

subjectPart and subjectOrientation Controlled Vocabularies User Guide

Date modified: 2023-01-10

Part of TDWG Standard: Not part of any standard

Abstract: This user guide provides information and examples showing how to use subjectPart and subjectOrientation Controlled Vocabularies and associated resources to describe the content of images depicting organisms.

Contributors: Steven J. Baskauf and Jennifer C. Girón Duque

Creator: TDWG Views Controlled Vocabularies Task Group

Bibliographic citation: Views Controlled Vocabularies Task Group. 2023. subjectPart and subjectOrientation Controlled Vocabularies User Guide. Biodiversity Information Standards (TDWG). https://github.com/tdwg/ac/blob/master/views/views_user_guide.pdf

Introduction

Biodiversity Information Standards (TDWG) is an international organization that develops standards for sharing biodiversity metadata. Audiovisual Core (formerly Audubon Core) is the TDWG standard for describing multimedia resources.

The Audiovisual Core properties *subjectPart* and *subjectOrientation* are used to categorize images for searching and display. Each possible value for these two terms is considered a *concept* and can be denoted either by a *controlled value string* or by an *internationalized resource identifier (IRI)*. The controlled value string resembles an English word or words (e.g. femaleCone) and is easy for humans to type. The IRI (e.g. <http://rs.tdwg.org/acpart/values/p0007>) is globally unique and can be used for machine processing. Because a controlled value string and an IRI represent the same concept, they are interchangeable.

The purpose of this guide is to provide guidance and examples for selecting and applying the controlled values to images and regions of interest (designated parts of images).

This guide was created by modifying the 2022-02-09 Views Controlled Vocabularies testing notes for test implementers of the vocabularies.

Reference resources

- subjectPart controlled vocabulary: <http://rs.tdwg.org/ac/doc/part/>
- subjectOrientation controlled vocabulary: <http://rs.tdwg.org/ac/doc/orient/>
- Categorized lists:
 - a. https://github.com/tdwg/ac/blob/master/views/code/orient_collections.md
 - b. https://github.com/tdwg/ac/blob/master/views/code/part_collections.md

Additional resources (CSV and JSON files) are linked in the text below.

Terminology

View

A "view" is a non-technical term for a description of the organism part and orientation of an image.

Broader concepts

Some narrow concepts are linked to broader concepts. For example, left and right subjectOrientations have the broader concept "lateral". Because of this relationship, it is possible to infer that a "right" orientation is also a "lateral" orientation. Generally, it is preferable to provide the narrowest concept possible unless it is not possible for the user to differentiate between the narrower categories and the user needs to select the broader category.

To discover the broader relationships

- look in the skos_broader column of the two controlled value spreadsheets <https://github.com/tdwg/rs.tdwg.org/blob/master/acpart/acpart.csv> and <https://github.com/tdwg/rs.tdwg.org/blob/master/acorient/acorient.csv>
- look in the "Has broader concept" term metadata field in the two term lists <http://rs.tdwg.org/ac/doc/part/> and <http://rs.tdwg.org/ac/doc/orient/>
- look for skos:broader values in the machine-readable JSON-LD at <https://tdwg.github.io/rs.tdwg.org/cvJson/acpart.json> and <https://tdwg.github.io/rs.tdwg.org/cvJson/acorient.json>

Regions of Interest

A region of interest (ROI) is a designated part of a media item. Each ROI in an image can be assigned its own values for subjectPart and subjectOrientation. Thus an image with several regions of interest may have multiple records for these two terms. Here is an example (region of interest outlined in yellow):

