Syst. Biol. XX(XX):XX–XX, 2023

https://doi.org/10.1093/sysbio/syad025
Advance Access Publication April 24, 2023

Formalizing Invertebrate Morphological Data: A Descriptive Model for Cuticle-Based Skeleto-Muscular Systems, an Ontology for Insect Anatomy, and their Potential Applications in Biodiversity Research and Informatics

Jennifer C. Girón[1,2,*], Sergei Tarasov[3], Luis Antonio González Montaña[4], Nicolas Matentzoglu[5], Aaron D. Smith[1], Markus Koch[6], Brendon E. Boudinot[7,8,16], Patrice Bouchard[9], Roger Burks[10], Lars Vogt[11], Matthew Yoder[12], David Osumi-Sutherland[13], Frank Friedrich[14], Rolf G. Beutel[8], and István Mikó[15]

[1]*Department of Entomology, Purdue University, West Lafayette, IN, USA*
[2]*Natural Science Research Laboratory, Museum of Texas Tech University, Lubbock, TX, USA*
[3]*Finnish Museum of Natural History, University of Helsinki, Pohjoinen Rautatiekatu 13, FI-00014 Helsinki, Finland*
[4]*Facultad de Ciencias Básicas e Ingeniería, Universidad de los Llanos, Villavicencio, Meta, Colombia*
[5]*Semanticly Ltd., London, UK*
[6]*Institute of Evolutionary Biology and Ecology, University of Bonn, An der Immenburg 1, 53121 Bonn, Germany*
[7]*Department of Entomology & Nematology, University of California, Davis, One Shields Ave, CA, USA*
[8]*Institut für Zoologie und Evolutionsforschung, Friedrich-Schiller-Universität Jena, Erbertstraße 1, 07743 Jena, Germany*
[9]*Biodiversity and Bioresources, Canadian National Collection of Insects, Arachnids and Nematodes, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario, K1A 0C6, Canada*
[10]*Entomology Department, University of California, Riverside, 900 University Ave. Riverside, CA, USA*
[11]*TIB Leibniz Information Centre for Science and Technology, Welfengarten 1B, 30167 Hannover, Germany*
[12]*Illinois Natural History Survey, University of Illinois, Champaign, IL, USA*
[13]*European Bioinformatics institute (EMBL-EBI) WellcomeTrust Genome Campus, Cambridge UK*
[14]*Institut für Zell- und Systembiologie der Tiere, Universität Hamburg, Martin-Luther-King-Platz 3, 20146, Hamburg, Germany and*
[15]*Department of Biological Sciences, University of New Hampshire, Durham, NH, USA*
[16]*Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington DC, USA*
**Correspondence to be sent to: Jennifer C. Girón, Museum of Texas Tech University, 3301 4th St, Lubbock, TX 79415, USA; E-mail: entiminae@gmail.com.*

Received 16 February 2022; reviews returned 17 April 2023; accepted 21 April 2023
Associate editor: Lauren Esposito

Abstract.—The spectacular radiation of insects has produced a stunning diversity of phenotypes. During the past 250 years, research on insect systematics has generated hundreds of terms for naming and comparing them. In its current form, this terminological diversity is presented in natural language and lacks formalization, which prohibits computer-assisted comparison using semantic web technologies. Here we propose a Model for Describing Cuticular Anatomical Structures (MoDCAS) which incorporates structural properties and positional relationships for standardized, consistent, and reproducible descriptions of arthropod phenotypes. We applied the MoDCAS framework in creating the ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM). The AISM is the first general insect ontology that aims to cover all taxa by providing generalized, fully logical, and queryable, definitions for each term. It was built using the Ontology Development Kit (ODK), which maximizes interoperability with Uberon (Uberon multispecies anatomy ontology) and other basic ontologies, enhancing the integration of insect anatomy into the broader biological sciences. A template system for adding new terms, extending, and linking the AISM to additional anatomical, phenotypic, genetic, and chemical ontologies is also introduced. The AISM is proposed as the backbone for taxon-specific insect ontologies and has potential applications spanning systematic biology and biodiversity informatics, allowing users to: 1) use controlled vocabularies and create semiautomated computer-parsable insect morphological descriptions; 2) integrate insect morphology into broader fields of research, including ontology-informed phylogenetic methods, logical homology hypothesis testing, evo-devo studies, and genotype to phenotype mapping; and 3) automate the extraction of morphological data from the literature, enabling the generation of large-scale phenomic data, by facilitating the production and testing of informatic tools able to extract, link, annotate, and process morphological data. This descriptive model and its ontological applications will allow for clear and semantically interoperable integration of arthropod phenotypes in biodiversity studies. [Biodiversity research; insects; morphology; ontology development.]

The ubiquitous distribution and stunning species richness of insects has generated a great diversity of phenotypes that fuel research in biodiversity, systematics, and various other biological fields. Roughly 90% of studies describing insect anatomy deal with structures related to the skeleto-muscular system (Deans et al. 2012a; Iyer et al. 2016; Adachi et al. 2020; Sommer 2020; Gotoh et al. 2021); the remaining 10%, in general, deal with the nervous system (e.g., Loesel et al. 2013), the midgut (e.g., Monteiro et al. 2014), the endocrine system (e.g., Page and Amdam 2007), the fat body (e.g., de Oliveira and Cruz-Landim, 2003), etc. Thousands of morphological terms referring to the insect skeleto-muscular system have historically emerged due to several general processes: 1) most basic terms (e.g., head, wings, and legs, etc.) have been borrowed from vertebrate anatomy due to functional or positional similarity (Snodgrass 1963); 2) some terms have been created *de novo* to name exclusive insect (or arthropod) structures (i.e., sclerite and tergite; Snodgrass 1963); 3) many terms have been repeatedly adopted across distant insect lineages to name similar structures located in similar areas of the body (e.g., cercus in Diplura vs. cercus in Hymenoptera; Snodgrass 1993); 4) the continuous reassessment of insect morphology in light of new comparative or phylogenetic data, constantly changes terms and their definitions; and 5) often, the definition of a term (which has been given in natural language) in subsequent studies, as in the "telephone game", suffers from interpretational deviations, thereby, producing a significantly different meaning that may eventually become widely adopted.

The interplay of these term-generating processes brings two major persisting problems. First, numerous terms in the corpus seriously suffer from semantic ambiguities such as homonymy (the same term is used for unrelated structures), polysemy (the same term is used for different but related-similar-structures), and synonymy (different terms with the same meaning) (Bolshoy and Lacková 2021). In addition, semantic ambiguities in morphological nomenclature have been historically reinforced by taxon-specific development and a lack of communication and agreement among morphological specialists across taxa (Vogt 2008; exemplified by genital terminology, Tuxen 1970). Second, many terms and definitions reflect the history of their usage rather than accurate anatomical concepts. For instance, the term "cercus" originally referred to appendages at the end of the abdomen, but it refers to different specific morphological entities depending on the taxonomic group (e.g., an appendage (with muscular attachment) composed of a single cercomere in Dermaptera vs. a cuticular protrusion (without muscular attachment) of the dorsal region of the postabdomen, composed of several cercomeres in Archaeognatha). Moreover, some terms refer to common spatio-structural properties, others refer to a common function or a common developmental or presumed common evolutionary origin, and some terms even refer to a mixture of these categories (Vogt et al. 2010). Consequently, interpreting and analyzing phenotypic data becomes unnecessarily difficult for nonexperts and integrating phenotype data with other sources of data in the life sciences is very difficult and time-consuming.

These problems are compounded by our tendency to see and characterize elements and developmental/evolutionary processes of the insect exoskeleton similarly to those of vertebrate anatomy (Snodgrass 1963), which has resulted not only in a misunderstanding of insect evolution and development but also in an overcomplicated system that worsens the above-mentioned issues of insect morphological terminology. Bones, the main elements of the vertebrate endoskeleton, develop from well separated cell clusters into a complex 3D scaffold of cells with different function and origin (Bitsch and Bitsch 2002; Wang et al. 2017; Blumer 2021). Bones are connected to each other by different types of joints, whose accurate functioning requires the interplay of unrelated elements, including ligaments, articular cartilages, and synovial fluid (Blumer 2021). In comparison, the insect exoskeleton, formed by chitinous cuticle, is an acellular product of the single-layered outer epithelium, the epidermis (Hall 1975; Adler 2017; Denk-Lobnig and Martin 2020) and its stiff elements, the sclerites (red/orange areas in Fig. 1), are only more rigid regions of the cuticle that are surrounded by more flexible ones of the same origin (conjunctivae), granting mobility (green areas of the cuticle in Fig. 1). Therefore, the insect skeleto-muscular system is suitable for a comparatively simple model, using clearly identifiable and consistently organized building blocks of the continuous cuticle (Fig. 2). These building blocks—sclerites, conjunctiva, and formative elements (Klass 2008; Klass and Matushkina 2012; Table 1)—can also be used as anatomical landmark entities (i.e., disjointed intrinsically identifiable anatomical entities; Young 1993) for identifying units of comparison across different species (i.e., nonevolutionary comparative homology assessment; Vogt 2017).

It is worth pointing out that the distinction between building blocks is not always clear, because semimembranous areas also occur (e.g., often parts of the epipharynx, the wing articulation), so everything is more or less a continuum (Fig. 2), with (gradually) different degrees of sclerotization. The concept of "building blocks" (i.e., unambiguously defined sclerites), may be justified for pragmatic reasons in most cases, but it is still a simplification.

Ontologies have become a fundamental technology for semantic management (i.e., assigning logical definitions to terms; see Vogt et al. 2010) and inference with biological knowledge (Smith et al. 2007; Balhoff et al. 2010; Deans et al. 2015; Dahdul et al. 2018; Tarasov 2019). An ontology is a logic-based representation of concepts and their relationships across a domain for modeling complex interactions in data (Deans et al. 2012a; Balhoff et al. 2013; Deans et al. 2015). In biology, ontologies serve two major purposes: they can be used as controlled vocabularies for stabilizing terminology and facilitate communication between scientists (Deans et al. 2012b), and as engines for inferring new complementary knowledge out of the encoded data. Because the definition associated with each term in an ontology is logic-based, these definitions are stable through time and free of the interpretational issues carried by natural language definitions. Therefore, ontology is a suitable technology for addressing the problems of understanding and interpreting terminology in arthropod morphology.

