapping depends on how target is assembled. E.g., some features (orders of magnitude) are implied in qudt.

KnowWhere graph: <https://knowwheragraph.org/>

reference is done thru the id. The ontology repo will have a

Example: dimensionality, e.g., in qudt, a quantity kind collects all the units that have the same dimensionality (similar function to EML unitType, although that is not consistently used):

https://github.com/qudt/qudt-public-repo/blob/master/vocab/quantitykinds/VOCAB_QUDT-QUANTITY-KINDS-ALL-v2.1.ttl#L124

Action items:

Margaret: contact Mark Schild - ask about his comparison of qudt and uo.

Mary: look for qudt users

Margaret - start a spreadsheet on unit mappings, for recording units that are commonly used and their ids in different systems.

<https://docs.google.com/spreadsheets/d/1RnzuWH8ObXNtUqUSOMOrO4UxvYDQbtOJHM6jCNrLKmU/edit#gid=0>

All - populate the spreadsheet.

2022-05-11

Attending: O'Brien, Martin, Zollo-Venecek, Porter

Regrets: An, Stevan

Agenda

1. ASM Workshop proposal submitted:
 - a. https://docs.google.com/document/d/1fUfckWg0y_27DnHYszVbW4Hqytx6orwgdIVMS6Goc14/edit
 - b. Evolving plans for that event
2. Continue discussion of handling new units -
 - a. Examine CF-conventions (<https://github.com/cf-convention>) can we use this type of system?

Notes

- Adopting a GH Issue-type request system for new units (see the CF conventions page)

Are there mature units management systems out there that we could hook up to?

We need a list of units models, features, strong points. Evaluate how well the EML and contributed (customUnits) match to what these hold.

Candidates:

<https://www.unidata.ucar.edu/software/udunits/>

<https://bioportal.bioontology.org/ontologies/UO>

<https://www.ebi.ac.uk/ols/ontologies/om> (<https://github.com/HajoRijgersberg/OM>)
<http://www.qudt.org/> (<http://qudt.org/2.1/vocab/unit>)
<https://github.com/nvs-vocabs/P06>

Possibilities

NERC vocabularies:

There may be some other unit lists here: https://vocab.nerc.ac.uk/search_nvs/

Elie Arnaud/Yvan LeBras group (they are heavy users of EDI and R utils)

Comment from Elie Arnaud, on units that could be added to the EML list.

<https://github.com/NCEAS/eml/issues/378#issuecomment-1125101282>

The longer thread is with the EML-R package (<https://github.com/ropensci/EML/issues/343>)

Which links to the udunits (ucar) list:

<https://www.unidata.ucar.edu/software/udunits/udunits-2.2.28/udunits2-common.xml>

STORET replacement?

What took the place of STORET? It seemed to be pretty sophisticated, but is considered vestigial. - Reporting Units List was only 26 (A-Z). unit-content was mixed, some codes contain multiple entries, some are quantities.

NEON/Christine Laney

1. Probably related to NEON's goals of standardizing some names (measurements, etc)
<https://ucum.org/ucum.html>

Action item:

Come up with a list of places we should look for:

Potential partners

Potential data (unit lists)

WG will Evaluate:

- Managers availability to collaborate

- Programmatic access

- Entries

- Compare entries to ours (EML-std, unit registry, other custom)

Custom units from EDI (Kyle brought up this need):

These should be culled down to units which follow rules in:

Units BP:

https://github.com/lter/IMC_WG_inactive/blob/master/unitsWG/attachments/LTERunitBestPractices_V14.pdf

See other pages here:

https://github.com/lter/IMC_WG_inactive/tree/master/unitsWG

2022-04-13

Attending: John, Stevan, Kyle, Margaret

Regrets: (Mary will be late), Chris

Agenda:

