Implementation Experience Report for Controlled Vocabularies Used with the Audubon Core Terms subjectPart and subjectOrientation

Steven J. Baskauf[‡], Jennifer C. Girón Duque[§], Matthew Nielsen[|], Neil S. Cobb[¶,#], Randy Singer[¤], Katja C. Seltmann[«], Zachary Kachian[»], Mervin Pérez[^], Donat Agosti[ˇ], Anna M. L. Klompen[¦]

‡ Vanderbilt University Libraries, Nashville, Tennessee, United States of America
§ Natural Science Research Laboratory, Lubbock, Texas, United States of America
| University of Oulu, Oulu, Finland
¶ Northern Arizona University, Flagstaff, United States of America
# Biodiversity Outreach Network, Flagstaff, United States of America
¤ University of Michigan Museum of Zoology, Ann Arbor, Michigan, United States of America
« Cheadle Center for Biodiversity and Ecological Restoration, University of California - Santa Barbara, Santa Barbara, California, United States of America
» Keller Science Action Center, Field Museum of Natural History, Chicago, Illinois, United States of America
^ Instituto de Investigaciones, Centro Universitario de Zacapa, Universidad de San Carlos de Guatemala, Zacapa, Guatemala
ˇ Plazi, Bern, Switzerland
¦ Dept. of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas, United States of America

Abstract

The Audubon Core vocabulary terms subjectPart and subjectOrientation are used to describe the depicted part of an organism and its orientation in an image. We describe the criteria and process for developing controlled vocabularies for these two terms. The vocabularies take the form of Simple Knowledge Organization System (SKOS) concept schemes and their terms are categorized using SKOS collections to allow users to select from particular sets of values appropriate for particular organism groups and their parts. We also report the results of implementation testing used to determine the usability of the proposed terms with actual images of living organisms and preserved specimens.

Keywords

Introduction and background

Publications in biology, especially those related to biodiversity, generally include images to portray habitus and morphological structures of the organisms that are the subject of those works. An increasing number of repositories are making available large collections of digital images of organisms and their parts, drawing on sources ranging from natural history collection digitization to citizen science platforms. These images are increasingly used as a source for many forms of research, including morphological measurements, identification of individual organisms, and automated taxonomic identification. Especially as repositories become larger and research more automated, usefulness of these images for research will increase greatly because of the ability to describe the depicted part of an organism and its orientation using a controlled vocabulary.

Morphbank :: Biological Imaging*[1] is an online repository project that began in 1998 and expanded after receiving funding from the National Science Foundation in 2003 to eventually include over 216 000 images of biological specimens and living organisms*[2]. A key feature of Morphbank was the ability to define characteristics of the image according to the part of the organism depicted, the orientation of the part (or the organism) in the image, and the imaging technique. The set of characteristics of the image was called a "view". Image uploaders could select from an existing view, or define a new view if the characteristics of their image differed from available existing views. This provided great flexibility, but also resulted in a great proliferation of views, eventually reaching over 17 000 possible views*[3]. Recognizing this problem of proliferation of views, Baskauf and Kirchoff (2008) attempted to standardize views for live plant images, resulting in the creation of collections of recommended Morphbank views for four categories of live plants*[4].

The Audubon Core Multimedia Resources Metadata Schema, often referred to simply as "Audubon Core" and abbreviated as AC, is a set of metadata vocabularies for describing biodiversity-related multimedia resources and collections (Morris et al. 2013a, Morris et al. 2013b). Morphbank was a key stakeholder in the development of Audubon Core, and when Audubon Core was originally ratified as a standard, it contained two terms that captured key aspects of Morphbank views: ac:subjectPart (http://rs.tdwg.org/ac/terms/subjectPart), defined as "The portion or product of organism morphology, behaviour, environment, etc. that is either predominantly shown or particularly well exemplified by the media resource." and ac:subjectOrientation (http://rs.tdwg.org/ac/terms/subjectOrientation), defined as "Specific orientation (= direction, view angle) of the subject represented in the media resource with respect to the acquisition device." When these terms were introduced, it was recognized that there was no formal controlled vocabulary for values of the terms, and it was assumed that the values would be text strings.

Soon after the formation of the Audubon Core Maintenance Group in 2018, it chartered a Views Task Group*[5] whose task was to develop controlled vocabularies that conformed to the TDWG Standards Documentation Specification (SDS; Baskauf et al. 2017b) for the terms ac:subjectPart and ac:subjectOrientation. One consequence of following the guidelines of the SDS was distinguishing between controlled strings and the internationalized resource identifiers (IRIs) that denote the concepts associated with those strings. The convention established in Audubon Core was to have pairs of analogous terms where one term intended for use with IRI values had a simple local name (e.g., ac:reviewer) and an analogous term intended for use with controlled strings had the same local name with "Literal" appended (e.g. ac:reviewerLiteral). To conform to this convention, the 2022-02-23 version of Audubon Core revised the definitions of ac:subjectPart and ac:subjectOrientation to clarify that their values are denoted by IRIs and created new terms ac:subjectPartLiteral and ac:subjectOrientationLiteral to be used with controlled string values*[6].