To date, there are eight ontologies dealing with the anatomy of different arthropod lineages, five of them dedicated to hexapods: the Hymenoptera Anatomy Ontology (Yoder et al. 2010); the *Drosophila* Anatomy Ontology (Osumi-Sutherland et al. 2013); the

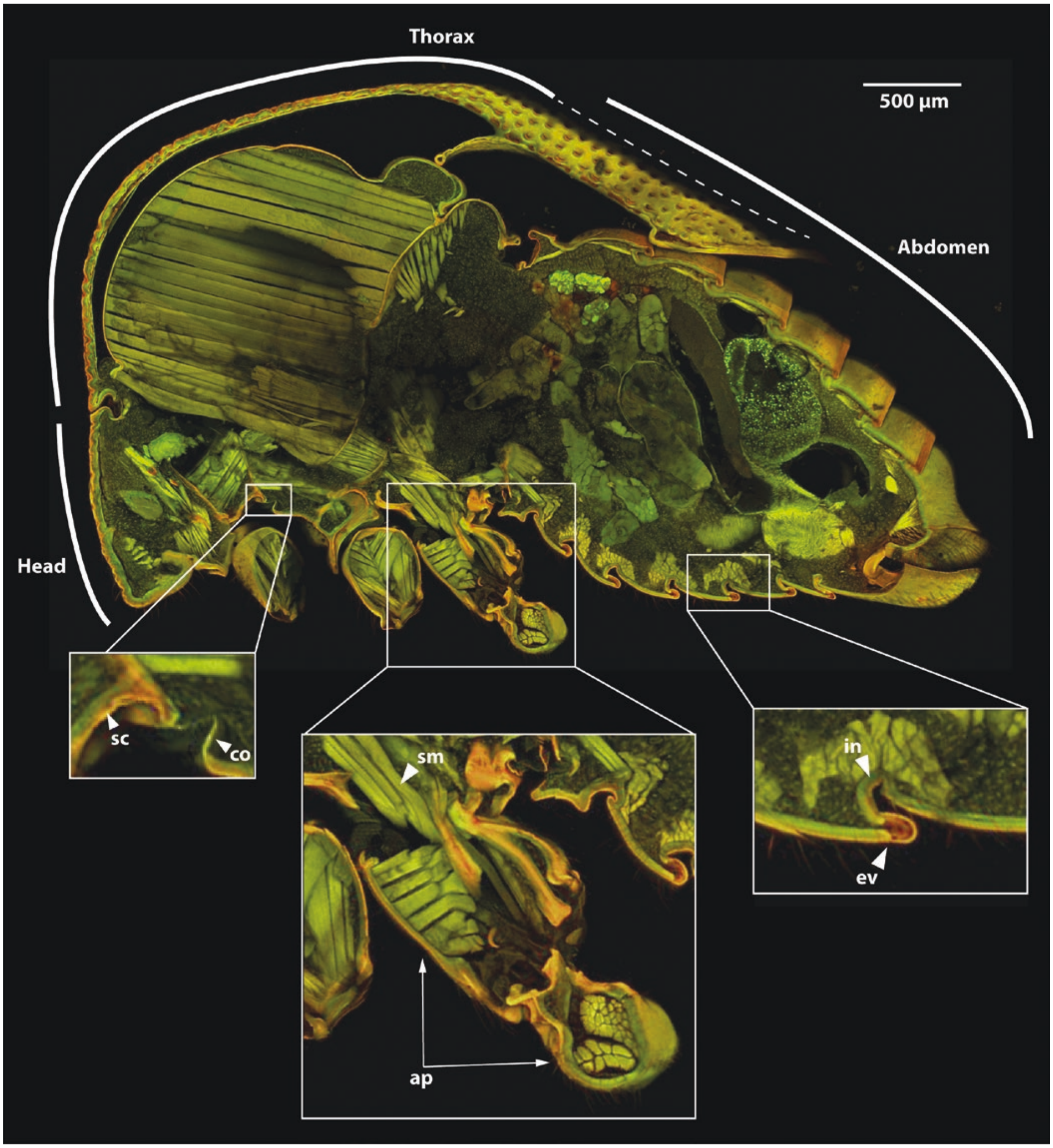

FIGURE 1. Autofluorescence-based CLSM (confocal laser scanning microscopy) micrograph showing the general structure of a sagittal section of the insect integument in an adult treehopper, *Ceresa* sp. (Membracidae). Excitation wavelength: 488. Emission wavelengths: 500–580 pseudocolor green for conjunctivae (more flexible cuticle), muscles, and other soft structures and 580–700 pseudocolor red for sclerotized components (orange surfaces, more stiff, less flexible cuticle). Abbreviations: sc: sclerite; co: conjunctiva; sm: skeletal muscle; ap: appendage; in: invagination; ev: evagination.

Tribolium Ontology (Dönitz et al. 2013); the Mosquito Ontology (Topalis et al. 2008), the Collembola Anatomy Ontology (González-Montaña 2023a), the Spider Anatomy Ontology (Ramírez and Michalik 2019), the Tick Anatomy Ontology (Topalis et al. 2008), and the Ontology of Arthropod Circulatory Systems (Wirkner et al. 2017). However, given the narrow scope and purpose of each, none of them can be generally applied to insects as a whole (e.g., Bertone et al. 2013). For the most part, these existing ontologies do not consider the interconnectedness of the whole cuticular system in their definitions, and those definitions tend to be idiosyncratic in the sense that they are taxon-specific and provide only textual/natural language definitions without much of a logical description, which prevents ontology-wide reasoning and inference. The properties of the cuticle are fundamental to arthropods and Ecdysozoa more broadly, and they underpin a substantial portion of animal biodiversity, thus are critical to account for in descriptive, experimental, and phylogenetic studies.

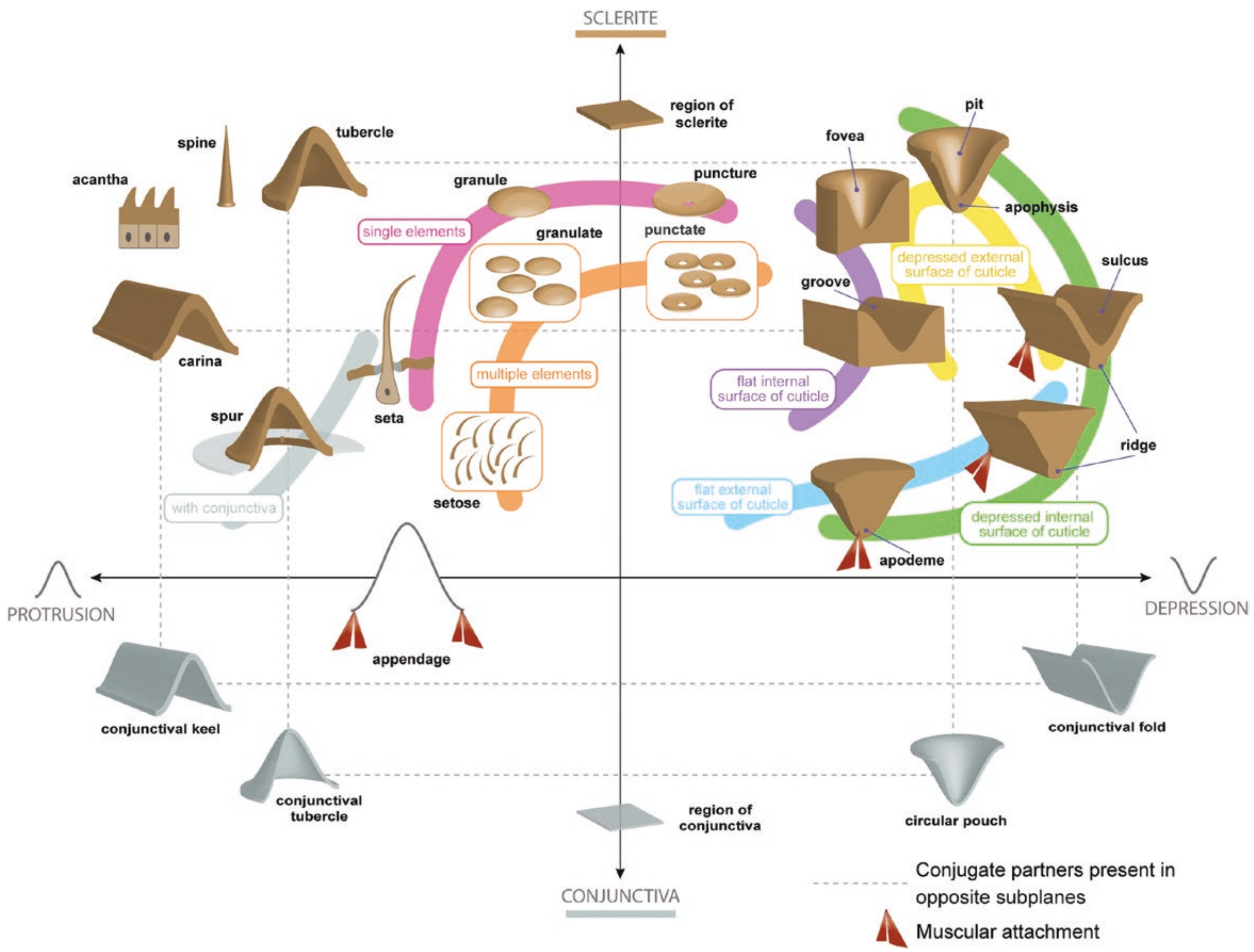

Figure 2. Schematic representation of some of the structural components of MoDCAS: a conceptual model for describing cuticular anatomical structures. The two principal structural properties that characterize the cuticle are: 1) degree of flexibility (y-axis), ranging from sclerite (stiff, at the top) to conjunctiva (flexible, at the bottom), and 2) degree of curvature (x-axis; protrusion –left– to flat –center– to depression –right–); the degree of curvature of the external and internal surfaces of cuticular protrusions and depressions can be different: when both run in parallel, they form hollow protrusions (top left subplane, e.g., carina, tubercle) or hollow depressions (top right subplane, yellow band, e.g., pit and sulcus); the external surface can be depressed with the internal surface flat (top right subplane, purple band, e.g., fovea and groove); the external surface can be flat with the internal surface depressed (top right subplane, blue band, e.g., ridge and apodeme). Additional properties: quantity (single vs. multiple elements; pink and orange bands, respectively); shape can be observed throughout each subplane (e.g., sclerotized protrusions can range from elongated –carina– to rounded –tubercle–); same for depressions (elongate –groove– vs. rounded –fovea–). The color bands in the top subplane indicate that different stages are not discrete; a myriad of intermediate cuticular elements exist.