1. Update on data package, an archive of current DB
 - a. (temp link to package in ezeml. packageld will be edi.1116.x)
 - i. <https://tinyurl.com/y5e7b3rb>
 - b. someone please check that you can reinstall (from sql). Thx.
 - i. DOES NOT import: ERROR 1146 (42S02) at line 317: Table 'unitregistry.unitscope' doesn't exist
 - ii. May want to try using "export" option in phpmyadmin
 - c. Contents:
 - i. 9 tables, 1 view (CSVs)
 - ii. ERD
 - iii. Documentation (PDFs)
 - iv. SQL dump (entire DB)
 - v. BP doc
2. how to handle documentation? BPs?
 - a. One possibility is: <https://github.com/EDlorg/unit-registry>
 - b.
3. Use cases: which do we have the material to (see March notes)
4. Outline some eventual goals
 - a. What other material would be needed (ie, for other use cases)
 - b. how much work is it to get additional material there
 - c. How do we store the unit-resource? (xml, db, text, data package?)

Notes

Location of old LTER IMC Units WG materials. PDFs are in the folder called "attachments":
https://github.com/lter/IMC_WG_inactive/tree/master/unitsWG/

E.g.,

https://github.com/lter/IMC_WG_inactive/blob/master/unitsWG/attachments/LTERunitBestPractices_V14.pdf

Discussion

What had been done with old unit dictionary

How to manage for the future? Tematres (used for controlled vocab - requires PHP and MySQL or Maria)?

Challenge with unit types and quantities - density of animals vs density of fluid,

UDUNITS (need key, won't work with moles or anything where the specific substance is needed) - is part of EML 2.2

Workshop at ASM?

Prep: Analyze EML units and unit dictionary units

Does EzEML help?

Synonyms of standard units

Could import unit dictionary into Tematres?

Or have Unit Dictionary up (as a container?)

Goals: Lay out what has been done

Work on Cleaning SI parents - Google sheets?

How do we create a framework for new units going forward?

How does community want to manage this?

Tools for updating existing EML documents' units or search tools for units that automatically include unit aliases (cm vs centimeter). Can we use annotation for target of unit (moles annotation: CARBON)

FOLLOW UP NOTES (added by margaret)

2. for mgt process, look at what CF conventions people do. That is a living vocab of measurement names and definitions. In git. And units (? only base units, I think)
 - a. <https://github.com/cf-convention>
3. Christine Laney sent this link. Probably related to NEON's goals of standardizing some names (measurements, etc) <https://ucum.org/ucum.html>
- 4.

2022-03-09

Attending: Margaret, Hillary, An, Kristin, John, Mary

Regrets: (Chris Turner)

Agenda

1. Update on data package, an archive of current DB
2. Use cases for units tool(s)
3. Review/list existing units tools

Notes

Use cases:

1. A Reference for how to construct units properly. E.g., is it cm-squared or square-cm
 - a. There is a doc (pdf) in the old units material, but often an example of an existing unit is enough.
 - b. The unitList in EML2.2 can be filtered to units that are not deprecated. We hope that ezeml (confirmed) and EAL uses that feature (to be confirmed)
2. Better browsing: most forms (ezeml) start from the beginning of the string. So if you need a unit for "m2", you might start typing "square..." and you would not find the correct unit (meterSquared), which is already there. So either you have to browse through the entire list, or try again. Some people will just create a custom unit for "square meter".
3. Easier synchronization with relational databases. E.g., Units in core-metabase are in a table(s). So some kind of export to text. (code to work with the syncing DB is up to the individual doing the importing)
4. A way to see units that go with quantities. E.g., return all the units we have for "population density". Units that might be returned could be numberPerMeterSquared, numberPerLiter, numberPerCentimeterSquared.
 - a. We've often used "count" for this ("countPerKilometerSquared") - it seems more specific than "number."
5. Aid in automated conversion of units to aid in integration of datasets
 - a. Challenge, for some the UNIT is NOT sufficient, since you sometimes need the target of the unit - e.g., grams carbon to moles carbon requires knowledge of the molecular weight of carbon. This might be fixed by developing standard ways of defining an annotation giving the substance in question (or even its molecular weight).
 - b. Udunits synonyms are part of the eml 2.2 list (for some units). It is not part of the old mysqlDB unit dictionary.
 - c. It might be interesting to do a survey of existing units to see which ones ARE interconvertible. Identification of frequently used units would help prioritize identification of conversions.