As a coordinated addition to Audubon Core, the Views Task Group has followed the process outlined in Section 4 of the TDWG Vocabulary Maintenance Specification (VMS; Baskauf et al. 2017a). Specifically, the Task Group generated a Feature Report based on input from the community and this paper comprises the Implementation Experience Report produced by the Task Group to satisfy the requirements of Section 4.2.2 of the VMS. The purpose of this report is to demonstrate that the desired features of the vocabularies generated from use cases submitted by the community were implementable by the contributing testers. The hope is that this report will assist the Audubon Core Maintenance Group and the TDWG community in the review process during ratification of the controlled vocabularies.

Development of the controlled vocabularies

After acceptance of the Task Group charter, the group began to hold regular meetings. The work carried out at those meetings was documented in a series of meeting notes*[7].

The initial task of the group was to solicit use cases from the community. Responses were received from image producers and consumers (including biodiversity researchers) and representatives of image aggregators. Contributors followed a template that was provided and the submitted use cases were organized by categories*[8]. The assembled use cases were discussed, then used by the Task Group to generate a candidate list of requirements for the controlled vocabularies*[9]. In this initial screening, some of the desired outcomes identified from the use cases were determined to be out of scope or not practical to satisfy within the time frame of the Task Group's work. This candidate list of requirements comprised a first version of the Feature Report described in Section 4.2.1 of the VMS.

During the course of the Task Group's work, it became apparent that subjectPart designations could apply to only part of an image if the image contained multiple parts or even multiple organisms. Standardizing the description of regions of interest (ROIs) was not in the scope of this Task Group, but the eventual addition of an ROI vocabulary to

Audubon Core *[10]facilitated the Task Group's work by making it possible to apply subjectPart and subjectOrientation terms to specific ROIs within an image rather than only to the image as a whole.

Following the generation of candidate requirements, the Task Group reviewed existing approaches to describing subject parts and orientations, and identified ontologies that could be used to define parts and orientations unambiguously. Then work began on the actual construction of the two controlled vocabularies in the form of Simple Knowledge Organization System (SKOS) concept schemes (Isaac and Summers 2009), one describing parts and the other describing orientations.
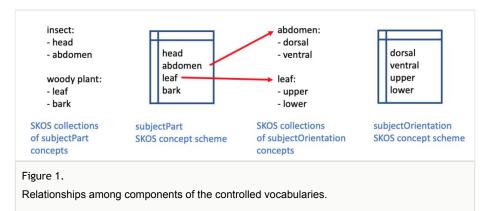
The concept schemes were relatively unstructured, except for a few cases where concepts had a broader relationship to other concepts (for example, "right" orientation had the broader concept "lateral" orientation; Table 1). One of the challenges in developing the concept schemes was meeting two of the candidate requirements involving categorization of the concepts: 1.1 "Subject part values are grouped appropriately for broad categories of organisms (e.g., trees, quadrupeds)." and 5.2 "Subject orientations are grouped appropriately for subject parts." Categorization was established by creating SKOS collections (Isaac and Summers 2009), where one set of collections indicated which subject part concepts were suitable for various organism groups, and another set of collections indicated which orientations were suitable for various subject parts (Fig. 1).

Table 1.

Example of metadata for a subjectOrientation value.

Term Name acorient:r0004	
Term IRI	http://rs.tdwg.org/acorient/values/r0004
Modified	2022-01-01
Term version IRI	http://rs.tdwg.org/acorient/values/version/r0004-2022-01-01
Label	right side
Definition	view of the right side of a whole bilaterally symmetric organism
Definition derived from:	http://purl.obolibrary.org/obo/BSPO_0000007
Controlled value	right
Has broader concept	acorient:r0003
Type	Concept

During the construction of the controlled vocabularies, it became apparent that although satisfying some of the candidate requirements would be desirable, doing so would not be practical. In some cases, meeting a requirement would have made the vocabularies too complex for most users. In other cases, it would have made some of the concepts so granular that few users would use them and the size of the vocabularies would be impractically large. As a result, a large number of the candidate requirements were dropped prior to the creation of the final list of requirements (i.e., the final version of the

VMS Feature Report). There were two cases where the candidate requirements from the Task Group helped drive other developments within Audubon Core:

- The term sawsdlrdf:modelReference was borrowed from the W3C Semantic Annotations for WSDL and XML Schema Recommendation*[11] as a means to link terms with external definitions. In the context of the Views Task Group, this term was used to link concepts to terms in the external ontologies. (Candidate requirement 6-ANATOMY-1)
- The creation of the ac:RegionsOfInterest class and the associated terms used to describe its instances made it possible to apply subjectPart and subjectOrientation properties to parts of images in cases where the image depicted multiple organisms or multiple parts. (Candidate requirement 2-FILTER-1)

Figure 1.

Relationships among components of the controlled vocabularies.

The final requirement list is presented in Appendix A. Of the seventeen potential requirements derived from the submitted use cases, seven were included in the final requirements.

