

Planteome Ontology Working Group Zoom Meeting

Time: Feb 21, 2017 8:15 AM Pacific Time (US and Canada)

Join from PC, Mac, Linux, iOS or Android: <https://zoom.us/j/370267127>

We are currently having some issues with our wiki site, but hopefully they will be resolved by tomorrow morning.

You can preview the latest ontology and database release here: Planteome Release 2.0 on <http://dev.planteome.org/amigo>

Agenda:

Discuss Workplan for Year 3 for the Ontologies group

- What are the goals for the coming months?
- Review what we listed in the last annual report...

PJ: We need to review what we proposed doing in the grant, and see what is still pending

- Building community interfaces and engagement on adding ontology terms and annotations
- Plant Stress Ontology
- Enrichment of TO for terms where we do not have a one-to-one match, in the mappings, for example from CO to TO
 - Enrich TO and other ontologies for terms for crops- should have 1-1 mappings
- CM: the ontologies need to be more integrated in annotations and need to be pulling in more annotations, like we are doing with GRIN- this will bring in more ontology terms (?)
- PJ: Issues with PATO as well- need to address some of the outstanding problems...
 - Should we continue to use it? Yes
 - Some definitions need major revisions
 - Who can make changes- CM and George are the main editors right now
 - Need to make sure we have open tickets for the issues....
 - Need appropriate terms to describe leaf morphology, shapes, margin types

- Need big effort in making sure these are in PATO-
<https://github.com/pato-ontology/pato/issues>
- Need to set a platform for next round- proof of concept
- Field is moving into the HPT phenotyping- lessons from Phenome 2017..
 - Automated learning, deep learning...finding new terms
 - HPT phenotyping traits are mathematical expressions, remote sensing data
 - How to engage with the HPT phenotyping community?
 - EA has met with Jesse Poland (KSU) - he is happy to discuss with us how to adapt the ontologies to the needs of the HTP community
 - Advances in the area towards farm to fork-- working with Industry partners, storage and management, processing plants,
 - Need vocabs and data standards on that side
 - what is the major impact of this work on the community
 - Data collection by sensors- alert farmers to water, App
 - Guiding automated processes
 - EBI meeting in April: industrial partnerships- food, quality, measurements , analytics
- CM: Need to also address the issue with EO ?? (difficult to hear on the recording...)
- Vision for wider focus and impact- idea for new grant?
 - Discussions with private start up company: Ripe.io - manages quality, measures
 - USDA NIFA - RFPs for RCN type projects...
 - Community workshops, take inputs, generate white papers, develop standards and analytics
 - Involve Matthew Lang- food ontology
 - difficulties getting response from him...
 - Vision and deliverables for industry, ontologies on back end
 - Applying for PhenRCN type grant,
 - BS: involved with NIST?
 - NIH grant MOTORPAK: Biochemistry of exercise/ ontology of physical activity, food relevant dimensions, looking for partners to get involved or contribute

- What activities will be left to do by November 2017?
- Possibility of a 1 year no cost extension? Have to have the \$
 - All the partners should review their budgets
 - Need to have a clear plan
 - Review goals in grant - see annual report 2016

CM: need to get Triple store up and running

- Tools for viewing the ontologies
 - APIs
- From CO: new crops coming: Brassica, sorghum, yam, sugarcane
 - Goal should be to use automated mappings and follow up with manual check
 - At what stage should the mappings take place?
 - Can we have the automated mappings when the TDs are created?

Areas of TO need to be addressed:

- canopy/population traits
- Eg. Yam ontology- differences in size of tubers
 - They want to have different classes for the weight of small, medium and large tubers- how to do this in TO?
 - Issue of overlapping equivalence axioms...
- Need to flesh out the EA in the TO- agree on the patterns
- Roseae traits coming up

From PJ: Semi automated approach should work, need to have people to contribute in an open way, understand what an ontology is,

- Manual creation of terms
- Need automated way to introduce new terms, bottlenecks
- NYBG contribute to TO??

MAL:

- We have started to do the mappings in a semiautomatic way, followed by manual checking
 - Based on EQ statements
 - Sorting out how to automatically add the terms to the TO
 - Issues with the TO is making the EQ patterns difficult
 - By instituting the design patterns, then the format of the definitions will change- no long will be genus differentia format and will not include the term IDs as we have been doing.
 - *Why is this the case? Can't we set up the design patterns to formulate them however we wish?*
- A given trait can be classified under different super classes and using the genus differentia format, you have to make a choice which one to use. Using the DPs, the classification is automatic.

PJ: need to make sure the TO terms are consistent with the EQ

LC: Need to work with PATO and make sure we are using the qualities in a consistent manner

CM: TermGenie will be replaced with a new better tool this year. IncaTools is the foundation for that

EA: Need guidelines for people developing the trait dictionaries-

- eg Elixir project, developing TD for tree species
- What are the recommended best practices?
 - Guidelines are in place for TD development, but need the next steps
 - Key step is the mapping to TO...
 - One TD per tree species
- PJ: Species specific traits should be added to the appropriate TD for folks from outside the CG system, then can be consumed by the planteome
 - Are they open to contributions from outside groups via GH?
 - Same TD file for everyone would be best
 - EA- would not be acceptable by project managers, project specificity needed
 - Select level of granularity
 - PJ: suggest adding new column for comments- which project added specific terms, with different IDs for these
 - Keeping them project specific
 - Private companies do not want to share their info- need tools that they can use for the mapping
- PJ: Concern that versions of TDs on GH are not the same as ones used in field??
 - Recommended: Use the same platform? Work from GH..?

- Demonstration of community involvement and buy-in

EA: Example issue the Brassica group- want to have their own GitHub

- Separate group, separate project
- Need written guidelines for the mapping part
- Need to agree on the workflow

- Make a few 5 minute videos to explain best practices and what an ontology is
 - Start at the beginning to collaborate

PJ: TO needs lots of work, would be great to have input from the NYBG group

- Need to expand the fruit types into TO
- Need to make sure the CO terms have a i:i mapping in the TO
- NYBG group: Focus should be on the inflorescence and the grain fruit
- Higher granularity in the CO vocabs

- Site visit in 2017? Probably not, PJ will follow up with the NSF folks
 - By Aug/Sept should be thinking about new proposal in 2018- PI?
 - Proposal should be ready to go in Jan

- Possible Workshop in fall? Or come to
 -

GO meeting in Corvallis <https://sites.google.com/view/goc2017>

- GO meeting June 1-5
- Protein Ontology sub meeting- BS organizing, will let PJ know
- Possibility of having GC folks come

DWS: what should the next project be - seeds?

- Inflorescence types for next meeting
- Commercial application- seed qualities
- Quality measures of inflorescences as well

Deep learning app- leaf shape recognition- Being developed by Sinisa and Dimitri etc

-

Phenome 2017 update:

- HTP traits- how to address in TO
- Think about bringing in the PCO
- Prepare a short summary of the traits- couple of slides
- Share google doc of notes
- Send PJ a short summary