6. Search by unit. See HFR example below. Might be of help in determining data that is fit for a particular use. Or understand how to use an odd unit like meterSquaredPerKilogram
7. Develop a reference system that lists frequent synonyms of standard units and the standard unit they refer to. Could be used to search for units that aren't standard, but should be. Could even be used to do search and replace to convert non-standard units to standard.
8. aadf

Existing tools:

HFR search by unit:

Mary pointed out this link: HFR has put their custom units into a searchable format (eXist DB).
<https://harvardforest1.fas.harvard.edu/exist/apps/datasets/da-list-custom-units.html>

It searches datasets by unit, and then returns (not the dataset - but) the description of the unit in that dataset. So it helps illustrate how/when a unit has been used.

<https://harvardforest1.fas.harvard.edu/exist/apps/datasets/da-unit-details.html?id=HF004&unit=micromolePerMeterSquaredPerHour>

2022-02-09

Regrets: Anna LaBry (rmlabry@gmail.com)

Present:

Suzanne Remillard (Andrews LTER)

Stevan Earl (CAP LTER)

Kyle Zollo-Venecek (EDI, kylezollo@gmail.com)

Chris Turner (NGA)

Kristin Vanderbilt (FCE LTER)

Sarah Elmendorf (NWT LTER)

Mary Martin (HBR LTER)

Mark Servilla (EDI)

Margaret O'Brien

An T. Nguyen (BLE)

Agenda

1. Introductions (all)

2. Review status (margaret)
3. Assess interest in a working group for Units
4. Find a regular meeting time
 - a. Poll: <https://www.when2meet.com/?14633588-2b7H7>
5. Topics for next meeting

Notes

Lots of material in the git repo: <https://github.com/EDlorg/unit-registry>
includes mysql DB dump

Also see issues associated with the creation of the EML 2.2 units list:

<https://github.com/NCEAS/eml/issues?q=is%3Aissue++units>

Mention of things like unitType, dimensions, and quantities there.

General discussion:

Stevan: We need a method to understand how to create new units, We don't much need a way to browse stuff that is already in EML.

Sarah: what audience do we want to serve?

New IM, who needs training (primarily in EML)

Someone who wants to learn about any unit

Suzanne: method to understand complex units (browsing, search was nice). Note: AND DB units/quants tables are similar to these.

Sven: do we know update frequency? Yes, see table called change log

Kyle: having a unit-interface might help people submitting data to EDI to create a "good" unit, and save Kyle some time. He thinks main conversation is about adding new units

Use case: Reading?

Writing: If list was a simple XML file (w or w/o an interface,) one model could be that to add a new unit, submit a git issue

Need a gatekeeper

Mary - we would want a dynamic list to be readable via EAL.

If a new unit was added (ie, via git pull), then we would want them to be available to EAL.
(sounds like the web service from the old PHP).

Question for Chris: what does axiom do? The use external dictionaries. LTER dictionary, UDUnits, (UDUnits are good for gridded). NetCDF std names are good too (Physical oc). but there are too many edge cases for one dictionary to meet all needs.

Future topics:

Possible back ends:

sqlite db.

XML

JSON.

Action items

Send out follow up email, try for another meeting time.