As a starting point towards a more stable, understandable, and interoperable terminology in arthropod morphology, in this study we provide a conceptual Model for Describing Cuticular Anatomical Structures (MoDCAS) at any developmental stage, in both formal (i.e., interpretable by machines) and natural (i.e., interpretable by humans) languages, and solely based on the structural properties and topological relationships of each anatomical structure. This model can be applied to create arthropod-specific ontologies, thus enhancing the translation of biological knowledge among major clades. As a case study, we applied MoDCAS to generate the first universally applicable anatomy ontology for insects, the Anatomy Ontology of Skeleto-Muscular system (AISM), which is a formalized representation of MoDCAS that incorporates general terms for insect anatomy, including generalized definitions, although integrating them with other relevant ontologies. We provide ontology reasoning examples using the AISM and demonstrate its robustness and extensibility using the Ontology Development Kit (ODK; Matentzoglu et al. 2022; https://github.com/INCATools/ontology-development-kit). The AISM provides a computer-parsable controlled vocabulary for the insect skeleto-muscular system with a broad range of applications, including service as a backbone for taxon-specific ontologies, the provision of opportunities for mining data from existing literature, as well as producing semantically enhanced descriptions. It also has the potential for integration in evo-devo research, phenotype to genotype mapping, and logical homology assessment analyses.

TABLE 1 Basic classes and spatial terms for ontologies of cuticle-based systems

Subclass	Definition	URI
region of cuticle	The region of the insect integument (UBERON:6007284) that is part of chitin-based cuticle (UBERON:0001001)	AISM:0000174
sclerite	The region of the cuticle (AISM:0000174) that is less flexible than the neighboring conjunctiva(e) [conjunctiva(e) (AISM:0000004) that the sclerite is continuous with]	AISM:0000003
conjunctiva	The region of the cuticle (AISM:0000174) that is more flexible than the neighboring sclerite(s) (AISM:0000003) [sclerite(s) that the conjunctiva is continuous with]	AISM:0000004
cuticular depression	The region of the cuticle that corresponds to a concave surface	AISM:0000005
cuticular invagination	The region of cuticle (AISM:0000174) that corresponds with an invagination of the single layer epidermis (epithelial fold; UBERON:0005157). The cuticular invagination sometimes corresponds to a cuticular depression (concavity on the surface of the cuticle; AISM:0000005)	AISM:0000006
cuticular protrusion	The region of the cuticle that corresponds to a convex surface	AISM:0000008
cuticular evagination	The region of cuticle (AISM:0000174) that corresponds with an evagination of the cuticle and the single layer epidermis (epidermal fold; UBERON:0005157). The cuticular evagination usually corresponds to a cuticular protrusion (convexity on the surface of the cuticle; AISM:0000008)	AISM:0000027
anatomical region	A 3D region in space without well-defined compartmental boundaries; for example, the dorsal region of an ectoderm. [e.g., anterior region (BSPO:0000071); lateral region (BSPO:0000082); ventral margin (BSPO:0000684)]	BSPO:0000070
somatic muscle	A muscle structure (UBERON:0005090) of invertebrates whose origin and insertion sites are in basal side of the epidermis or structures derived from it. The simplest somatic muscles consist of a single cell and associated extracellular structures.	UBERON:0014895

MATERIALS AND METHODS

Model for Describing Cuticular Anatomical Structures (MoDCAS)

The Model for Describing Cuticular Anatomical Structures (MoDCAS) incorporates structural properties and topological relationships to characterize the anatomical structures used in morphological descriptions that involve the arthropod skeleto-muscular system. We adopted anatomical concepts from Richards and Richards (1979), Snodgrass (1963), and Klass (2008) regarding the structural properties of the insect cuticle to define elementary building blocks (Fig. 2). For the topological relationships, we used terms referring to the relative position of a given block along the body axes (dorsal, lateral, and distal, etc.) and its connectedness to other structures (e.g., continuity and attachment). Using this model, each cuticular anatomical structure can be described and defined as one or more building blocks that are specifically related to other building blocks.

Creating and Editing the AISM

The ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM.owl) and accompanying file system were generated using the Ontology Development Kit (ODK, Matentzoglu et al. 2022; https://github.com/INCATools/ontology-development-kit) and edited with Protégé version 5.5.0 (Musen 2015). All the files are available on GitHub at https://github.com/insect-morphology/aism (Girón et al. 2023a). The ODK uses ROBOT-based workflows (Jackson et al. 2019; http://robot.obolibrary.org/) to automatically generate imports from related external ontologies including Uberon (Mungall et al. 2012) and the OBO relations ontology (Smith et al. 2005), and to drive quality control tests under continuous integration. It also provides a semiautomated release process, supporting the generation of release products enhanced by the results of OWL reasoning, preventing duplication of terms and logical definitions.

Throughout this text, we use **bold lettering** to indicate ontology classes, *italics* when referring to object properties and use ID numbers for the specification of each. ID numbers are composed of the ontology prefix followed by a colon and a seven-digit number, e.g., AISM:0000003. This ID represents an OBO persistent, unique identifier (e.g., http://purl.obolibrary.org/obo/AISM_0000003) that links to online versions of the encoded information in ontology repositories including OntoBee (http://www.ontobee.org/), OLS (https://www.ebi.ac.uk/ols/ontologies), and BioPortal (https://bioportal.bioontology.org/ontologies).

Following the principles proposed by MoDCAS, we created terms referring to the elementary building blocks of the insect skeleto-muscular system (Table 1), as well as generalized terms from the glossary presented by Beutel et al. (2014). Each term has a label and a series of specific annotation properties including sensu, *definition* and *contributor*, and *has exact synonym* (Table 2).

Each AISM term is unambiguously labeled using the annotation property OBO foundry unique label (IAO:0000589) and formally represented by as many subclasses of descriptors as necessary to clearly characterize the term using object properties and associated classes. Terms from existing general anatomy ontologies [e.g., Uberon multispecies anatomy ontology (Mungall et al. 2012); BFO: Basic Formal Ontology (Spear et al. 2016)], and supporting ontologies [e.g., BSPO: Biological Spatial Ontology (Dahdul et al. 2014); CARO: Common Anatomy Reference Ontology (Haendel et al. 2008); PATO: Phenotype And Trait Ontology (Gkoutos et al. 2005); RO: Relations Ontology (Mungall et al. 2023)] were imported using the ODK.

TABLE 2 Main annotation properties used in the AISM (ontology for the anatomy of the insect skeleto-muscular system)

Annotation property	AISM usage	URI
rdfs:label	A term indicated by a word or set of words to unambiguously name an insect anatomical structure	https://www.w3.org/2000/01/rdf-schema#label
definition	A natural language statement to describe an insect anatomical structure, constructed by articulating the appropriate subclass of descriptors	http://purl.obolibrary.org/obo/IAO_0000115
has exact synonym	Alternative labels applied to the defined insect anatomical structure. Should be accompanied by a sensu annotation	http://www.geneontology.org/formats/oboInOwl#hasExactSynonym
OBO foundry unique label	An alternative name for a class or property which is unique across the OBO Foundry	http://purl.obolibrary.org/obo/IAO_0000589
sensu	Bibliographic reference with its corresponding DOI (or other link to it), and the textual definition of the term according to that reference	http://purl.obolibrary.org/obo/AISM_0000171
dc:contributor	The person who composed the definition or added the subclass of descriptor	http://purl.obolibrary.org/dc/elements/1.1/contributor
creation_date	The date when the definition was composed in year-month–day format	http://geneontology.org/formats/oboInOwl#creation_date
date_modified	Date on which the resource was changed	http://purl.org/dc/terms/modified
curator note	Additional comments to clarify or expand on the presented definition. Should be accompanied by contributor and creation date.	http://purl.obolibrary.org/obo/IAO_0000232
foaf:depiction	Associated image or images illustrating the structure being defined, linked by a DOI or URL.	http://xmlns.com/foaf/0.1/depiction

An effort was made to maximize the inclusion of existing logically defined terms, avoiding duplication of existing object properties and general higher classes; we made sure that the class definitions offered in existing ontologies were compatible with the intended usage in the AISM before importing a class. Each term of the AISM is accompanied by a verbatim logical definition that translates each set of subclasses of descriptors into natural language. Each definition in the AISM is intended to be broad enough to be applicable across Insecta, in a similar way as Uberon provides generalized definitions for animals (Mungall et al. 2012). As a convention, labels for muscles and conjunctivae are given in English, using their sites of attachment from proximal to distal, anterior to posterior, or dorsal to ventral.

We used *continuous with* (RO:0002150) for sclerite-conjunctiva attachments, whereas *adjacent to* (RO:0002220) for sclerite-sclerite articulations. We also propose the object properties *encircles* (AISM:0000078) and *encircled by* (AISM:0000079) to annotate the relationship between ring sclerites and their corresponding conjunctivae (e.g., femur, antennomere).

Reasoning

We used a ROBOT template (https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv; Jackson et al. 2019) to create AISM-based instances and definitions to demonstrate how the terms and generalized definitions provided in the AISM can be used to fit insect taxon-specific definitions more closely. In this template, we represented different paired cuticular structures of the abdominal tergites as individuals (instances), for the orders Archaeognatha, Zygentoma, Dermaptera, Ephemeroptera (Baetidae), Hemiptera (Aphididae), Psocodea, and Coleoptera (Carabidae larva) (Table 3; see Fig. 3 for a schematic representation of these definitions). Using ROBOT (Jackson et al. 2019; http://robot.obolibrary.org/) we generated an OWL file from this template, which included terms from the AISM and other ontologies. This template-based OWL file was then merged with the AISM. We ran a series of DL queries in Protégé, using ELK 0.5 as a reasoner on this merged ontology to verify the fit of the provided taxon-specific definitions with the terms and definitions available in the AISM (Table 4). The expectation was that the queries would return the appropriate instances, depending on the properties indicated in the template.

In addition, to demonstrate the interoperability of the AISM with existing ontologies, we provide an example of how to describe a particular insect species phenotype (the yellow profemur on a chalcid wasp), by concatenating classes and object properties. The ability to relate structures across additional ontologies was also illustrated by linking structures of the AISM with the circulatory system using relationships from the Relation Ontology (RO) and terms from the Ontology of Arthropod Circulatory Systems (OArCS; Wirkner et al. 2017).

Results and Discussion

MoDCAS: Model for Describing Cuticular Anatomical Structures

MoDCAS is a descriptive model: a set of principles based on traditional approaches to classify and characterize cuticular elements based on their structural properties and topology. It is proposed as a baseline to generate consistent and reproducible descriptions of cuticle-based skeleto-muscular structures across Arthropoda.

