

2018-03-06 Ontology Improvement Call

Tuesday, March 3, 10 AM US Eastern Time

Connection Info

Join from PC, Mac, Linux, iOS or Android: <https://duraspace.zoom.us/j/952326581>

Or iPhone one-tap :

US: +16468769923,,952326581# or +16699006833,,952326581#

Or Telephone:

Dial(for higher quality, dial a number based on your current location):

US: +1 646 876 9923 or +1 669 900 6833 or +1 408 638 0968

Meeting ID: 952 326 581

International numbers available: <https://duraspace.zoom.us/j/952326581>
[more details»](#) [copy to my calendar](#)

Attendees

- [Mike Conlon](#)
- [Damaris Murry](#)
- [Violeta Ilik](#)
- [Muhammad Javed](#)
- [Christian Hauschke](#)
- [Marijane White](#)
- [Tatiana Walther](#)
- [Andrei Tudor](#)
- [Brian Lowe](#)

Agenda

1. vivo.owl and applicationConfiguration.owl – current state and next steps
 - a. ApplicationConfiguration ontology terms <https://goo.gl/vayV0D>
 - b. Namespace
 - c. JIRA here: Unable to locate Jira server for this macro. It may be due to Application Link configuration.
2. Working ontology issues
 - a. create a JIRA (or a GitHub issue – Mike keeps them synced)
 - b. fork the ontology lab VIVO repo (we're not quite ready to work directly in the VIVO repo)
 - c. create a branch named for the JIRA (vivo-xxxx)
 - d. Do the work in your fork
 - e. Make a minor pull request
 - f. Review other's PRs and approve/request-changes

Notes

Mike: Welcome, everyone.

Update on issues we've been working on. Vivo.owl and the applicationconfiguration.owl, VIVO-1447, I encourage everyone to do whatever they like to inspect these files. We're about to open a PR and pull it into the VIVO repository. We keep finding things in the file that have always been there, another issue was opened because 27 copies of the ontology source assertions, which is redundant. But we're just trying to figure out if the file contains everything it needs to run VIVO. We're interested in any other comments about the file in the next week or two, and then I will issue a PR. Then future work will be on that file.

The file is in the vivo-ontology-lab repo, if you build that VIVO starts, we think it's ready to pull into the main repo. And then future work can be in the development branch using the normal VIVO development process.

We had a bit of a false alarm about the application configuration ontology, it turns out there never was one. The assertions we found in the software had been gathered into a file, and we can post that file, but it is a separate issue, VIVO and Vitro work fine without it. But we have found all the assertions and we have a working version of the application configuration file, but that is a separate issue, VIVO-1446.

Marijane: I'm confused, the file never existed, the applications work without it, so why are we creating a file?

Mike: Because we want to know what the application configuration is

Javed: The file is in the vivo-ontology-lab repo, the assertions were pulled out of the VIVO ontology.

Mike: Yes, you pulled them out, I thought there were others, we pulled them out into their own file.

That reminds me, create an action item for me, there should be a separate issue around ontologies.owl -- We have discussed here which ontologies should be in the VIVO ontology list, but we need an issue for that, and it's not any of the issues we've been talking about.

Tatiana: Is it possible that some classes that are not in VIVO now could be imported?

Mike: Yes, there can be plenty of issues opened about things that should be added to VIVO

Javed: I think that should be a separate commit.

Mike: Yes, I agree, and that is a good segue into the next agenda item, which is how we will make these changes going forward. Each change should have its own issue. For example adding a Spinoff Company class could be useful. So if we were to add...

Tatiana: The issue is already there.

Mike: Yes. So now anyone can begin to work any issue they like. You can notify people in Slack or change the JIRA status to let others know you're working on it, that's fine. You can work in the ontology lab repo for now, it might take us a couple weeks to get into the development branch because we will have to wait for our big PR to be reviewed, it will be extensive and it will take some time.

Javed: I agreed with Andrew, I am going to lead the first sprint, I'm thinking we should include this PR in that sprint, so we can assign it to people working on the sprint.

Mike: Remind us when the sprint is?

Javed: April, I'm not sure when.

Christian: Last two weeks

Mike: That's a long way out.

Javed: But we can start reviewing now.

Mike: So we can open the PR now. And we can have branches off that commit to do additional ontology work.

The branch and the PR will be in place by our next meeting, then we can discuss how to work in the meantime. Meanwhile, if you are creating assertions for new stuff, you can do that in your own copy, and we'll figure out how to pull that in later.