Find notes from ~2010? Rules for how to construct a unit (from old IMC website)

DONE: Units BP:

https://github.com/lter/IMC_WG_inactive/blob/master/unitsWG/attachments/LTERunitBestPractices_V14.pdf

See other pages here:

https://github.com/lter/IMC_WG_inactive/tree/master/unitsWG

Follow up - HFR search by unit:

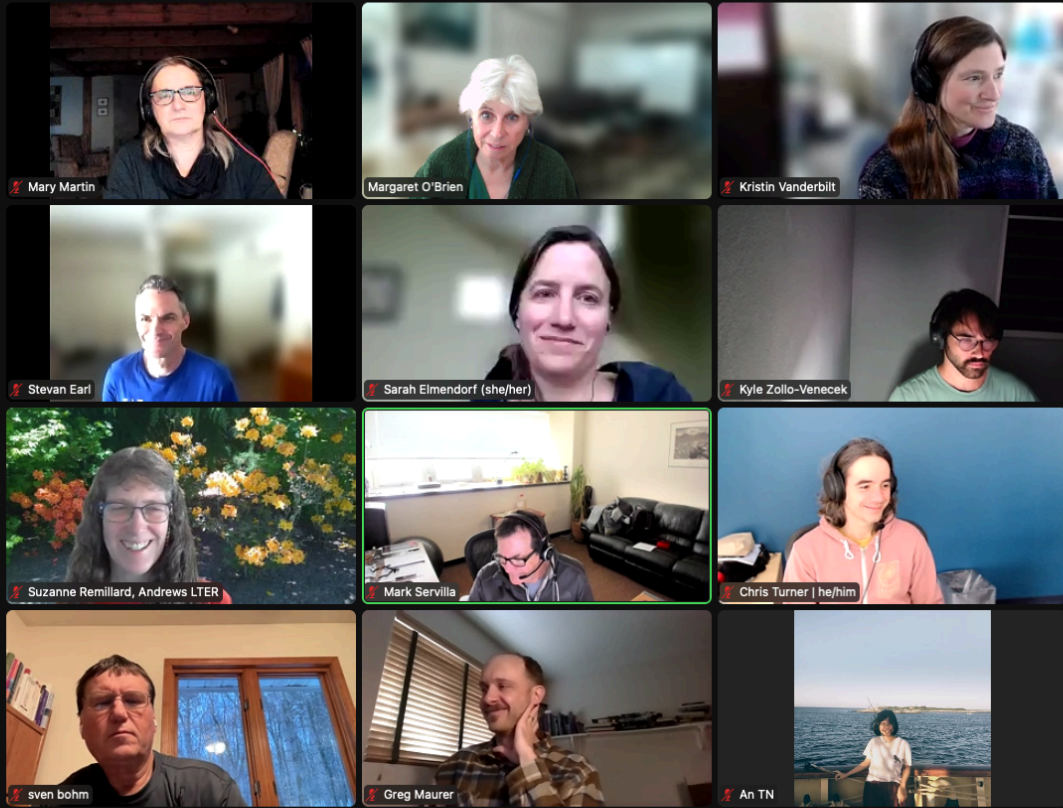
Mary pointed out this link: HFR has put their custom units into a searchable format (eXist DB).

<https://harvardforest1.fas.harvard.edu/exist/apps/datasets/da-list-custom-units.html>

It searches datasets by unit, and then returns (not the dataset - but) the description of the unit in that dataset. So it helps illustrate how/when a unit has been used.

<https://harvardforest1.fas.harvard.edu/exist/apps/datasets/da-unit-details.html?id=HF004&unit=micromolePerMeterSquaredPerHour>

Who we are:



John's analysis, unit element missing

When QUDT Unit Elements Aren't Found in QUDT as Units

Some complex units (containing more than one unit element) contain unit elements that are not found by themselves as individual QUDT units. For example, the complex unit KiloM3-PER-SEC2 contains the unit elements KiloM3 and SEC2. However, only SEC2 is found in QUDT as a QUDT unit by itself. KiloM3 is only found as part of a unit, never as a unit itself. QUDT units can be identified because they do not contain a dash (“-”).

Units raised to powers (e.g., M3 for cubic meter), having prefixes (e.g., Kilo, Mega) and containing qualifiers (e.g. GAL_US) are treated as distinct unit elements in this analysis.