The arthropod endo- and the exo-skeleton is a continuous entity that can be considered as a single

TABLE 3 Example of a template to specify new terms to include in the AISM or AISM-derived ontologies

Term suggestion	Definition	Type of cuticular element	Location	Laterality	***Anterior to***	***Lateral to***	***Has part***
cercus_ Archaeognatha	The paired protrusion of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	**'cuticular protrusion'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		**cercomere**
cercus_ Zygentoma	The paired appendage of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	**'insect appendage segment'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		**cercomere**
cercus_ Dermaptera	The paired appendage of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	**'insect appendage segment'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		**cercomere**
tergalius_ Ephemeroptera	The paired appendage of the dorsal region of the preabdomen that is lateral to the abdominal tergite.	**'insect appendage segment'**	**preabdomen**	**'bilaterally paired'**		**'abdominal tergite'**	
cercus_ Ephemeroptera	The paired appendage of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	**'insect appendage segment'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		**cercomere**
cornicle_ Aphididae	The paired cuticular protrusion of the dorsal region of the preabdomen.	**'cuticular protrusion'**	**preabdomen**	**'bilaterally paired'**			
cercus_ Psocodea	The paired region of the dorsal region of the postabdomen that is anterior to the anus and composed of a collection of setae.	**'region of cuticle'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		**'setose cuticle'**
urogomphus_ Carabidae	The paired cuticular protrusion of the dorsal region of the postabdomen that is anterior to the anus	**'cuticular protrusion'**	**'dorsal postabdomen'**	**'bilaterally paired'**	**'insect anus'**		

Note: An extended version of this template with additional descriptors is available at https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv. In a template, terms composed of more than one word need to be indicated in single quotation marks for ROBOT to be able to recognize the appropriate term from the ontology. Class identifiers: **abdominal tergite** (AISM:0004057); **cuticular protrusion** (AISM:0000008); **dorsal postabdomen** (AISM:0000523); **bilaterally paired** (PATO:0040024); **cercomere** (AISM:0004199); **insect appendage segment** (AISM:0004284); **insect anus** (AISM:0004197); **preabdomen** (AISM:0004055); **region of cuticle** (AISM:0000174); **setose cuticle** (AISM:0000530). Figure 3 illustrates the taxon-specific terms.

anatomical structure (Klass 2008; similar to the skin of a vertebrate). Cuticular elements can be defined and distinguished from each other by variations along five key properties: 1) degree of flexibility (i.e., stiffness or resistance to deflection: sclerite vs. conjunctiva; Fig. 2, *y*-axis), 2) degree of surface curvature (i.e., deviations from a flat surface: depression vs. protrusion; Fig. 2, *x*-axis), 3) presence of muscular attachments, 4) quantity (single vs. multiple, repeated cuticular elements—generating sculpture and pilosity; Fig. 2, pink vs. orange bands, respectively), and 5) shape (circular and elongate). The interplay of these properties determines the features of each elementary building block that, together with its topological relations and connectedness, allows for the modeling of the entire structural diversity of the arthropod skeleto-muscular system. Similar categorical sets of properties have been employed in other semantic descriptive models of anatomical systems (e.g., OArCS, Ontology of Arthropod Circulatory System; Wirkner et al. 2017) to allow for better data structuring.

Sclerites and conjunctivae have been historically used for describing insect morphological structures, although their definitions and the parameters used to differentiate them have been inconsistent through time. Even though these different regions of the cuticle are often characterized by their histological properties (sclerites with thick exocuticle with sclerotin and endocuticle vs. conjunctiva with thin exocuticle and thick endocuticle without sclerotin or entirely devoid of exocuticle; Beutel et al. 2014), these are not discernible without histological sections and are not possible to obtain from dry specimens. Therefore, we use the degree of flexibility to differentiate these cuticular elements. Based on their relative degree of flexibility there are two main regions of the cuticle: 1) sclerites, which are relatively stiff, and 2) conjunctivae, which are relatively flexible and provide mobility (Klass and Matushkina 2012; Fig. 2, *y*-axis).

The degree of curvature of a cuticular region can mirror changes in the single-layered epithelium, but also depends on the properties of the cuticle and is

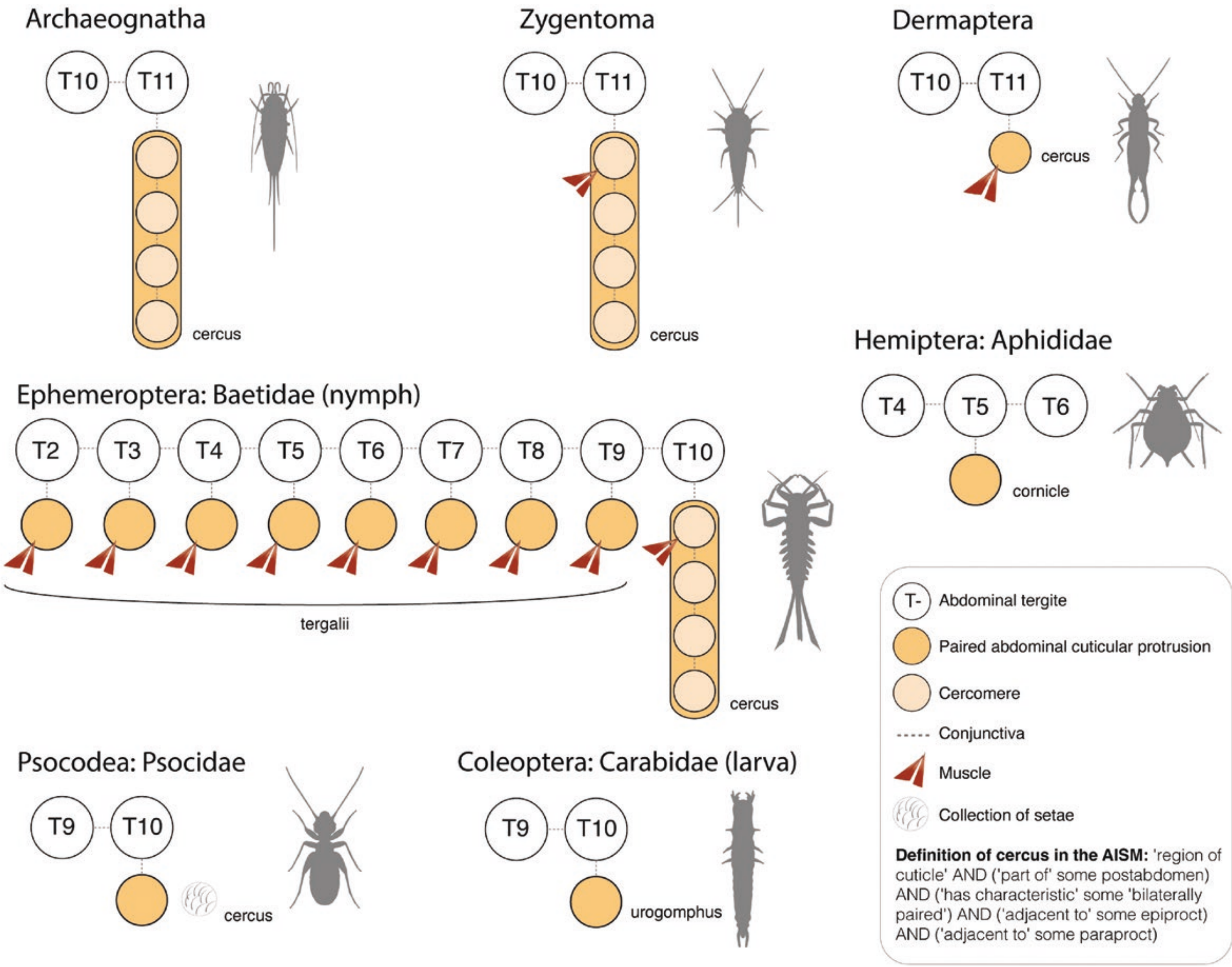

FIGURE 3. Schematic representation of taxon-specific definitions for paired cuticular structures of the abdominal tergites (not including structures of the genitalia) across different insect orders. Textual definitions for each structure are provided in Table 3.

independent of the epithelial geometry. To classify superficially similar cuticular specializations (e.g., spine, spur, and seta; Fig. 2) we must explicitly define the causes of the change in cuticular morphology (cuticular or epithelial). Therefore, we classify regions of the cuticle by the degree of curvature of their internal and external surfaces (Fig. 2, *x*-axis and color bands on top right subplane), accounting for the morphology of the single layer epithelium and the external surface of the cuticle. The external surface can be flat, convex (cuticular protrusions) or concave (cuticular depressions).

Cuticular protrusions, if they correspond to evaginations of the cuticle (i.e., cuticular protrusion corresponding to an internal cuticular depression; Fig. 2, top left subplane), can correspond to either the evagination of a single cell membrane (e.g., seta) or to the evagination of a region of the single-layered outer epithelium (e.g., spurs, lobes) (Richards and Richards 1979). Appendages differ from other cuticular protrusions (e.g., spurs or lobes) in that they are connected to the rest of the body by somatic muscles (Fig. 2). When individual elements like a carina or a seta are repeated across a region of the cuticle (Fig. 2, orange band), they generate texture on that particular surface, forming sculpture or pilosity, respectively. Ring sclerites often represent repetitive subdivisions of appendages that can be either musculated (appendage segments) or nonmusculated (meres). Cuticular depressions (Fig. 2, top right subplane) vary in constitution depending on the orientation of the external and the internal surfaces of the cuticle: when both run in parallel, they form hollow depressions (e.g., pit, sulcus; Fig. 2, top of the green band); the external surface can be depressed, with the internal surface flat (e.g., fovea, groove; Fig. 2, purple band); or the external surface can be flat, with the internal surface depressed (e.g., ridge, apodeme; Fig. 2, blue band, the bottom of green band); this particular kind on cuticular depression forms strengthened areas across the body, providing mechanical stability, and frequently constitute sites for muscle attachment (Klass and Matushkina 2012; Beutel et al. 2014).