During the development process, ac:subjectPart and ac:subjectOrientation were added as custom metadata fields in Zenodo. Controlled string values appropriate for insects from the preliminary vocabularies were used to categorize images of fly specimens when they were submitted. Fig. 2 shows an example specimen*[12] image*[13] that includes these fields. Including values for these fields makes it possible to search images to retrieve all images of a particular subjectPart (e.g., https://zenodo.org/search?page=1&size=20&custom=%5Bac:subjectPart%5D:%5Bhead%5D) or a particular subject part in a particular orientation (e.g., https://zenodo.org/search?page=1&size=20&custom=%5Bac:subjectPart%5D:%5Bhead%5D&custom=%5Bac:subjectOrientation%5D:%5Bventral%5D). At that time, IRIs had not yet been assigned to candidate terms nor had the "Literal" term analogs been adopted. So this was a preliminary test that did not include all of the features of the final submitted draft vocabularies. Nevertheless, this demonstrated the practicality of using the controlled vocabulary terms to sort out images having a particular view from a larger set of specimen images.
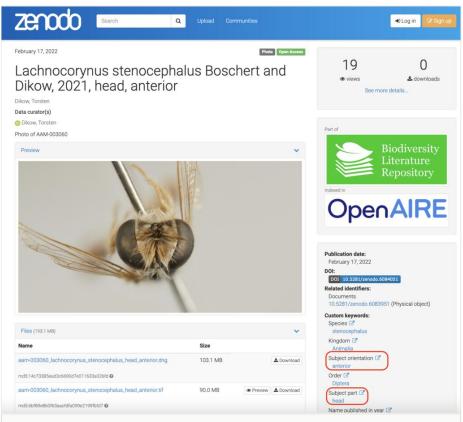
Figure 2.

Specimen image https://zenodo.org/record/6084051 showing an anterior view of the head of a fly, *Lachnocorynus stenocephalus*. Image used under a CC BY license. Dikow, Torsten. (2022). *Lachnocorynus stenocephalus* Boschert and Dikow, 2021, head, anterior. Zenodo. https://doi.org/10.5281/zenodo.6084051

Implementation testing

After the completion of the draft controlled vocabularies, the Task Group began planning for implementation testing by identifying potential testers and creating an implementation testing guide (Suppl. material 1). The group invited participation from researchers specialized in diverse groups of organisms that included non-vascular plants and non-arthropod invertebrate animals in addition to the vascular plants and arthropods that were the primary focus of the vocabulary development, with the intention of judging how extensible the terms were.

The group recognized that not every tester would use the vocabularies in the same way, so the guide described testing by manual entry, machine-guided entry, and machine processing. The guide also included the questions that would be asked on the feedback

form at the conclusion of testing so that the participants would have a better idea of what to consider while carrying out the testing. The vocabularies were also made available to the testers as CSV tables*[14], [15], text documents*[16], [17], and machine-readable JSON-LD files*[18], [19].

After potential testers were identified, the Task Group held an optional workshop that would give the testers an opportunity to try applying the terms to a small number of images while they could ask clarifying questions to the workshop presenters. Testers then carried out testing on a larger sample of images.

After completing the testing, implementers submitted feedback using a Google form (Suppl. material 2). The Task Group reviewed the feedback and in some cases considered modifications to the vocabularies to address the concerns raised by the test implementers. When the revisions were complete, the labels and definitions were translated into Spanish.

Implementation test results

There were five institutions participating in the implementation testing (Table 2).

Table 2.

Implementation testing participants.

Organization	Taxonomic coverage	Number of images in test	Image type	Testing type
Field Museum (Field)	plants	33	live organisms, digitized specimens	manual entry
Dept. of Ecology and Evolutionary Biology, University of Kansas (Kansas)	marine invertebrates	14	live organisms	manual entry
Bioimages (Bioimages)	seed plants	25	live organisms	manual entry, machine processing
UC Santa Barbara, Cheadle Center for Biodiversity and Ecological Restoration (California)	Anthophila	21	digitized specimens	manual entry
Universidad de San Carlos de Guatemala -CUNZAC- (Guatemala)	plants	8	live organisms	manual entry

Manual entry

The detailed testing results are presented in Appendix B. All testers carried out testing manually by having a human refer to the human-readable lists of concepts or CSV tables, then entering controlled value strings into a spreadsheet. All users selected orientation concepts from a collection appropriate for a particular part. Some users also used the

collections of parts appropriate for organism groups, although this wasn't generally necessary for groups whose images were only part of one group.

The testers applied the controlled vocabularies to a variety of types of organisms with several kinds of photographing circumstances: preserved specimens with fixed orientations and live organisms with uncontrolled and controlled orientations (Table 2).

Generally, the problems that users encountered had less to do with the vocabularies themselves, but more with difficulties caused when an image was not restricted to a single organism, part, or orientation. In theory, problems related to inclusion of multiple organisms or parts could be addressed by defining regions of interest within the image and then applying the terms to those specific regions, but without machine assistance to demarcate those regions, record their bounds, and associate the values with those regions, that solution wasn't practical for a human assessing an entire image. This assistance could range from a simple manual "click and drag" tool for demarcating the regions to a fully automated system for detecting parts and orientations. Two of the testers noted that sometimes both upper side and lower side orientations are purposefully included in the same specimen or image, making this problem a frequent occurrence for photographs that include plant leaves. The problem of uncontrolled orientation was noted by two testers who photographed live organisms. Due to difficulties of photographing organisms that were not in a fixed position, or organisms whose parts pointed in multiple directions, the chosen orientation represented a "best estimate". One tester noted that users might have trouble selecting an appropriate part if they lacked the technical expertise to do so.

Machine processing

None of the testers used machine-guided selection. That is not surprising since it would require development of new software or customization of existing software to make use of the machine-readable SKOS. However that is likely to be an important use case in the future after adoption of the vocabularies.