Of the 2821 non-deprecated units defined as QUDT units in release 3.2.1, there are 940 (33%) single-element QUDT units and 1881 (67%) complex units. The complex units are comprised of 389 distinct unit elements and 4564 unit elements overall (allowing repeats).

Combining the QUDT units with the distinct unit elements derived from complex units yields 969 distinct units. Most elements are used only as single-element QUDT units (62%, 580/969). An additional (37%, 360/969) appear both as QUDT units and as parts of complex units. Only 3% (29/969) of unit elements are found only as parts of complex units, but not as single-element QUDT units.

Of the 29 base units not found as QUDT units in QUDT, all involve exponentiation (e.g., M3), 6 also contain prefixes (e.g., Kilo, Centi, Milli) and 2 contain qualifiers.

Despite the small number of base units not found as QUDT units, 51 of the 1881 complex units use one or more unit elements missing as single-element QUDT units. Thirty-three of the base units were used only once, and only 10 of the 29 unit elements were used two or more times, and only 2 more than 5 times: IN6 and L2 (both used 6 times).

Unit Elements that are Only used in Complex Units

Unit	Number of Complex Units Using This Unit Element
A2	1
C2	1
C3	1
C4	1
CentiM5	1
EV2	1
FT6	3
GAL_UK2	3
GAL_US2	3
GigaEV2	1
HZ2	2
HZ3	1
HZ4	1
IN6	6
J2	2
J3	1
K4	1
KiloGM2	3
KiloM3	1
L2	6
M0dot5	2
MicroMOL2	1
MicroS2	1
PA2	2
PlanckMass2	1
SEC3	1
T2	1
V2	1
W0dot5	1

List of Complex Units Incorporating Elements Only Used in a Complex Element

unit	element
A2-SEC	A2
C2-M2-PER-J	C2
C3-M-PER-J2	C3
C4-M4-PER-J3	C4
DYN-SEC-PER-CentiM5	CentiM5
PER-EV2	EV2
LB-HR-PER-FT6	FT6
LB-MIN-PER-FT6	FT6
LB-SEC-PER-FT6	FT6
LB-HR-PER-GAL_UK2	GAL_UK2
LB-MIN-PER-GAL_UK2	GAL_UK2
LB-SEC-PER-GAL_UK2	GAL_UK2
LB-HR-PER-GAL_US2	GAL_US2
LB-MIN-PER-GAL_US2	GAL_US2
LB-SEC-PER-GAL_US2	GAL_US2
PER-GigaEV2	GigaEV2
M2-HZ2	HZ2
M2-PER-HZ2	HZ2
M2-HZ3	HZ3
M2-HZ4	HZ4
LB-HR-PER-IN6	IN6
LB-MIN-PER-IN6	IN6
LB-SEC-PER-IN6	IN6
OZ-HR-PER-IN6	IN6
OZ-MIN-PER-IN6	IN6
OZ-SEC-PER-IN6	IN6
C3-M-PER-J2	J2
PER-J2	J2
C4-M4-PER-J3	J3
W-PER-M2-K4	K4
KiloGM2-PER-SEC2	KiloGM2
N-M2-PER-KiloGM2	KiloGM2
PER-KiloGM2	KiloGM2
KiloM3-PER-SEC2	KiloM3
GM-HR-PER-L2	L2
GM-MIN-PER-L2	L2
GM-SEC-PER-L2	L2
MilliGM-HR-PER-L2	L2
MilliGM-MIN-PER-L2	L2

MilliGM-SEC-PER-L2	L2
MegaPA-M0dot5	M0dot5
PA-M0dot5	M0dot5
MicroMOL2-PER-M4-SEC2	MicroMOL2
MicroS2-PER-CentiM2	MicroS2
PA2-PER-SEC2	PA2
PA2-SEC	PA2
	PlanckMas
PER-PlanckMass2	s2
KiloGM-PER-SEC3-K	SEC3
J-PER-T2	T2
V2-PER-K2	V2
N-M-PER-W0dot5	W0dot5