TABLE 4 Example DL queries and their results using the AISM

DL query	Resulting subclasses and individuals
Internal queries	
'cuticular protrusion' and ('*part of*' some '**insect head**')	Classes: '**insect mandible**', '**insect maxilla**', '**labial palpus**', **'maxillary palpus'**, **antenna**, **antennifer**, **galea**, **glossa**, **labium**, **labrum**, **lacinia**, **ligula**, **mouthpart**, **paraglossa**
'**insect appendage**' and ('*part of*' some '**insect thorax**')	Classes: '**fore leg**', '**fore wing**', '**hind leg**', '**hind wing**', '**insect leg**', '**insect wing**', '**mid leg**', **mesopretarsus**, **metapretarsus**, **pretarsus**, **propretarsus**
'*part of*' some '**fore leg**'	Classes: '**procoxal-protrochanteral conjunctiva**', '**profemoro-protibial conjunctiva**', '**protibio-protarsal conjunctiva**', '**protrochantero-profemoral conjunctiva**', **procoxa**, **profemur**, **propretarsus**, **protarsus**, **protibia**, **protrochanter**
'**appendage segment**' and '*part of*' some **antenna** and '*adjacent to*' some '**head capsule**'	Classes: **scapus**
Queries for taxon-specific definitions	
'*part of' some* '**insect abdomen**' and '**region of cuticle**' and '*has characteristic*' some '**bilaterally paired**'	Classes: '**gonocoxa IX**', '**gonocoxa VIII**', '**gonostylus IX**', '**gonostylus VIII**', **cercus**, **paramere** Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Psocodea, cercus_Zygentoma, cornicle_Aphididae, tergalius_ Ephemeroptera, urogomphus_Carabidae
'*part of*' some **preabdomen** and '**region of cuticle**' and '*has characteristic*' some '**bilaterally paired**'	Instances: cornicle_Aphididae, tergalius_Ephemeroptera
'*part of*' some **postabdomen** and '**region of cuticle**' and '*has characteristic*' some '**bilaterally paired**'	Classes: '**gonocoxa IX**', '**gonocoxa VIII**', '**gonostylus IX**', '**gonostylus VIII**', **cercus**, **paramere** Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Psocodea, cercus_Zygentoma, urogomphus_Carabidae
'*part of*' some **postabdomen** and '**cuticular protrusion**' and '*has characteristic*' some '**bilaterally paired**'	Classes: **paramere** Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma, urogomphus_Carabidae
'*part of*' some **postabdomen** and '**cuticular protrusion**' and '*has characteristic*' some '**bilaterally paired**' and '*anterior to*' some '**insect anus**'	Classes: none Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma, urogomphus_Carabidae
'*part of*' some **postabdomen** and '**insect appendage**' and '*has characteristic*' some '**bilaterally paired**' and *anterior_to* some '**insect anus**'	Classes: none Instances: cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma

Note: In Protégé, terms composed of more than one word need to be indicated in single quotation marks to be able to run a DL query. Each of the queries tested returned the expected outcomes in terms of subclasses and individuals. Class identifiers: *adjacent to* (RO:0002220); **antennifer** (AISM:0000190); *anterior to* (BSPO:0000096); **bilaterally paired** (PATO:0040024); **appendage segment** (AISM:0000063); **antenna** (AISM:0000032); **cercus** (AISM:0004165); **cuticular protrusion** (AISM:0000008); **fore leg** (AISM:0000034); **fore wing** (AISM:0000037); **galea** (AISM:0000023); **glossa** (AISM:0000049); **gonocoxa VIII** (AISM:0004068); **gonocoxa IX** (AISM:0000200); **gonostylus VIII** (AISM:0004076); **gonostylus IX** (AISM:0004198); *has characteristic* (RO:0000053); **head capsule** (AISM:0000019); **hind leg** (AISM:0000036); **hind wing** (AISM:0000038); **insect abdomen** (AISM:0000109); **insect anus** (AISM:0004197); **insect appendage** (AISM:0000029); **insect head** (AISM:0000107); **insect leg** (AISM:0000031); **insect mandible** (AISM:0000043); **insect maxilla** (AISM:0000044); **insect thorax** (AISM:0000108); **insect wing** (AISM:0000033); **labial palpus** (AISM:0000024); **labium** (AISM:0000087); **labrum** (AISM:0000042); **lacinia** (AISM:0000026); **ligula** (AISM:0000048); **maxillary palpus** (AISM:0000051); **mesopretarsus** (AISM:0004195); **metapretarsus** (AISM:0004196); **mid leg** (AISM:0000035); **mouthpart** (AISM:0000165); **paraglossa** (AISM:0000050); **paramere** (AISM:0004064); **procoxa** (AISM:0000066); **profemur** (AISM:0000070); **pretarsus** (AISM:0000047); **propretarsus** (AISM:0004194); **protarsus** (AISM:0004190); **protibia** (AISM:0000067); **protrochanter** (AISM:0004171); **scapus** (AISM:0000113); **postabdomen** (AISM:0004056); **preabdomen** (AISM:0004055); **procoxal-protrochanteral conjunctiva** (AISM:0000123); **profemoro-protibial conjunctiva** (AISM:0004179); **protibio-protarsal conjunctiva** (AISM:0004183); **protrochantero-profemoral conjunctiva** (AISM:0004175); **region of cuticle** (AISM:0000174).

Each of these elementary building blocks with their particular features can be specifically characterized by their connections to and spatial relations regarding other elementary building blocks, including topological relationships (e.g., dorsal, ventral, distal, proximal, medial, lateral), connectedness (e.g., continuous with, encircled by, adjacent to) and further phenotypic descriptors (color, relative size). This specific characterization results in accurate, consistent, and reproducible descriptions of cuticular anatomical structures. If employed correctly, MoDCAS-based natural language definitions should be easily translated into logical definitions and instance-based semantic phenotype descriptions of individual specimens using ontologies, so that information of cuticular skeleto-muscular systems can be accessible for machine processing.

MoDCAS can be applied to a broad range of organisms where the movable elements and the basis of motion are built on features of the cuticle. Based on MoDCAS, ontologies can be created for skeleto-muscular systems in ecdysozoans that bear cuticles as the major component of the exoskeleton and somatic muscles that are moving those parts (Table 1). Here we present a specific application of MoDCAS for the insect anatomical system as an example of how those ontologies can be built using the Ontology Development Kit (ODK, Matentzoglu et al. 2022; https://github.com/INCATools/ontology-development-kit). We restrict this example ontology to insects because of the expertise of the authors of this contribution, but the ontology is easily expandable and interoperable with other ontologies. We provide a manual to edit and create AISM-based

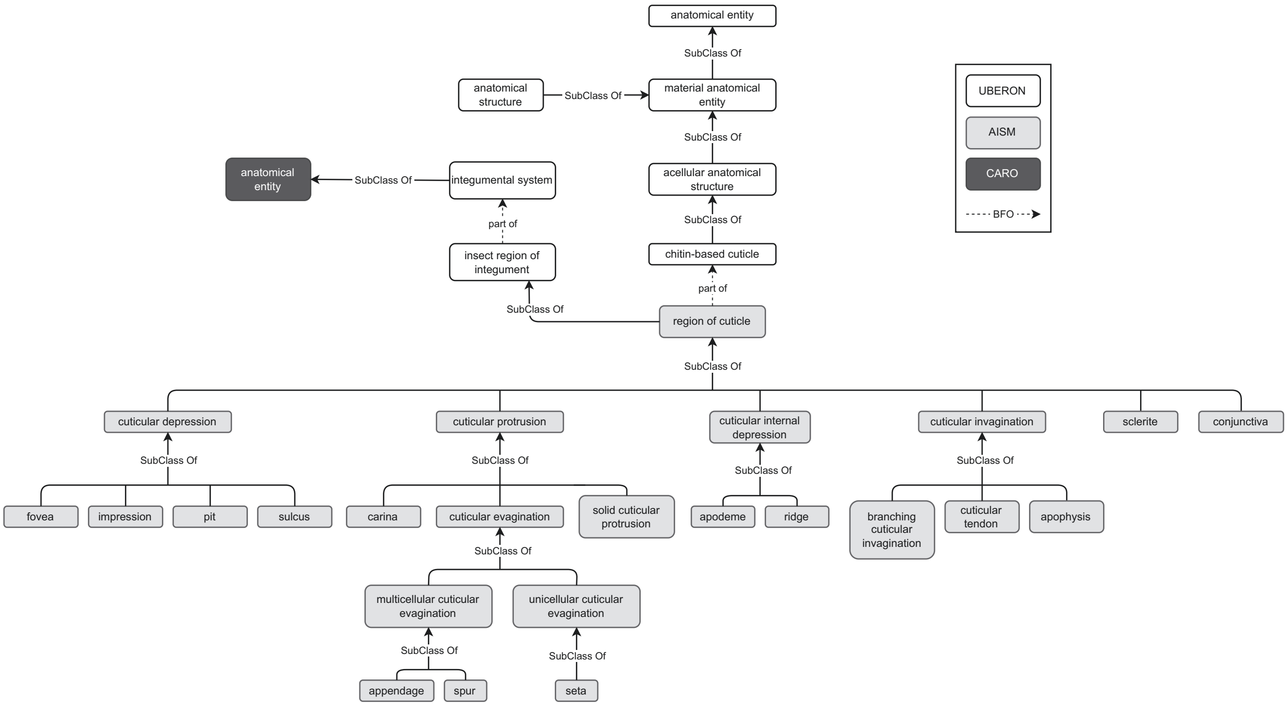

FIGURE 4. Graphic representation of high-level classes for ontologies of cuticle-based systems and some of their children, including hierarchy and elements from other ontologies.

ontologies (Girón et al. 2021), as well as a template system to create additional terms.

AISM: Ontology for the Anatomy of the Insect Skeleto-Muscular System

In its version v2023-04-14 (see https://github.com/insect-morphology/aism; via Dryad at https://doi.org/10.5061/dryad.7sqv9s4w6), the ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM) contains 7443 classes, where 586 are AISM terms; it uses 33 object properties and 22 annotation properties. All other terms have been imported from existing ontologies as part of the basic imports using the Ontology Development Kit (ODK; see methods), which not only brings the specified terms but also all their hierarchically associated terms to preserve the logical integrity of the ontology and maximize interoperability.

The insect integument, as a continuous structure, is composed of **chitin-based cuticle** (UBERON:0001001); therefore, every component of this continuous structure is designated as a **region of cuticle** (AISM:0000174), which is the parent class for all skeletal anatomical structures in the AISM (Fig. 4). Interpreting the skeleto-muscular system of insects as a set of consistently organized components and following the framework proposed by MoDCAS, each class included in the AISM is defined logically in OWL by some combination of: 1) kind of cuticular element (e.g., **sclerite** [AISM:0000003], **conjunctiva** [AISM:0000004], **cuticular depression** [AISM:0000005], **cuticular protrusion** [AISM:0000008], **skeletal muscle tissue** [UBERON:0001134], among others); 2) location of structure in the body (e.g., *part of* [BFO:0000050] the **insect thorax** [AISM:0000108], *anterior to* [BSPO:0000096] the **abdominal tergite I** [AISM:0000021]); 3) connected structures indicated by specific relations and spatial descriptors (e.g., *adjacent to* [RO:0002220] **posterior margin** [BSPO:0000672] of **abdominal sternite III** [AISM:0004105]). In this way the continuous nature of the cuticular integument is considered, making explicit statements about connectivity between parts and providing positional/spatial localization for each structure.