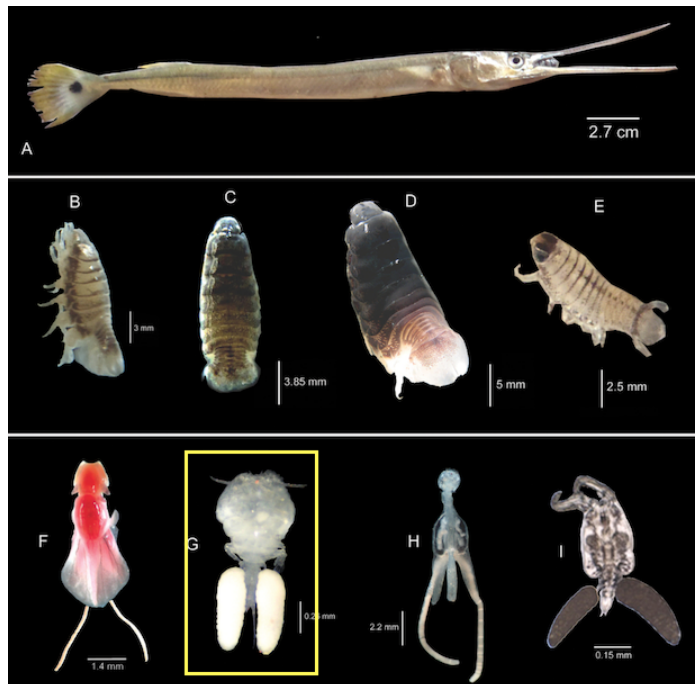


Image IRI: <https://zenodo.org/record/959321>

Image file name: big_34616.jpg

ROI IRI: <https://zenodo.org/record/959321#g>

subjectPartLiteral: entireOrganism

subjectOrientationLiteral: dorsal

Ways to apply values.

1. Manual entry.

The users examine an image and use the lists or spreadsheets of concepts to determine the controlled value appropriate for the image, then enter the value in a spreadsheet containing metadata about that image. Here are some examples:

Image of a live fish



The user refers to the spreadsheet of controlled values for subjectPart at <https://github.com/tdwg/rs.tdwg.org/blob/master/acpart/acpart.csv> and selects the appropriate controlled value string for an entire organism: "entireOrganism".

subjectPart_cv.csv

Liberation Sans 10 pt B / U A

A4:AMJ4 fx Σ = p0001

	A	B	J	
1	term_localName	label	controlled_value_string	type
2	p	subject part concept scheme		http://www.w3.org/2
3	p0000	unspecified part	unspecifiedPart	http://www.w3.org/2
4	p0001	entire organism	entireOrganism	http://www.w3.org/2
5	p0002	bark	bark	http://www.w3.org/2
6	p0003	twig	twig	http://www.w3.org/2
7	p0004	stem	stem	http://www.w3.org/2
8	p0005	leaf	leaf	http://www.w3.org/2
9	p0006	strobilis (cone)	strobilis	http://www.w3.org/2
10	p0007	inflorescence	inflorescence	http://www.w3.org/2
11	p0008	fruit	fruit	http://www.w3.org/2
12	p0009	seed	seed	http://www.w3.org/2
13	p0010	male cone	maleCone	http://www.w3.org/2
14	p0011	female cone	femaleCone	http://www.w3.org/2
15	p0012	flower	flower	http://www.w3.org/2
16	p0013	head	head	http://www.w3.org/2
17	p0014	thorax	thorax	http://www.w3.org/2

The user enters this value into the spreadsheet:

Liberation Sans 10 pt B / U A

B2 fx Σ = entireOrganism

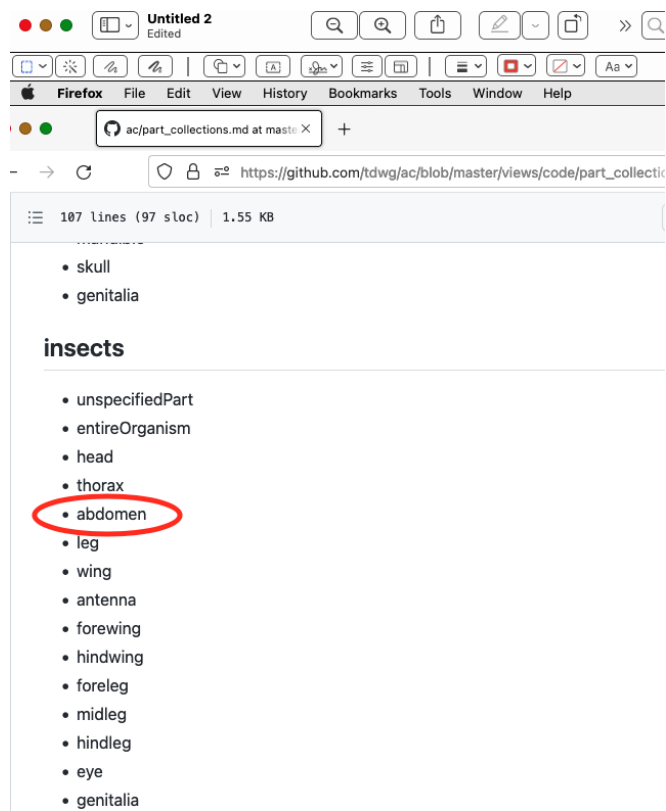
	A	B	C	D
1	image	ac:subjectPartLiteral	ac:subjectOrientationLiteral	
2	https://zenodo.org/record/907755	entireOrganism		
3				
4				

The user refers to the spreadsheet of controlled values for subjectOrientation at <https://github.com/tdwg/rs.tdwg.org/blob/master/acorient/acorient.csv> and selects the appropriate controlled value string for the right side of the organism: "right".

