

Views Controlled Vocabularies (VCV) Task Group
Zoom meeting
2021-04-21 15:00 UTC

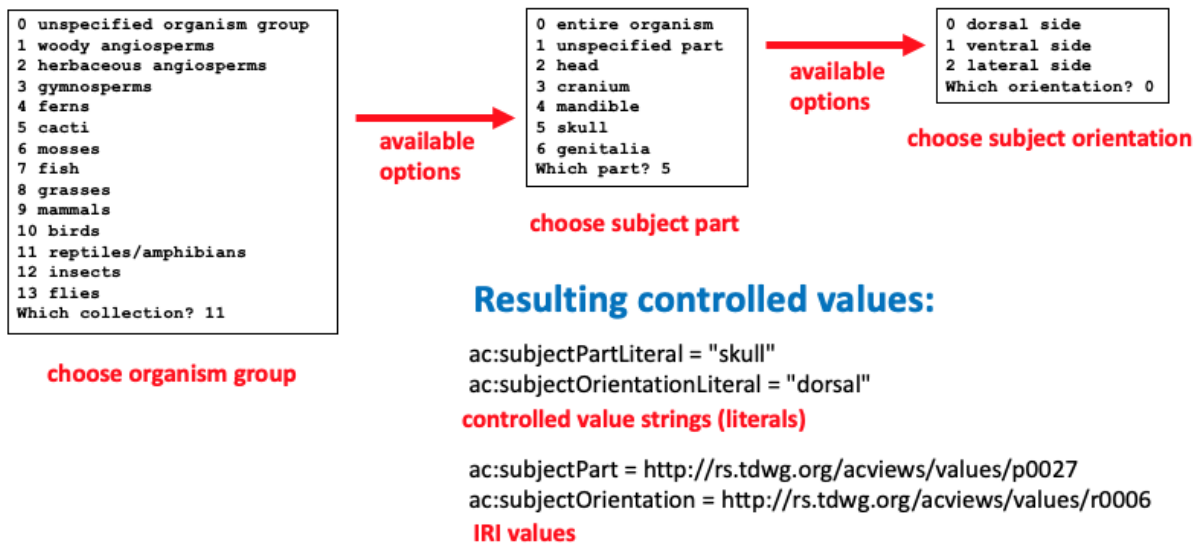
Attending: Steve Baskauf, Matthew Nielsen, Neil Cobb, Jennifer Giron

Regrets: Donat Agosti, Torsten Dikow

Meeting notes:

Notes added during the meeting and from the chat are in red.

- I. Steve meeting with Tanya Berger-Wolf (AI for wildlife ID). She confirms that standardized views and definitions for ROIs ("Region Of Interest") could be useful for machine learning. However, it seems unlikely that anyone from her group would actually participate in any of our discussions. We might be able to get them to try implementing something, though. Links to publicly accessible resources her team provided in Appendix A.
- II. Steve cleanup actions since last meeting:
 - A. In the interest of simplification, I deleted ferns, moss, cacti, and grass. When people want views for them, we can create them.
 - B. See definition and ontology links in subjectPart spreadsheet (unfinished task listed in last meeting notes, now done). Couldn't find a few parts.
 - C. New versions of everything pushed out to
<https://github.com/tdwg/ac/tree/master/views/code>
- III. Regions Of Interest report from AC Maintenance Group meeting
 - A. Example fish parasites image:
https://github.com/tdwg/ac/blob/master/fragments/transformation_examples/big_34616.jpg
 - B. Example ROI description in JSON (change terminology: "fragment" to "ROI"):
https://github.com/tdwg/ac/blob/master/fragments/transformation_examples/image_graph_zookeys.json
- IV. Status of Zenodo metadata custom keywords
 - A. Experiment with uploading to API:
https://github.com/HeardLibrary/linked-data/blob/master/publications/zenodo/zenodo_api.ipynb <https://sandbox.zenodo.org/record/750176>
 - B. Pending "ticket" on Fabricius ant holotypes.



V. Holdovers from last meeting

- A. Steve clarification about diagram above: "organism group" categories are for convenience only. No one is required to use any particular category -- they are simply for guidance. So if there is a more specific category, no one is forced to use it if they prefer a broader category. subjectPart and subjectOrientation are the actual controlled vocabularies that will be part of the standard.
- B. Comment from Torsten: I think there is no need to divide up larvae from adult insects or any subgroup of insects. Larvae have the same parts as adults such as head, thorax, abdomen, and legs even though they are often morphologically changed. Photographing larvae will also result in the same views such as dorsal, lateral, and ventral. A nice web-site with examples is <https://www.macroinvertebrates.org> that focuses on aquatic larvae (caddisflies, alderflies, flies, beetles, moths) and nymphs (mayflies, stoneflies, dragonflies, damselflies, and true bugs).