In addition, each class (i.e., label and descriptors) is accompanied by a set of annotation properties including a natural language definition that has been created from the annotated descriptors (or vice versa). These definitions include the contributor who constructed the definition and the date of creation and date of modification in format year-month-day. When available, references for textual definitions from the literature have been annotated on each label using the annotation property sensu (AISM:0000171), which includes the full citation of the reference, a DOI or link when the reference is available online, and the verbatim definition provided in the text, in quotation marks (see Yoder et al. 2010). When explicitly mentioned in the literature, synonyms are added using the *has exact synonym* object property indicating the sensu where this synonymy is proposed.

Using Templates to Curate and Extend the AISM

In order to make the AISM maximally accessible and reusable, the AISM aims for MoDCAS-based definitions that follow consistent and simple patterns that allow users to map terms or link them to available terms in the AISM. To ensure that users of the AISM can also easily generate MoDCAS/AISM-compliant descriptions of anatomical structures, we provide a template system for composing definitions (i.e., https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv; see also Table 3). This template system can be used to provide formal descriptions of cuticular anatomical structures, or for extending the AISM with taxon-specific terms (subclasses). Even if the template is not directly used, it provides guidance for the types of definitions that are compatible with the MoDCAS/AISM approach to defining terms.

The aim of the template is to ensure that users provide the specific type of cuticular element (Fig. 2) and its appropriate location within the arthropod body. Users may further refine the location by specifying the structure's relative position via multiple statements using relations such as *adjacent to* (RO:0002220), *posterior to* (BSPO:0000099), and *dorsal to* (BSPO:0000098). The template also includes a free text comment column allowing additional information to be provided in a less formal manner. Once this detailed, MoDCAS-compliant description is provided, users may also propose a commonly used term for the described structure, such as "cercus." The advantage of this approach is that it forces users to provide an accurate description of the structure's properties and location not captured by the generally used term.

Even without additional processing, the filled-out template constitutes a shareable and accessible controlled description of anatomical structures. Because the template corresponds to a standard OWL template system, it can also be used to generate new terms or instances for cuticle-based systems, including the AISM, or for extending the AISM with taxon-specific terms.

In Table 3 we present a few examples of terms to refer to different paired structures of the abdomen of different insect taxa. For instance, the cercus of Archaeognatha, defined as the paired protrusion of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres, as indicated in the template specifying its type of cuticular element (**cuticular protrusion**, AISM:0000008), its location (**dorsal postabdomen**, AISM:0000523), its laterality (**bilaterally paired**, PATO:0040024), its position regarding other structures (*anterior to* [BSPO:0000096] **insect anus** [AISM:0004197]), and its composition (**cercomere**, AISM:0004199). As examples are specified, OWL reasoning can be used to provide a list of candidate terms in the AISM that conform to the definition (see for example the query for '*part of*' some **postabdomen** and '**region of cuticle'** and '*has characteristic*' some '**bilaterally paired'**, which results in the classes '**gonocoxa IX'**, '**gonocoxa VIII'**, '**gonostylus IX'**, '**gonostylus VIII'**, **cercus**, **paramere**; Table 4).

The current implementation of the templates relies on users following the specification. It is possible to use the CEDAR template system (https://more.metadatacenter.org/tools-training/cedar-template-tools/#design-template) to provide auto-completion and constraints on column content, guiding and constraining users to ensure that the correct types of terms are added in each column. It is also possible to integrate a term suggestion option to avoid replication and detect potential synonyms.

Reasoning with the AISM

Each of the queries tested returned the expected outcomes in terms of subclasses and individuals (Table 4). Across Insecta, abdominal protrusions are highly variable in position, shape, and components, and in many cases, the morphological interpretations of these structures and their features have been problematic over time. The terms and broad definitions presented in the AISM have the capability to incorporate the broad variation presented in our example taxa. By adding subclasses and relationships to AISM terms it is possible to characterize taxon-specific structures. For instance, the cornicles of Aphididae (Hemiptera) are paired cuticular protrusions located on the dorsal surface of the abdominal tergite 5 (sometimes abdominal tergite 6); the existing terms and definitions incorporated in the AISM allow for accommodating all these details into a definition for cornicle in a potential Hemiptera-specific ontology. Similarly, the tergalii of Ephemeroptera, paired appendages of the preabdomen located on the lateral region of the abdominal tergites (Kluge 2004), can be easily defined and the particular abdominal tergites where the tergalii are present could be specified.

The different kinds of cerci present in our example taxa were also easily characterized, as they follow the generalized definition proposed in the AISM (the bilaterally paired region of the cuticle of the postabdomen, that is anterior to the anus; see Table 4, third query for taxon-specific definitions): in Dermaptera the cercus was characterized as an appendage (with muscular attachment) and composed of a single cercomere (Table 3 and Fig. 3). In Psocodea it was defined as a region of the cuticle that is anterior to the anus and bears a collection of setae (Table 3 and Fig. 3); we followed the definition presented by Yoshizawa (2005), even though the definition of this particular surface (anterior to the anus) in Psocodea as cercus has been contentious. In Zygentoma and Ephemeroptera the cercus was characterized as a paired appendage (with muscular attachment) of the dorsal region of the postabdomen, composed of cercomeres (Table 3 and Fig. 3), whereas in Archaeognatha, the cercus is a cuticular protrusion (without muscular attachment) of the dorsal region of the postabdomen, composed of cercomeres (Table 3 and Fig. 3). Our queries also returned the Coleoptera urogomphus as a cercus, as this undivided cuticular protrusion satisfies the requirements of the AISM definition for cercus. This demonstrates the power of the AISM's homology-free approach, as these urogomphi are structurally equivalent to cercus, but not homologous, a similarity that would be obscured if we only relied on homology-biased terminology. On the other hand, the same query (Table 4, third query for taxon-specific definitions: '*part of* ' *some* '**cuticle of insect abdomen'** and '**region of cuticle**' and '*has characteristic*' some '**bilaterally paired**') did not recover the cornicles, as these are paired projections of the abdomen, but located in a different abdominal region. The amount of detail incorporated into each definition will depend on the intended use of the ontology. Indeed, the number and sequence of cercomeres can be specified, along with the presence of setae, scales, or other relevant features.

Describing Phenotypes with the AISM

The AISM has been conceived as the backbone ontology for insect taxon-specific ontologies. For instance, in the AISM-based Coleoptera-specific ontology (COLAO; Girón et al. 2023b), the class **elytron** (COLAO:0000000) is a subclass of the class **fore wing** (AISM:0000037) and includes descriptors related to specific properties of elytron, such as a subclass of the class **sclerite** (AISM:0000003). Similarly, in an AISM-based Diptera-specific ontology the class **haltere** would be a subclass of the class **hind wing** (AISM:0000038). In these examples, the broadly applicable terms contained in the AISM are superclasses for terms defining more taxon-specific anatomical structures.

In general, for describing specific insect phenotypes with the AISM, a series of Entity-Quality statements can be used (e.g., Washington et al. 2009), taking advantage of the high interoperability of the AISM and the broad range of available existing ontologies including those for the phenotype (PATO: Phenotype And Trait Ontology), taxonomy (NCBITaxon: National Center for Biotechnology Information organismal classification), and spatial relationships (BSPO: Biological Spatial Ontology), among others. The template system proposed here can accommodate additional descriptors and relationships to better define structures within the AISM. For instance, it is possible to represent phenotypes like a yellow profemur on a chalcid wasp: [(Chalcididae [NCBITaxon:92425] AND *has part* [BFO:0000051] some **profemur** [AISM:0000070]) AND (*has characteristic* [RO:0000053] some **yellow** [PATO:0000324])].

In our exercise linking AISM to OArCS, no new terms were required, just additional linkages between existing terms and relationships, for example, the **alary muscle** (OARCS:0000151) is *attached to* (RO:0002371) both, the **heart** (OARCS:0000253) and the **abdominal tergite** (AISM:0004057). The term **alary muscle** would be imported using the ODK, bringing the necessary hierarchically linked terms and properties to be able to construct the logically appropriate axioms.

Taxon-specific ontologies can be linked to specialized taxonomic ontologies if those were available (e.g., see Stucky 2019). An example of taxonomic ontologies is the Vertebrate Taxonomy Ontology (VTO; Midford et al. 2013), which provides a comprehensive taxonomic hierarchy for vertebrates. It incorporates classes from the Taxonomic rank vocabulary (http://obofoundry.org/ontology/taxrank.html) and the NCBI organismal classification (National Center for Biotechnology Information; http://www.obofoundry.org/ontology/ncbitaxon.html).

There are various examples of semantically enhanced taxonomic descriptions for insects (e.g., Mullins et al. 2012). Mikó et al. (2021) used PhenoScript (https://github.com/sergeitarasov/PhenoScript/wiki), which is an ontology-based programming language for describing species. The use and improvement of this kind of tool will create species descriptions that are logical in origin, allowing for direct data processing.

All these approaches to phenotypic descriptions can be implemented using and extending the proposed template system. Furthermore, integration between the AISM and existing ontologies like Uberon, the *Drosophila* Anatomy Ontology (FBBT; Costa et al. 2013), and the Hymenoptera Anatomy Ontology (HAO; Yoder et al. 2010) can be improved over time by adding cross-reference annotations to each shared term. Ontologies for arthropod lineages other than insects can also be easily created by taking advantage of the current availability of MoDCAS-based basic classes and spatial terms for ontologies of cuticle-based systems (Table 1).

Taxonomy, Morphology, and Evo-Devo: MoDCAS-Based Ontologies on Different Granularity Levels

Similar surface modifications of the cuticular skeleto-muscular system can correspond with cardinally different epithelial modifications: multicellular invaginations/evaginations of the epidermal cell layer (e.g., cuticular depressions, spurs, pits, and appendages), invaginations/evaginations of a single cell membrane (e.g., cuticular components of sensilla), changes in the thickness of the cuticle (i.e., modifications that do not correspond to any epithelial fold, e.g., impression, acantha, or carina), and in some cases the combination of these categories.