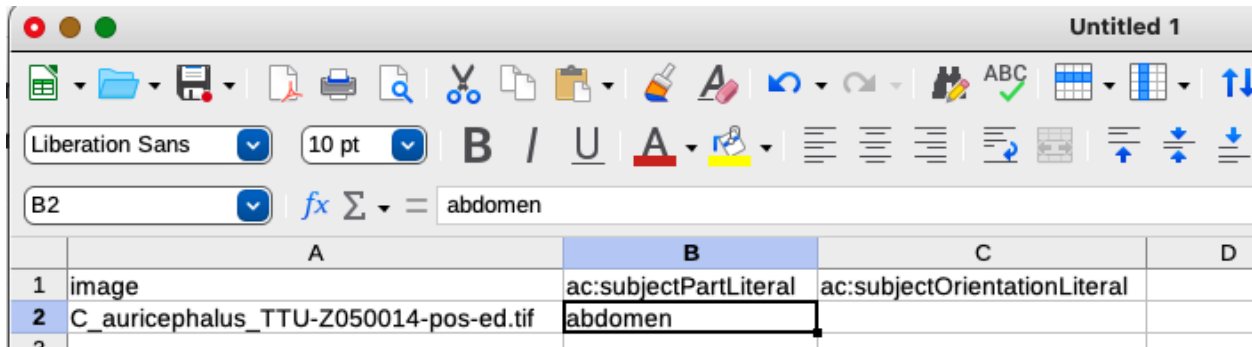
Image of an insect specimen



The user refers to the web page listing controlled value strings appropriate for different organism groups at https://github.com/tdwg/ac/blob/master/views/code/part_collections.md and finds the list for insects.

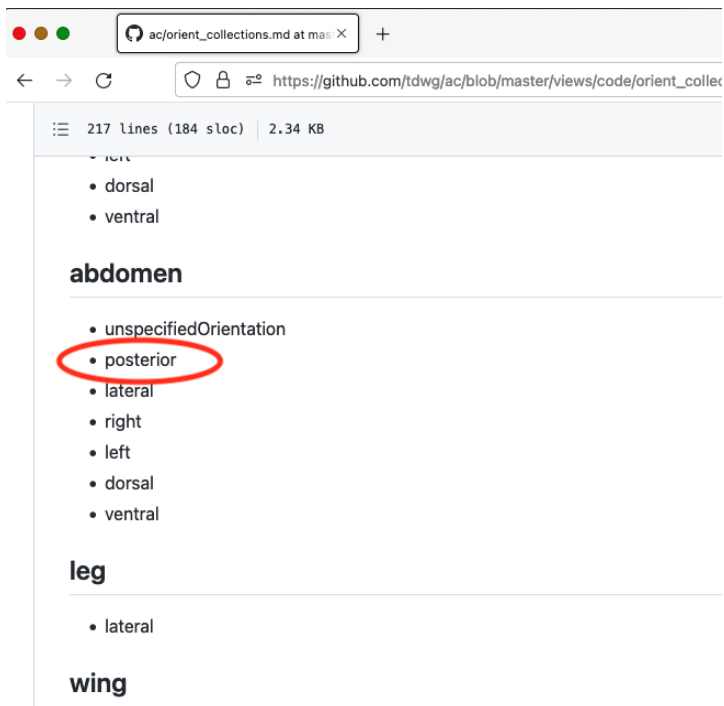


The user selects the value "abdomen" and enters it into a spreadsheet.

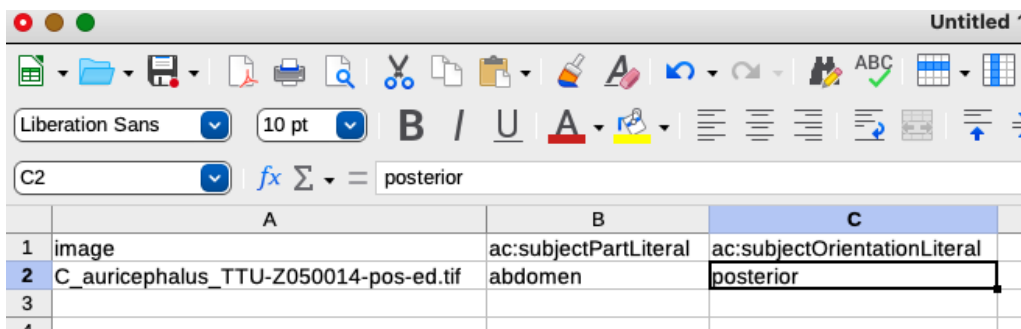


	A	B	C	D
1	image	ac:subjectPartLiteral	ac:subjectOrientationLiteral	
2	C_auricephalus_TTU-Z050014-pos-ed.tif	abdomen		