One implementer used machine processing to assign values based on existing text descriptions of the view. The process and results were described as follows: "For the automated conversion, I queried the Bioimages image dataset to create a spreadsheet of the view descriptor IRIs we use along with the associated part and 'view' labels. Each IRI was mapped in a table to an ac:subjectPartLiteral and ac:subjectOrientationLiteral value appropriate for that IRI. In some cases there wasn't a specific subject orientation, so I used 'unspecifiedOrientation'. See https://github.com/baskaufs/msc/blob/master/bioimages_views/stdviews_table.csv for the mappings. I then queried the database for the 25 images that were used for the human test and used the mapping table to assign AC controlled value strings for each of the images based on its descriptor IRI value. The values derived by a human were compared to those generated by the automated mapping. In all images, the subjectParts corresponded completely. Where a specific subject orientation could be assigned via mapping, that orientation agreed with the human assessment, but there were many cases where the mapping generated an 'unspecifiedOrientation' value when the human was able to make an assignment. This is just a limitation of our existing system to

capture complete information about the orientation. See https://github.com/baskaufs/msc/blob/master/bioimages_views/mapping_ test.csv for the results. The Python script to do the querying and mapping is at https://github.com/baskaufs/msc/blob/master/bioimages_views/bioimages_views.ipynb ."

This test was probably more successful than would be likely for a random provider since the Bioimages images were organized using the same Morphbank view categories that influenced the construction of the controlled vocabularies. Nevertheless, it showed that automated conversion was possible, although an "unspecifiedOrientation" assignment is likely to be the result when existing view descriptors cannot be perfectly mapped to the controlled vocabularies.

Unimplemented requirements

There were three requirements (Appendix A) that testers did not implement:

1.4 Specify multiple parts in an image by applying subjectPart concepts to Regions of Interest within an image. (2-FILTER-1)

1.5 Distinguish between single and aggregate parts (e.g., one vs. several leaves) by applying multiple subjectPart concepts of the same type to Regions of Interest within a single image. (7-CLARITY-2)

2.2 For some organism groups, filter orientations so that selection is only possible if the feature is visible for a particular subject part. (8-ORIENT-1)

Requirements 1.4 and 1.5 depend on the implementation of Regions of Interest, which is not a feature of these vocabularies on their own, but rather a separate technology that none of the implementers had (yet) implemented. So the inability to implement was not a deficiency of the vocabularies per se.

Requirement 2.2 depends on software development that was beyond the scope of what was expected of the testers. So again, the inability to implement during testing does not indicate a deficiency of the vocabularies themselves. Although the structural design of grouping concepts for filtering purposes was not used in a machine-assisted way (i.e., through software consuming the vocabularies as JSON-LD or tabular data), the division into SKOS collections was used to generate the human-readable documents to which most implementers referred when selecting concepts.

Use of vocabularies beyond the original organism groups

The testing included some organism groups that were not included in the original set: cnidarians and bryophytes.

The testing suggested that the existing terms were not adequate when applied to cnidarians. Although "lateral side" was an appropriate description, "dorsal side" and

"ventral side" did not make sense when applied to non-bilaterally symmetric animals. Therefore, two additional orientations: "oral side" and "aboral side" were added to make the vocabularies useable with groups like cnidarians and echinoderms.

The existing terms developed for seed plants were not adequate for all bryophyte images. During the testing period, additional terms were proposed for bryophytes, ferns, and fungi. Since there was insufficient time for these new terms to be tested in order for them to be included in the initial submission, they were designated as candidate terms to be considered as future additions to the vocabularies*[20].

One result of the discussion about using the vocabularies with these other groups was changing the subjectOrientation terms for "adaxial side" and "abaxial side" to "upper side" and "lower side". That made them more broadly usable in organism groups that did not have a clearly defined central axis. It also made the terms easier for human users to select without error, since the strings "adaxial" and "abaxial" are more difficult to distinguish due to their visual similarity.

Conclusions

Controlled vocabularies are necessary for standardizing the way we manage and process information. By developing these views vocabularies, we provide a means to describe the content of images by manual entry, machine-guided entry, and machine processing. Test implementers were able to use the terms proposed without major issues. Although the controlled vocabularies for subjectPart and subjectOrientation were not able to handle every scenario imagined in the original request for use cases, they fulfilled most of the final requirements established by the Task Group (the "Feature Report" required by the SDS).

The usefulness of these controlled vocabularies would be increased if they were incorporated within tools for manually demarcating or automatically detecting regions of interest that correspond to subjectParts depicted in an image. We hope that developers will take advantage of this opportunity to associate parts of images with machine-readable metadata for describing what is depicted in those parts.

Moving forward, when the vocabularies are adopted and broadly used, we expect that these will expand over time to include subjectParts that were not tested in this implementation. Through the TDWG term change process (Baskauf et al. 2017a), we expect suggestions from a diversity of users spanning across biological groups that have not been tested yet. Those additions should be submitted for testing, but we also acknowledge that finding test implementers can represent a challenge.

One reason for using SKOS to describe the controlled vocabularies is that it provides a mechanism for enriching them to make them more broadly usable. As we noted, SKOS collections have been used to group the concept terms in meaningful ways and skos:prefLabel has been used to add labels and definitions in Spanish. In the future, labels and definitions in other languages may be added as translations are completed and skos:altLabel may be used to document alternative labels for which users may search.

Overall, the results of implementation testing demonstrated that the vocabularies are ready for adoption and inclusion as part of the Audubon Core standard.

Appendix A. Final requirements

Note: source use cases*[8] follow a requirement (in parentheses).