Changes in the geometry of the epithelial sheet that results in invaginations and evaginations are governed by genes that define changes in the shape of epidermal cells or regulate cell proliferation (Zartman and Shvartsman 2010; Hannezo et al. 2014; Gotoh et al. 2021), although those genes that are involved in the reorganization of the cytoskeleton are governing similar geometrical changes of the membrane of a single cell (Lees and Waddington 1942; Bitan et al. 2012; Djokic et al. 2020). A third set of genes are involved in surface characteristics that are related to cuticle thickness, which are related to processes regulating cuticle deposition (Adler 2017; Jan et al. 2017; Tajiri 2017; Zhao et al. 2017). These processes are also separated in time and space; evaginations and invaginations happen during the last larval and early pupal stages, although cuticle deposition starts in the late pupal stage (Andersen 2012).

It is evident that differentiating these superficially similar structures will be key for the accurate understanding of phenotypic diversity and morphological evolution. However, the differences between practical approaches to anatomy across different knowledge domains represent a huge communication gap that hinders progress towards a more integrative view of anatomy (Richards and Richards 1979): 1) morphology aims to interpret the structural identity and connectivity of anatomical structures (Snodgrass 1951); it uses dissections and section-based methods ranging from histology to µ-CT and usually focuses on a handful of specimens in each study; 2) taxonomy focuses for the most part on externally visible structures with diagnostic value; each study can involve thousands of specimens in a comparative framework; 3) evo-devo studies gene expression on developing structures; the taxonomic breadth is usually limited to model organisms that are reared under laboratory conditions.

These knowledge domains refer to anatomy at different granularity levels and from different frames of reference (Vogt 2019) across different shared themes (i.e., taxonomy, individual count, developmental stage), which causes interoperability problems and misunderstanding among disciplines, due to the shifting of concepts for anatomical entities. The AISM provides a controlled vocabulary to facilitate communication, by using an interconnected hierarchy of superficial cuticular elements (anatomical surfaces) and the hierarchy of deeper structures that reveal developmental and structural properties of the single-layered outer epithelium. The AISM provides an opportunity to link insect phenotypes to genotypes across developmental stages and taxonomic groups via the Gene Ontology (Ashburner 2000), and to metabolic processes via the Protein Ontology (PRO; Natale et al. 2017) and the ontology for Chemical Entities of Biological Interest (CHEBI; Hastings et al. 2016). With the support of developmental biology specialists, the construction of AISM-based order-specific ontologies would allow for the accommodation of additional terms for specific developmental stages as needed for work in evo-devo.

Homology and the AISM

Evolutionary homology is a central concept in biology, whereby structural similarity has evolved through shared ancestry in different taxa (Minelli and Fusco 2013; Wagner 2014). The definitions of the classes included in the AISM are descriptive in anatomical terms and may serve to assess the primary criteria of position and similarity (see comparative homology, units of comparison; Vogt 2017), so that instead of asking whether the cercus in Archaeognatha is homologous to the cercus in Psocodea, we can ask if a multisegmented appendage on the 11th tergite in one taxon is homologous with a setose patch on the 10th tergite of the other (see our examples for abdominal cuticular protrusions, Table 3). There are data models such as the one proposed by Mabee et al. (2020), where homology relationships can be logically formalized between anatomical structures of different taxonomic units. This approach requires elements from anatomy ontologies, taxonomic ontologies, and the Evidence and Conclusion Ontology (ECO, Chibucos et al. 2014). Under this scenario, the AISM would serve as one of the components required to assess homology statements across different taxonomic groups of insects.

Bringing Insects into the Phenomic Era

Nomenclatural rules require that the establishment of an animal taxon new to science "be accompanied by a description or definition that states in words characters that are purported to differentiate the taxon" or "by a bibliographic reference to such a published statement" (Article 13.1 in ICZN 1999). Taxonomists have described over one million species of insects worldwide (Stork 2018). These descriptions constitute vast amounts of information that could be efficiently mined, compared, interpreted, and analyzed, just like any large molecular data set nowadays. However, these phenotypic descriptions are presented in nonstandard natural language form, and are therefore, inaccessible for machine interpretation (Balhoff et al. 2010; Dahdul et al. 2010). Ontologies and knowledge graphs offer systems to represent entire knowledge domains in an organized, standardized, consistent, and logical manner, so that information can be processed and quality-checked by computers (Arp et al. 2015).

There are informatic tools that allow data extraction from the literature, based on XML markup, which has been used primarily for extracting taxonomic information from PDF files (Penev et al. 2011). For instance, GoldenGATE-Imagine (Sautter et al. 2007; https://github.com/plazi/GoldenGATE-Imagine), which is used by Plazi (http://plazi.org). There are also tools that use ontologies for annotating anatomical, phenotypic, and taxonomic data (Phenex; Balhoff et al. 2010). Lücking et al. (2021) provide an overview of methods for semantic annotation of bibliographic records and introduce a system to use multiple annotations for terms; the authors also introduce the BIOfid-portal (https://www.biofid.de/en/search/), which is an online tool for accessing the semantics of biodiversity texts in German. The annotation method proposed by Lücking et al. (2021) is partly based on the MATTER conceptual framework for annotations (Model, Annotate, Train and Test, Evaluate, and Revise; Pustejovsky and Stubbs 2012).

The AISM provides the key to annotating insect phenotypic information that is extracted from the literature. Combining or expanding these and similar informatics tools can generate large-scale phenotypic data sets, unlocking multiple avenues of research including, among others, genotype to phenotype associations, evo-devo studies, and the use of Artificial Intelligence and ontological inference (Jackson et al. 2018) to analyze morphological evolution across insects. Phenotypic data generated with the aid of the AISM would greatly contribute to increasing links in the Biodiversity Knowledge Graph (Page 2013).

Availability

The AISM is available on GitHub at https://github.com/insect-morphology/aism (Girón et al. 2023a) as well as on the OBO Foundry at http://www.obofoundry.org/ontology/aism.html. All the released versions of the AISM are archived via ZENODO (https://doi.org/10.5281/zenodo.4660322). The GitHub repository is open for collaborative editing. We provide a manual on how to edit the AISM (https://github.com/insect-morphology/Manual; Girón et al. 2021), including the use of templates, and how to use the AISM as the starting point for developing taxon-specific ontologies. The AISM is available as an OWL file at https://github.com/insect-morphology/aism/blob/master/aism.owl.

Conclusion

Here we provided a Model for Describing Cuticular Anatomical Structures (MoDCAS) that incorporates structural properties and topological relationships to define anatomical structures of cuticle-based systems, independent of developmental stage, homology assumptions, or taxonomic group. Following the set of principles established by MoDCAS, we created the first universally applicable anatomy ontology for insects, the Ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM). The AISM provides a basic backbone of generalized and unambiguously labeled and defined terms for the anatomy of the skeleto-muscular system of insects. Each term is accompanied by natural language definitions translated into sets of subclass of descriptors to provide logical definitions in the ontology. Built using the Ontology Development Kit, which is a free, open-source, and OBO Foundry-supported system, the AISM is interoperable with existing ontologies in the biological sciences, open for editing and refinement, and extensible to tackle arthropod taxon-specific ontologies.

The AISM opens new opportunities for phenomic-scale research in biology by providing computer-parsable formalization and a controlled vocabulary for insect anatomy. The potential application of MoDCAS and the AISM spans all biological domains, including phenotype comparison and description, and ontology-informed phylogenetic methods (Tarasov 2019).

Supplementary Data

Supplementary material, including the ontology files in owl and obo format, along with the template example can be found in the Dryad data repository (https://doi.org/10.5061/dryad.7sqv9s4w6).

Funding

Funding was provided by the NSF ARTS Program (DEB #1523605 and DEB #2009247 to Aaron Smith). Sergei Tarasov was supported by the Academy of Finland grant: 339576, and a three-year grant from the University of Helsinki. Markus Koch was supported by grants of the German Research Foundation (DFG-project MO 412/54-1) for the Specialized Information Service Biodiversity Research (BIOfid). Brendon Boudinot was supported by the Alexander von Humboldt Stiftung (Germany).

Conflict of interests

The authors declare no conflict of interest.

Acknowledgments

We are grateful to the Editor in Chief, Dr. Isabel Sanmartín, the Associate Editor, Dr. Lauren Esposito, and two anonymous reviewers for their comments and suggestions that contributed to improve the manuscript.

References

Adachi H., Matsuda K., Nishida K., Hanson P., Kondo S., Gotoh H. 2020. Structure and development of the complex helmet of treehoppers (Insecta: Hemiptera: Membracidae). Zool. Lett. 6(1):3. doi:10.1186/s40851-020-00155-7.

Adler P.N. 2017. Gene expression and morphogenesis during the deposition of *Drosophila* wing cuticle. Fly 11(3):194–199. doi:10.1080/19336934.2017.1295188.

Andersen S.O. 2012. 6 - Cuticular sclerotization and tanning. In: Gilbert L.I., editor. Insect molecular biology and biochemistry. Elsevier, Amsterdam, p. 167–192. doi:10.1016/B978-0-12-384747-8.10006-6.

Arp R, Smith B, Spear AD. 2015. Building ontologies with basic formal ontology. Cambridge, MA, USA: MIT Press, p. 248 doi:10.7551/mitpress/9780262527811.001.0001.

Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J. C., Richardson J.E., Ringwald M., Rubin G.M., Sherlock G. 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25(1):25–29. doi:10.1038/75556.

Balhoff J.P., Dahdul W.M., Kothari C.R., Lapp H., Lundberg J.G., Mabee P., Midford P.E., Westerfield M., Vision T.J. 2010. Phenex: ontological annotation of phenotypic diversity. PLoS One 5(5):e10500. doi:10.1371/journal.pone.0010500.

Balhoff J.P., Mikó I., Yoder M.J., Mullins P.L., Deans A.R. 2013. A semantic model for species description applied to the ensign wasps (Hymenoptera: Evaniidae) of New Caledonia. Syst. Biol. 62(5):639–659. doi:10.1093/sysbio/syt028.

Bertone M.A., Mikó I., Yoder M.J., Seltmann K.C., Balhoff J.P., Deans A.R. 2013. Matching arthropod anatomy ontologies to the Hymenoptera Anatomy Ontology: results from a manual alignment. Database 2013:bas057. doi:10.1093/database/bas057.

Beutel R.G., Friedrich F., Yang X.-K., Ge S.-Q. 2014. Insect morphology and phylogeny: a textbook for students of entomology. De Gruyter, p. 516. doi:10.1515/9783110264043.