The above will also be applicable to pupae of holometabolous insects (beetles, flies, wasps, moths etc.) as well as subadults (Marcus will know the correct term) of all other insects.

Dividing insects up into subgroups is also not necessary as they all share the same basic morphology and the vast majority have four wings even though these can be morphologically changed in some groups. Halteres are technically the hindwing in flies and just received a special term rather than saying 'stub-like hindwing'. Flies couldn't fly without halteres ...

This deals very specifically with candidate requirements 2.1 and 2.2 at <https://github.com/tdwg/ac/blob/master/views/candidate-requirements.md>. Do we

just eject those from the requirements? Neil brought 2.1, his thoughts? Neil was in favor of taking the simple approach.

- C. Steve thinks: don't need separate organism group/part/whatever for flattened specimens. The mix of orientations for whole pressed herbarium specimens is similar to whole plant photos in general. We can know it's a specimen by preparations or some other DwC term.
- D. For mammals and herps, is "body" (separate from "whole organism") a thing? Notes from last meeting: There was a lot of discussion about this issue. We really are more concerned about functional definitions rather than theoretical ones. An "organism group" is really more about kinds of organisms that share the same parts and less about taxonomy. "subjectParts" that have the same possible orientations can be considered the same "part" even if they technically are different: for example, a skin that is wrapped around a stick and has all of the same orientations as a whole organism should probably just be called a whole organism even though technically it's only part of the organism.
Steve proposes that we treat these as "whole organism" and break them down into categories when somebody shows that they need it.
- E. Add male/female genitalia subcategories to genitalia leave general for now, sort by dwc:sex.
- F. perhaps a 'part of' term (?). Didn't discuss last time. I can't remember what this means, different than the generic skos:broader relations currently using? this was with respect to only part of a part being in view. Let's leave it.
- G. Antero-lateral/ oblique view? [see [Plate9_Heads.jpg](#) and [C_auricephalus_TTU-Z_219296-obl-ed.tif](#)]. What do we call this??? probably close enough to a lateral view that we should probably not add it and add it later if there is demand.
- H. What to do when orientation is uncertain from the image? Clarify definition of "unspecified orientation" -- is that a placeholder for images whose orientation has yet to be determined? Is it different from "we can't tell what the orientation is", which might be different from "the orientation doesn't fall into one of our standard categories"? Whatever these categories are, they should probably be added to every part collection. use "ambiguous" rather than undetermined. If a value has not been determined, tell people to not provide a value. Action item: change existing value of "undetermined" to "ambiguous"

VI. Remaining unresolved tasks

- A. Review Candidate Requirements
(<https://github.com/tdwg/ac/blob/master/views/candidate-requirements.md>) to determine which are satisfied, which are not satisfied but should be, and which should be deleted because they can't/won't be satisfied. This will be the major focus of our next meeting. Core members should review the candidate requirements so that we can discuss them next time.

- B. Do we want to push off "best practice guides" as a second stage (but with this TG still around)? (chicken and egg again). Provide examples, recipes, but not an actual guide
 - C. Identify test implementations: we need to come up with a set of instructions. Explain how to use with CSV.
Suggestion: Provide a how-to guide for people who are going to test the terms:
 - How many images (place them in a shared folder?)
 - [Provide template with examples to fill out]
 - Choose terms from CSV files for subjectPart and subjectOrientation and fill out template file
 - Answer a few questions about the exercise
 - How complicated was it to match image to terms
 - Which images caused problems/confusion
 - Do the general terms provided satisfy basic documentation?
 - Do you *need* additional terms?
 1. Zenodo Fabricius ants
 2. Bioimages plant images
 3. iDigBio?
 - D. Think about 3-5 people who could test with different taxonomic focus
 1. Non-insect invertebrates
 2. non-vascular and non-flowering plants
- VII. Next meeting time: Wednesday, May 19. 15:00 UTC

Appendix A. Overview of terminology and workflow used by Wildbook:

Here's a good overview of what we call things:

<https://docs.wildme.org/docs/researchers/overview>

and our ML pipeline:

https://docs.wildme.org/docs/researchers/ia_pipeline

at the nitty gritty level for object definition of MediaAssets (e.g., photos) and Annotations (bounding boxes on MediaAssets as created by ML or manually drawn where ML misses one):

MediaAsset:

<https://github.com/WildMeOrg/Wildbook/blob/master/src/main/java/org/ecocean/media/MediaAsset.java>

Annotation (likely much more useful):

<https://github.com/WildMeOrg/Wildbook/blob/master/src/main/java/org/ecocean/Annotation.java>

And a visualization of a MediaAsset and a single Annotation (a whale_fluke)

<https://www.flukebook.org/obrowse.jsp?type=MediaAsset&id=2298343>