Subject part

1 Categorization

1.1 Subject part values are grouped appropriately for broad categories of organisms (e.g., woody angiosperms, insects). Selecting a SKOS Collection will allow a user to find a group of part concepts appropriate for a particular category of organisms. (1-CATEGORIZE-1)

1.2 Concepts are linked to well-known ontologies to clarify definitions and standardize labels. However, the actual concepts are TDWG-adopted terms, providing stability that might not exist in the source ontologies. (6-ANATOMY-1) Ontologies used were the Biological Spatial Ontology (BSPO)*[21], Phenotype and Trait Ontology (PATO)*[22], Common Anatomy Reference Ontology (CARO)*[23], Plant Ontology (PO)*[24], Uber-anatomy Ontology (UBERON)*[25], Drosophila gross anatomy Ontology (FBBT)*[26], and Ontology for the Anatomy of the Insect SkeletoMuscular system (AISM)*[27].

1.3 Concepts allow for distinguishing between sexes (if multimorphic) by selecting narrower categories of subjectPart. (added during discussion)

1.4 Specify multiple parts in an image by applying subjectPart concepts to Regions of Interest within an image. (2-FILTER-1)

1.5 Distinguish between single and aggregate parts (e.g. one vs. several leaves) by applying multiple subjectPart concepts of the same type to Regions of Interest within a single image. (7-CLARITY-2)

2 Relationship between part and orientation

2.1 Determine what orientations are appropriate for subject parts other than whole organism. (3-MEASURE-4, 1-CATEGORIZE-2)

2.2 For some organism groups, filter orientations so that selection is only possible if the feature is visible for a particular subject part. (8-ORIENT-1)

Appendix B. Test implementation notes

Results from specific testers to questions on the feedback form are included in Table 3.

Table 3.

Specific responses from test implementers.

Organiz.	Testing details	Difficulties in selecting concepts
Field	We used photos of live plants as well as herbarium specimens for this exercise. We had 3 people (2 data, 1 botanist) view a set of 33 images and independently assign values for subjectPartLiteral and subjectOrientationLiteral. We then compared the results and recorded any questions that arose.	This happened frequently. Most of our images were of whole organisms, without clearly defined regions of interest. In the case of herbarium specimens, for example, it is oftentimes the goal to articulate the plant in a way that shows both the adaxial and abaxial sides of the leaf. There were many times when a part or orientation of a plant was present in the image but may have not been the focus of the image or maybe was only partially visible so we were unsure if those parts/orientations should be recorded. Also there were times when clarity of the image makes it difficult to identify a part or orientation. It was sometimes hard for the "data" people in the team to identify SubjectPart and needed botanist's expertise. **Missing concepts**: For subjectPartLiteral: Suggest adding "petiole" and "stipules" (for plant's leaf); maybe useful to include the maturity of reproductive parts e.g., "bud" and "flower"; "trunk" should be added since sometimes people specifically take a photo of trunk to show certain characteristics; the term "bark" may not be useful if we have "twig" and "trunk"; technically, when inflorescence becomes a fruit, it is called "infructescence"; also the term "peduncle" came up - the botanist indicated it as a helpful term to include. For subjectOrientationLiteral, maybe add "transverse" and "longitudinal" for fruits. Often times, people take a photo of cut fruit.
Kansas	The types of images I chose were all live images, mainly from a marine field station and also images through a microscope from my lab, which works with live cnidarians. The field station images included a mix of specimens that were in the field (i.e. jellyfish in the water), in a glass jar or container, or specimens through a microscope. I am trained as a marine invertebrate and evolutionary biologist, so the types of images I would be taking would be for documenting species that we may bring back to the lab for various types of assays or sequencing experiments. For sequencing in particular, it is useful to have images of the specimens used, but getting "clean" images with easily definable	For radially symmetric animals or animals with multiple individuals (e.g. colonial animals, groups of animals), it was tricky to determine the most appropriate concepts. For multiple individuals I selected a focal "individual" centered on the image, but of course this is subjective to each viewer. **Missing concepts**: Many of the animals that I work with are radially symmetric (or similar - e.g pentameral symmetry of sea stars). While dorsal, ventral, and lateral were sufficient for most tasks, for cnidarians (and ctenophophores) their symmetry is described as oral-aboral, which would be more appropriate concepts for orientation.