The user refers to the web page listing controlled value strings appropriate for different subject parts at https://github.com/tdwg/ac/blob/master/views/code/orient_collections.md and finds the list for abdomen:



The user selects the value "posterior" and enters it into a spreadsheet.



	A	B	C
1	image	ac:subjectPartLiteral	ac:subjectOrientationLiteral
2	C_auricephalus_TTU-Z050014-pos-ed.tif	abdomen	posterior
3			
4			

2. Machine-guided entry

The users examine images that are presented by a content management system (CMS). The CMS guides the user to select concepts that are appropriate for a particular organism group or subjectPart using dropdown menus or some other system. Here is an example:

Bioimages Collection Manager - Data Entry

File View Add Help

Group selected thumbnails as an individual organism

Organism ID: vanderbilt.edu/ind-baskauf/39190

Scheme Manually Set

Cameo Set: it.edu/baskauf/39192

Organism Remarks: Organism remarks

Org. Lat, Long | Alt: 36.43553, -87.30967 | Alt.

Org. Geo. Remarks: dent GPS measurement.

Organism Name: Friendly name

Organism Scope: multicellular organism

HTML Note: HTML note

Collection Code:

Catalog Number:

Establishment Means: native

Determinations: 1 of 1 < > New

Identified By: baskauf

Date Identified: 2005-04-10

Source of Name: th of Mexico (fna1993)

ID Remarks: Identification remarks

Common Name: bloody butcher

Kingdom: Plantae

Class: Liliopsida

Order: Liliales

Family: Liliaceae

Genus: Trillium

Species: recurvatum

Infraspecific Epithet:

Taxon Rank: species

tsnID: 43059

Specimen Group: herbaceous angiosperms

Specimen Part: inflorescence

Specimen View: unspecified

Image Caption: whole plant

Image Caption: stem

Occurrence remarks: leaf

Latitude, Longitude: fruit

36.43553 seed

Uncertainty (m): 30

Elevation (m): Elevation

Locality: slope above Cumberland R

Image georeference remarks:

Location inferred from organism coordinates.

Reason for withheld information:

How data was generalized:

Displaying all thumbnails

Advanced filter Sort

eral9-1f39185.jpg eral9-st39187.jpg eral9-1f39189.jpg

trre5-flyellow391... trre5-wpyellow3... trre5-flyellow-cl...

asbi6-wpin-flow... asbi6-wpin-flow... asbi6-flside3924...

Selected Filename Image Date Time TZ

e5-flyellow-close39194.jpg 2005-04-10 14:32:14 -05:00

Resolution: 1440 2160 Focal Length: 50

Photographer: baskauf

Image title: rvatum (Liliaceae) - inflorescence - closeup of flower interior

Image description: rvatum (Liliaceae) - inflorescence - closeup of flower interior

Geodetic Datum: EPSG:4326

Geonames Admin: 4642445

Geonames Other: Geonames other

Copyright Owner: baskauf

Copyright Year: 2005

Copyright Statement: (c) 2005 Steven J. Baskauf

Usage Terms: CC BY 3.0

Credit: Steven J. Baskauf http://bioimages.vanderbilt.edu/

URL to High Res: http://www.morphbank.net/?id=816543&imgType=jpeg

The user selected the organism group label "herbaceous angiosperms" and the subjectPart collections data were used to determine which subjectParts were appropriate for that organism group. The CMS used this selection to assign the value

ac:subjectPart = <http://rs.tdwg.org/acpart/values/p0007>

to assign the IRI for "inflorescence" to the property ac:subjectPart, whose values are intended to be IRIs. The user would then be presented with a picklist of subjectOrientations that are appropriate for inflorescences, based on the subjectOrientation collections data. In this case, the user would select "lateral" from the dropdown and the CMS would assign the value

ac:subjectOrientation = <http://rs.tdwg.org/acorient/values/r0003>

since the property ac:subjectOrientation is expected to have an IRI value.