Brian: I have a general question here, I feel like I have missed something. Don't we want to do this in ISF and extract it?

Mike: Eventually.

Marijane: We're not ready for that yet. But I think we can backport changes in the VIVO repo into the ISF

Mike: Look at VIVO-1460: create an extraction process

Brian: As long as we can backport it, that's fine, I'm just nervous that things might get missed.

Marijane: ROBOT has a good diff tool, we should be able to compare an extracted vivo.owl with the vivo.owl in the VIVO repository.

Brian: I think I missed the ROBOT demo.

Javed: Who is managing the OpenRIF repo?

Marijane: OpenRIF is suffering from some absentee landlordism

Violeta: This is why we need to shepherd these changes into VIVO

Brian: So should we abandon OpenRIF and stop propping up this myth of the ISF?

Mike: I had a related thought, why can't we just load source.owl into VIVO?

The idea is, why would we extract and merge, why not just load the whole thing?

Brian: We tried that in 2013, it caused some pretty serious performance problems, such as browsing the ontology list.

Mike: OK, I will test that.

Marijane: Perf issues are why people started doing MIREOT.

Mike: I'm surprised, there just aren't that many triples. There may be things in the code that make it difficult, but in general it's not that much data.

Javed: I don't think we should pull in a bunch of stuff we don't use.

Mike: Yes, there are like 20,000 classes we'd never use.

Javed: I would suggest that if we get a request, we add a class, and then we look to see if it's in source.owl, pull it from there, if not, add it to vivo.owl and then push it back to the ISF later.

Mike: I'm fine with that, we just need merge and extract functionality.

Javed: I am sharing a doc about some work I did around namespaces and whether we need to keep them. <https://docs.google.com/spreadsheets/d/1Of5S9Dt-OqRfhAtFqeXLT-khCe4EFUkdA7RxNrkovNE/edit#gid=0>

There are some things in here that are not used in VIVO, and I'm thinking about what we should do. First, we could delete them from vivo.owl, but if one person is using it, how can we help them not break their instance. Perhaps we could make modules for them so people could load them optionally. But we don't currently provide those modules.

Items in red can be removed, but everything else is OK.

Marijane: I would love to know how <http://isf> got in there.

Javed: Yes, there is an equipmentFor property in that namespace that is deprecated, it can be removed.

Mike: Let's see what the concrete actions are around this. You would like to not have the things in rows 2,3,4,5 because we're not using much from them?

Javed: no, they're just not needed.

Mike: I'm not following.

Javed: let me share my screen. You can see they are defined in the file, but there's no need to define them in the file because they're imported.

Marijane: I agree with that.

Javed: They don't hurt anything, but they're not needed.

Brian: They could actually be dangerous, if you redefine something like maxCardinality, you could wipe out it's semantics.

Javed: Yes. So that is what I'm talking about in rows 2, 3,4, 5 of the Google spreadsheet. And then we should delete <http://isf>, and I already deleted the one in row 30.

Mike: And the remaining three assertions are the good annotations on the ontology itself, the other 27 were redundant.

Javed: Yes. And I committed that change.
Other changes:

- we should use only one version of SKOS.
- Analyze suggested ontologies to be deleted
- Can ontology definition/label/prefix reside in one file? Ontologies.owl
- VIVO ontology files will use vivo as prefix everywhere.

Marijane: Will changing the prefix have unintended consequences?

Mike: Both are in use, actually, and there are actually more usages of vivo than core in the code. I do think it should be vivo: in the file we produce. What's in vivo.owl now?

Javed: It's vivo.

Mike: But there are other references throughout the code where the core prefix is used, and that should be deprecated and removed from the codebase.

Tatiana: Core is too general.

Mike: Exactly.

Christian: So we have a new issue for the development call.

Mike: Yes.

I actually went through the code, and the actual usages in the code are rare, they're actually defined in these miscellaneous RDF file that are used to load triples.

Javed: I think it will have less impact in the VIVO code, but more impact if people are extracting from VIVO using SPARQL queries.

Christian: This might impact Karma.

Violeta: Yes.

Mike: Yes, there should be an announcement.

Violeta: Are we going to scare people?

Mike: Yes.

Tatiana: I think OpenRefine also uses the vivo prefix.

Mike: The vivo prefix was introduced years ago, we just never deprecated the core prefix.

So another action item: I will generate a statement and a strawman process.

Marijane: Isn't core in the URI?

key	summary	type	created	updated	due	assignee	reporter	priority	status	resolution
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										