Kansas	features or single organisms can be difficult. In selecting images, I wanted to get a range of body types and phyla, both ones that were easily identifiable in terms of orientation as well as unconventional images from the field that may be trickier to categorize. This is why I went over the 5-10 image range for manual testing.	**Insufficient granularity**: Colonial animals could perhaps require a specific identification, or at least a way to distinguish from other organism types, since "entire Organism" could mean entire colony or entire individual of the colony. Cnidarians or non-bilateral animals in general may also require additional terms.
Bioimages	I applied the controlled values to a broad range of plant parts across woody angiosperms, herbaceous angiosperms, and gymnosperms. I examined the image and selected appropriate values from the spreadsheet of available values. In some cases, I referred to the lists of orientations appropriate for parts to make sure that the orientation I was selecting was appropriate for the part. I copied the selected controlled value from the spreadsheet of values to the spreadsheet where I was recording the test results, along with the image filename and GUID (IRI) for the image. Since I was working alone, I didn't have anyone to check my work. However, I did crosscheck using the automated mapping.	Photos of leaves intentionally included both adaxial and abaxial sides to show the difference in surface characteristics. I chose the most prominent side, but when leaf margins were photographed, there wasn't really a predominant side. See http://bioimages.vanderbilt.edu/baskauf/41902 for an example. The orientation for inflorescences was sometimes difficult to determine when multiple infloreses were visible (e.g. clusters of catkins). In photos of dehiscing fruit, there wasn't really an appropriate view -- what would be the lateral side of the fruit was on the outside and not visible when the fruit interior was photographed. See http://bioimages.vanderbilt.edu/baskauf/24261 for an example. It was difficult to determine whether images of juvenile herbaceous plants were apical or lateral because they were usually photographed at an angle. The orientation of clusters of male cones was a similar situation to catkins. Usually the images did not include just a single cone and the cones were sticking out at differing orientations. In case where leaves were not laminate (for example pine needles and rounded leaves (for example sedum: http://bioimages.vanderbilt.edu/baskauf/21935), it wasn't clear which side of the leaf was adaxial or abaxial. For plants where the leaves emerged vertically in whorls (e.g. yucca http://bioimages.vanderbilt.edu/baskauf/14597) it is difficult to photograph one side of a single leaf.For some photos of a part of a plant part (trunk of a whole tree, internal parts of a flower), it was difficult to specify the orientation because the whole part was not visible and the image was taken at an angle to the feature that wasn't apical or lateral. In some cases where the inflorescence consisted of a single flower, it wasn't clear whether the part should be "flower" or "inflorescence". See for example http://bioimages.vanderbilt.edu/baskauf/50597 (For specific details, see Table 4.)

| California | The 1,243,540 are categorized in a general way to facilitate searching both by orientation and part. https://library.big-bee.net/portal/imagelib/search.php | It was generally easy, but difficult when labels are present or the only part of a specimen imaged. **Insufficient granularity**: Hymenoptera technically do not have a thorax and abdomen, these are actually mesosoma and metasoma due to the place of constriction of the waist. This is pretty technical difference and thorax/abdomen can be considered general terms for before and after the constriction. (For specific details, see Table 5.) |
| Guatemala | I tried to contribute with information for bryophytes, a group of plant that wasn't included at the first time. | There is still some work needed to complete parts and orientations for bryophytes, however, the data proposed seem to be functional for this group of plants. |

Specific comments about test images from the Bioimages test are included in Table 4.

Table 4.

Detailed results from Bioimages testing. The image_identifier is appended to a base IRI of http://bioimages.vanderbilt.edu/.

image_identifier	ac:subjectPartLiteral	ac:subjectOrien-tationLiteral no	notes
baskauf/25638	entireOrganism	lateral	
baskauf/12625	entireOrganism	lateral	The part is more like "trunk" than whole organism.
baskauf/41910	bark	lateral	
baskauf/63779	twig	lateral	
baskauf/41905	leaf	adaxial	Hard to remember that "adaxial" is the upper leaf surface.
baskauf/41902	leaf	adaxial	About half of the image is adaxial of one leaf and abaxial of another.
baskauf/41887	leaf	adaxial	Image shows many leaves, all of the adaxial side, but probably needs to use ROIs.
baskauf/42310	inflorescence	lateral	Orientation not clear as there are various inflorescences sticking out.
baskauf/50597	inflorescence	lateral	Not sure if I should call this a flower or inflorescence
baskauf/50743	flower	apical	
baskauf/50741	flower	apical	Just part of the flower
baskauf/41891	fruit	lateral	
baskauf/24261	fruit	lateral	Not really lateral, the fruit is dehiscing

baskauf/65236	stem	lateral	
baskauf/57859	entireOrganism	apical	Several plants, some more lateral than apical
baskauf/27473	flower	apical	
baskauf/61716	leaf	abaxial	not sure how you can tell the orientation for needles
baskauf/51363	femaleCone	lateral	
baskauf/51365	maleCone	lateral	several cones, a variety of orientations
baskauf/33496	leaf	adaxial	
baskauf/33504	inflorescence	lateral	
thomas/ 0627-01-01	entireOrganism	lateral	
baskauf/21935	leaf	abaxial	difficult to say if this is leaf or stem and the leaves are not really laminate, so orientation is difficult
baskauf/14597	leaf	adaxial	because of the whorled nature of this, both abaxial and adaxial orientations are present

Specific comments about test images from the California test are included in Table 5.

Table 5.

Detailed results from the California (Big Bee) project.

image	ac:subject PartLiteral	ac:subject Orientation Literal	Notes	Image description
https://mczbase.mcz.harvard.edu/ specimen_images/entomology/large/MCZ-ENT00610195_Agapostemon_texanus_hwg.jpg	hindwing	dorsal		Bee images from the Big-Bee project: https://library.big-bee.net/portal/taxa/index.php?taxon=3667
https://serv.biokic.asu.edu/ imglib/ecdysis/UCSB_IZC/UCSB-IZC00036/ UCSB-IZC00036938_1633713553.jpg	entire organism	dorsal	no way to recognize label which is part of a specimen	Bee images from the Big-Bee project: https://library.big-bee.net/portal/taxa/index.php?taxon=3667