The following JSON files may be useful for setting up the CMS to generate the dropdowns:

subjectPart collection providing parts appropriate for different organism groups:

https://tdwg.github.io/rs.tdwg.org/cvJson/acpart_collection.json

metadata about subjectPart concepts (labels, definitions, controlled value strings, broader relationships, ontology links): <https://tdwg.github.io/rs.tdwg.org/cvJson/acpart.json>

subjectOrientation collection providing orientations appropriate for different organism parts:

https://tdwg.github.io/rs.tdwg.org/cvJson/acorient_collection.json

metadata about subjectOrientation concepts (labels, definitions, controlled value strings, broader relationships, ontology links): <https://tdwg.github.io/rs.tdwg.org/cvJson/acorient.json>

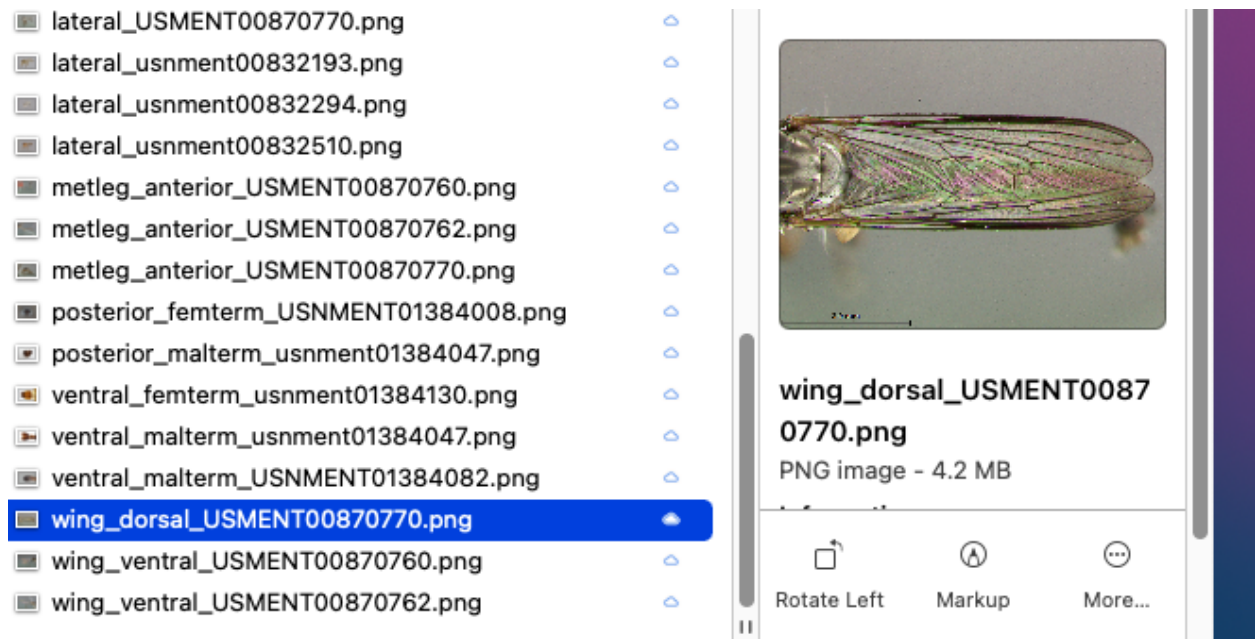
3. Machine processing

Data managers map the controlled value concepts to existing text-based schemes (e.g. text embedded in filenames). Values for subjectPart and subjectOrientation are generated for images using these mappings.

There is no specific algorithm for accomplishing this, but frequently existing identifiers or metadata contain sufficient information to assign at least some of the values automatically. In cases where values are ambiguous, the "unspecifiedPart" (acpart:p0000) and "unspecifiedOrientation" (acorient:r0000) values can be assigned, with later human-mediated processing to assign more specific values.

Example:

The filenames of these images embed information about the body parts:



The file name wing_dorsal_USMENT00870770.png can be split based on underscore positions:

part: "wing"

orientation: "dorsal"

rest of string: USMENT00870770.png

Mappings:

"wing" = <http://rs.tdwg.org/acpart/values/p0017>

"dorsal" = <http://rs.tdwg.org/acorient/values/r0006>

Metadata for image:

filename: wing_dorsal_USMENT00870770.png

ac:subjectPart: <http://rs.tdwg.org/acpart/values/p0017>

ac:subjectOrientation: <http://rs.tdwg.org/acorient/values/r0006>

For a more extensive example, see the "Machine processing" section of the Implementation Experience Report at <http://doi.org/10.3897/biss.7.94188> .