image	ac:subject PartLiteral	ac:subject Orientation Literal	Notes	Image description
https://serv.biokic.asu.edu/ imglib/ecdysis/UCSB_IZC/UCSB-IZC00028/ UCSB- IZC00028367_3d_2020-08-07-16.04.15_lg.jpg	entire organism	anterior	As part of a series of 2D images that go around a specimen it is not a specific orientation but really between orientations.	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667
https://monarch.calacademy.org/ mnt/target-images/CASTYPE/00001/ CASTYPE1503_h.jpg	head	anterior		Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667
https://ids.si.edu/ids/ deliveryService/id/ark:/65665/ m3642ca13b63774a5c9f907938471a d74f/1200	entire organism	lateral	Also includes a good image of the wing	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667
https://monarch.calacademy.org/ mnt/target-images/CASTYPE/00001/ CASTYPE1506_label.jpg	N/A	N/A	Nothing applies.	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667
https://monarch.calacademy.org/ mnt/target-images/CASTYPE/00001/ CASTYPE1506_l.jpg	entire organism	lateral	Also includes a good image of the wing	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667

image	ac:subject PartLiteral	ac:subject Orientation Literal	Notes	Image description
https://serv.biokic.asu.edu/ imglib/ecdysis/UCSB_IZC/UCSB-IZC00028/ UCSB-IZC00028367_3d_2020-08-07-18.02.18.jpg	entire organism	ventral	As part of a series of 2D images that go around a specimen it is not a specific orientation but really between orientations.	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=3667
https://symbiota.ccber.ucsb.edu/ content/specimenImages/UCSB_IZC/UCSB-IZC00036/ UCSB-IZC000364411xlateral1-edi_ 1583187181_lg.jpg	entire organism	lateral	Also includes a good image of the wing	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=6324
https://symbiota.ccber.ucsb.edu/ content/specimenImages/UCSB_IZC/UCSB-IZC00009/ UCSB-IZC00009308_lg.jpg	entire organism	dorsal	no way to recognize label which is part of a specimen	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=6324
https://serv.biokic.asu.edu/ imglib/ecdysis/UCSB_IZC/UCSB-IZC00042/ UCSB-IZC00042555_3d_2021-08-02-13.44.48.jpg	entire organism	lateral	As part of a series of 2D images that go around a specimen it is not a specific orientation but really between orientations.	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=6324
https://serv.biokic.asu.edu/ imglib/ecdysis/UCSB_IZC/UCSB-IZC00010/ UCSB-IZC00010327_3d_Folder_004_2021-07-22-20.20.33_lg.jpg	abdomen	dorsal	As part of a series of 2D images that go around a specimen it is not a specific orientation but really between orientations.	Bee images from the Big-Bee project: htt ps://library.big-bee.net/portal/ taxa/ index.php? taxon=6324

image	ac:subject PartLiteral	ac:subject Orientation Literal	Notes	Image description
https://ids.si.edu/ids/deliveryService/id/ark:/65665/m3be3a15601b514314a5737aa195fb9d36/1200	entire organism	dorsal	Two specimens in single image	Bee images from the Big-Bee project: https://library.big-bee.net/portal/collections/individual/index.php?occid=1673075
https://ids.si.edu/ids/deliveryService/id/ark:/65665/m377be30404e8546a988597bf6689bebc0/1200	N/A	N/A	Nothing applies.	Bee images from the Big-Bee project: https://library.big-bee.net/portal/collections/individual/index.php?occid=1673075
https://mczbase.mcz.harvard.edu/specimen_images/entomology/large/MCZ-ENT00017211_Melissa_porteri_fwg.jpg	forewing	unspecifiedOrientation		Bee images from Big-Bee project: https://library.big-bee.net/portal/taxa/index.php?taxon=17374
https://mczbase.mcz.harvard.edu/specimen_images/entomology/large/MCZ-ENT00017211_Melissa_porteri_hav.jpg	entire organism	ventral		Bee images from the Big-Bee project
https://mczbase.mcz.harvard.edu/specimen_images/entomology/paleo/large/PALE-7514_Apis_henshawi.jpg	entire organism	ventral		Bee images from the Big-Bee project
https://mczbase.mcz.harvard.edu/specimen_images/entomology/large/MCZ-ENT00610536_Lasioglossum_oblongum_thl.jpg	thorax	lateral	In Hymenoptera, this is called a mesosoma because one segment of the abdomen is part of the structure.	Bee images from the Big-Bee project
https://mczbase.mcz.harvard.edu/specimen_images/entomology/large/MCZ-ENT00000547_Halictus_albitarsis_hlg.jpg	leg	unspecifiedOrientation		Bee images from the Big-Bee project
https://mczbase.mcz.harvard.edu/specimen_images/entomology/large/MCZ-ENT00015720_Megachile_pseudoexilis_hlg.jpg	leg	unspecifiedOrientation		Bee images from the Big-Bee project

image	ac:subject PartLiteral	ac:subject Orientation Literal	Notes	Image description
https://serv.biokic.asu.edu/imglib/ecdysis/UCSB_IZC/UCSB-IZC00010/UCSB-IZC00010221.jpg	entire organism	dorsal	no way to recognize label which is part of a specimen	Bee images from the Big-Bee project

Acknowledgements

Author contributions

SJB, JCGD, and MN wrote the manuscript and were core task group members. NSC and RS were core task group members. KCS, ZK, MP, DA, and AMLK were implementation testers. DA also participated in the task group during the development phase.

References

- Baskauf S, Kirchoff B (2008) Digital plant images as specimens: toward standards for photographing living plants. Vulpina 7: 16-30. [In English]. URL: http://www.cals.ncsu.edu/plantbiology/ncsc/vulpia/pdf/Baskauf_&_Kirchoff_Digital_Plant_Images.pdf
- Baskauf S, Wieczorek J, Blum S, Morris RA, Rees J, Sachs J, Whitbread G (2017a) Vocabulary Maintenance Specification. Biodiversity Information Standards (TDWG). Review manager: Dag Endresen. URL: http://rs.tdwg.org/vms/doc/specification/2017-04-25
- Baskauf S, Hyam R, Blum S, Morris RA, Rees J, Sachs J, Whitbread G, Wieczorek J (2017b) Standards Documentation Specification. Biodiversity Information Standards (TDWG). Review manager: Dag Endresen. URL: http://rs.tdwg.org/sds/doc/specification/2017-04-25
- Isaac A, Summers E (Eds) (2009) SKOS Simple Knowledge Organization System Primer. World Wide Web Consortium (W3C). W3C Working Group Note 18 August 2009. URL: https://www.w3.org/TR/skos-primer/

- Morris RA, Barve V, Carausu M, Chavan V, Cuadra J, Freeland C, Hagedorn G, Leary P, Mozzherin D, Olson A, Riccardi G, Teage I (2013a) Audubon Core Introduction. Biodiversity Information Standards (TDWG). Review manager: Steve Baskauf. URL: http://rs.tdwg.org/ac/doc/introduction/
- Morris RA, Barve V, Carausu M, Chavan V, Cuadra J, Freeland C, Hagedorn G, Leary P, Mozzherin D, Olson A, Riccardi G, Teage I, Whitbread G (2013b) Discovery and publishing of primary biodiversity data associated with multimedia resources: The Audubon Core strategies and approaches. Biodiversity Informatics 8 (2). [In English]. https://doi.org/10.17161/bi.v8i2.4117

Supplementary materials

Suppl. material 1: Views Controlled Vocabularies testing notes `doi`

Authors: Steven J. Baskauf and Jennifer C. Girón Duque
Data type: text document
Brief description: These notes were provided to test implementers as a guide to carrying out the testing.
Download file (2.43 MB)

Suppl. material 2: Views Controlled Vocabularies Implementation Reporting Form `doi`

Authors: Steven J. Baskauf, Jennifer C. Girón Duque, and Matthew Nielsen
Data type: text document
Brief description: This document is an export of the questions to which implementers responded after testing.
Download file (63.56 kb)

Endnotes

[*1] Morphbank :: Biological Imaging (https://www.morphbank.net/, 15 July 2022). Florida State University, Department of Scientific Computing, Tallahassee, FL 32306-4026 USA.

[*2] https://www.morphbank.net/About/AboutMb/

[*3] https://www.morphbank.net/MyManager/?tab=viewTab

[*4] https://www.morphbank.net/Show/?id=470716

[*5] https://github.com/tdwg/ac/blob/9f173d69704afc8d9c11776377f8b94d758308e3/historical/views-tg-draft-charter-2019-06-25.pdf

[*6] Proposal to "Revise ac:subjectOrientation and ac:subjectPart and add ac:subjectOrientationLiteral and ac:subjectPartLiteral" https://github.com/tdwg/ac/issues/195

[*7] https://github.com/tdwg/ac/tree/28ceb904018ee54ea8d3a6f8f88e5e347e10a32c/views/historical

*8   https://github.com/tdwg/ac/blob/b0eb3f091557a86fe67cc207ff813493a4b3f4b8/views/
         submitted-use-cases.md

*9   https://github.com/tdwg/ac/blob/317bf0045f8ab865d3ebfa37151c36fd7045f7c5/views/
         candidate-requirements.md

*10 https://ac.tdwg.org/termlist/#711-region-of-interest-vocabulary

*11 https://www.w3.org/TR/sawsdl/#modelReference

*12 https://zenodo.org/record/6083951

*13 https://zenodo.org/record/6084051

*14 https://github.com/tdwg/ac/blob/e19fbfa12bf694ce36dacfb3371875093ccde421/views/
         code/subjectPart/subjectPart_cv.csv

*15 https://github.com/tdwg/ac/blob/e19fbfa12bf694ce36dacfb3371875093ccde421/views/
         code/subjectOrientation/subjectOrientation_cv.csv

*16 https://github.com/tdwg/ac/blob/022d5af69f4956a456ecd4380786024b71c28b4d/
         views/subjectPart.md

*17 https://github.com/tdwg/ac/blob/022d5af69f4956a456ecd4380786024b71c28b4d/
         views/subjectOrientation.md

*18 https://github.com/tdwg/rs.tdwg.org/blob/
         3618e0e8e71c4919da2d609486a3da233fa75529/cvJson/acpart.json

*19 https://github.com/tdwg/rs.tdwg.org/blob/
         3618e0e8e71c4919da2d609486a3da233fa75529/cvJson/acorient.json

*20 https://github.com/tdwg/ac/tree/b54222a821b3cd3451fe0c7667b5bb0e1705a4b6/
         views/candidate_terms

*21 BSPO, https://obofoundry.org/ontology/bspo.html

*22 PATO, https://obofoundry.org/ontology/pato.html

*23 CARO, https://obofoundry.org/ontology/caro.html

*24 PO, https://obofoundry.org/ontology/po.html

*25 UBERON, https://obofoundry.org/ontology/uberon.html

*26 FBBT, https://obofoundry.org/ontology/fbbt.html

*27 AISM, https://obofoundry.org/ontology/aism.html