Bitan A., Rosenbaum I., Abdu U. 2012. Stable and dynamic microtubules coordinately determine and maintain *Drosophila* bristle shape. Development 139:1987–1996. doi:10.1242/dev.076893.

Bitsch C., Bitsch J. 2002. The endoskeletal structures in arthropods: cytology, morphology and evolution. Arthropod Struct. Dev. 30:159–177. doi:10.1016/S1467-8039(01)00032-9.

Blumer M.J.F. 2021. Bone tissue and histological and molecular events during development of the long bones. Ann. Anat. 235:151704. doi:10.1016/j.aanat.2021.151704.

Bolshoy A., Lacková L. 2021. Illusions of linguistics and illusions of modern synthesis: two parallel stories. Biosemiotics 14:115–119. doi:10.1007/s12304-021-09429-9.

Chibucos M.C., Mungall C.J., Balakrishnan R., Christie K.R., Huntley R.P., White O., Blake J.A., Lewis S.E., Giglio M. 2014. Standardized description of scientific evidence using the Evidence Ontology (ECO). Database 2014:bau075. doi:10.1093/database/bau075.

Costa M., Reeve S., Grumbling G., Osumi-Sutherland D. 2013. The *Drosophila* anatomy ontology. J. Biomed. Semant. 4:32. doi:10.1186/2041-1480-4-32.

Dahdul W.M., Balhoff J.P., Engeman J., Grande T., Hilton E.J., Kothari C., Lapp H., Lundberg J.G., Midford P.E., Vision T.J., Westerfield M., Mabee P.M. 2010. Evolutionary characters, phenotypes and ontologies: curating data from the systematic biology literature. PLoS One 5(5):e10708. doi:10.1371/journal.pone.0010708.

Dahdul W.M., Cui H., Mabee P.M., Mungall C.J., Osumi-Sutherland D., Walls R.L., Haendel M.A. 2014. Nose to tail, roots to shoots: spatial descriptors for phenotypic diversity in the biological spatial ontology. J. Biomed. Semant. 5(1):34. doi:10.1186/2041-1480-5-34.

Dahdul W., Manda P., Cui H., Balhoff J.P., Dececchi T.A., Ibrahim N., Lapp H., Vision T., Mabee P.M. 2018. Annotation of phenotypes using ontologies: a gold standard for the training and evaluation of natural language processing systems. Database 2018:bay110. doi:10.1093/database/bay110.

Deans A.R., Mikó I., Wipfler B., Friedrich F. 2012a. Evolutionary phenomics and the emerging enlightenment of arthropod systematics. Invertebr. Syst 26:323–330. doi:10.1071/IS12063.

Deans A.R., Yoder M.J., Balhoff J.P. 2012b. Time to change how we describe biodiversity. Trends Ecol. Evol. 27:78–84. doi:10.1016/j.tree.2011.11.007

Deans A.R., Lewis S.E., Huala E., Anzaldo S.S., Ashburner M., Balhoff J.P., Blackburn D.C., Blake J.A., Burleigh J.G., Chanet B., Cooper L.D., Courtot M., Csösz S., Cui H., Dahdul W., Das S., Dececchi T.A., Dettai A., Diogo R., Druzinsky R.E., Dumontier M., Franz N.M., Friedrich F., Gkoutos G.V., Haendel M., Harmon L.J., Hayamizu T.F., He Y., Hines H.M., Ibrahim N., Jackson L.M., Jaiswal P., James-Zorn C., Köhler S., Lecointre G., Lapp H., Lawrence C.J., Novère N.L., Lundberg J.G., Macklin J., Mast A.R., Midford P.E., Mikó I., Mungall C.J., Oellrich A., Osumi-Sutherland D., Parkinson H., Ramírez M.J., Richter S., Robinson P.N., Ruttenberg A., Schulz K.S., Segerdell E., Seltmann K.C., Sharkey M.J., Smith A.D., Smith B., Specht C.D., Squires R.B., Thacker R.W., Thessen A., Fernandez-Triana J., Vihinen M., Vize P.D., Vogt L., Wall C.E., Walls R.L., Westerfeld M., Wharton R.A., Wirkner C.S., Woolley J.B., Yoder M.J., Zorn A.M., Mabee P. 2015. Finding our way through phenotypes. PLoS Biol. 13:e1002033. doi:10.1371/journal.pbio.1002033.

Denk-Lobnig M., Martin A.C. 2020. Divergent and combinatorial mechanical strategies that promote epithelial folding during morphogenesis. Curr. Opin Genet. Dev. 63:24–29. doi:10.1016/j.gde.2020.02.014.

de Oliveira V.P., Cruz-Landim C. 2003. Morphology and function of insect fat body cells: a review. Biociências 11:195–205.

Djokic S., Bakhrat A., Tsurim I., Urakova N., Rasgon J.L., Abdu U. 2020. Actin bundles play a different role in shaping scales compared to bristles in the mosquito *Aedes aegypti*. Sci. Rep-UK 10:14885. doi:10.1038/s41598-020-71911-0.

Dönitz J., Grossmann D., Schild I., Schmitt-Engel C., Bradler S., Prpic N.-M., Bucher G. 2013. TrOn: an anatomical ontology for the beetle *Tribolium castaneum*. PLoS One 8:e70695. doi:10.1371/journal.pone.0070695.

Girón J.C., González Montaña L.A., Mikó I. 2021. How to create AISM-based insect anatomy ontologies using the Ontology Development Kit (ODK) and edit them in Protégé. GitHub repository. doi:10.5281/zenodo.4642556.

Girón J.C., Lumen R., Smith A. 2023b. The Coleoptera Anatomy Ontology (COLAO), v2023-04-15. GitHub repository. doi:10.5281/zenodo.4777299.

Girón J.C., Mikó I., González Montaña L.A., Tarasov S., Matentzoglu N. 2023a. Ontology for the Anatomy of the Insect Skeleto-Muscular System (AISM), v2023-04-14. GitHub repository. doi:10.5281/zenodo.4660322.

Gkoutos G.V., Green E.C., Mallon A.-M., Hancock J.M., Davidson D. 2005. Using ontologies to describe mouse phenotypes. Genome Biol. 6(1):R8. doi:10.1186/gb-2004-6-1-r8.

González-Montaña L.A., Rueda-Ramírez D., Serna Cardona F.J., Gaigl A. 2023. An anatomical ontology for the Class Collembola (Arthropoda: Hexapoda). Braz. arch. biol. technol. 66:e23220682. doi:10.1590/1678-4324-2023220682.

Gotoh H., Adachi H., Matsuda K., Lavine L.C. 2021. Epithelial folding determines the final shape of beetle horns. Curr. Opin Genet. Dev. 69:122–128. doi:10.1016/j.gde.2021.03.003.

Haendel M.A., Neuhaus F., Osumi-Sutherland D., Mabee P.M., Mejino J.L.V., Mungall C.J., Smith B. 2008. CARO – the common anatomy reference ontology. In: Burger A, Davidson D, Baldock R, editors. Anatomy ontologies for bioinformatics: principles and practice. London: Springer. (Computational Biology). p. 327–349. doi:10.1007/978-1-84628-885-2_16.

Hall B.K. 1975. Evolutionary consequences of skeletal differentiation. Am. Zool. 15:329–350. doi:10.1093/icb/15.2.329.

Hannezo E., Prost J., Joanny J.-F. 2014. Theory of epithelial sheet morphology in three dimensions. P. Natl. Acad. Sci. USA 111:27–32. doi:10.1073/pnas.1312076111.

Hastings J., Owen G., Dekker A., Ennis M., Kale N., Muthukrishnan V., Turner S., Swainston N., Mendes P., Steinbeck C. 2016. ChEBI in 2016: improved services and an expanding collection of metabolites. Nucleic Acids Res. 44(D1):D1214–D1219. doi:10.1093/nar/gkv1031.

ICZN. 1999. International Code of Zoological Nomenclature. Fourth edition, adopted by the International Union of Biological Sciences. London: International Trust for Zoological Nomenclature.

Iyer J., Wang Q., Le T., Pizzo L., Grönke S., Ambegaokar S.S., Imai Y., Srivastava A., Troisí B.L., Mardon G., et al. 2016. Quantitative assessment of eye phenotypes for functional genetic studies using Drosophila melanogaster. G3 Genes Genom. Genet. 6(5):1427–1437. doi:10.1534/g3.116.027060.

Jackson R.C., Balhoff J.P., Douglass E., Harris N.L., Mungall C.J., Overton J.A. 2019. ROBOT: a tool for automating ontology workflows. BMC Bioinf. 20(1):407. doi:10.1186/s12859-019-3002-3.

Jackson L.M., Fernando P.C., Hanscom J.S., Balhoff J.P., Mabee P.M. 2018. Automated integration of trees and traits: a case study using paired fin loss across teleost fishes. Syst. Biol. 67:559–575. doi:10.1093/sysbio/syx098.

Jan S., Liu S., Hafeez M., Zhang X., Dawar F.U., Guo J., Gao C., Wang M. 2017. Isolation and functional identification of three cuticle protein genes during metamorphosis of the beet armyworm, *Spodoptera exigua*. Sci. Rep.-UK 7:16061. doi:10.1038/s41598-017-16435-w.

Klass K.-D. 2008. The female abdomen of ovipositor-bearing Odonata (Insecta: Pterygota). Arthropod Syst. Phylo. 66:45–142. https://www.zobodat.at/pdf/Arthropod-Systematics-Phylogeny_66_0045-0142.pdf.

Klass K.-D., Matushkina N.A. 2012. The exoskeleton of the female genitalic region in *Petrobiellus takunagae* (Insecta: Archaeognatha): insect-wide terminology, homologies, and functional interpretations. Arthropod Struct. Dev. 41:575–591. doi:10.1016/j.asd.2012.06.003.

Kluge N. 2004. The phylogenetic system of Ephemeroptera. Springer Science & Business Media: p. 442. https://link.springer.com/book/10.1007/978-94-007-0872-3.

Lees A.D., Waddington C. 1942. The development of the bristles in normal and some mutant types of *Drosophila melanogaster*. P. Roy. Soc. Lond. B Bio. 131(862):87–110. doi:10.1098/rspb.1942.0019.

Loesel R., Wolf H., Kenning M., Harzsch S., Sombke A. 2013. Architectural principles and evolution of the arthropod central nervous system. In: Minelli A., Boxshall G., Fusco G. editors. Arthropod biology and evolution: molecules, development, morphology. Berlin, Heidelberg: Springer, p. 299–342.

Lücking A., Driller C., Stoeckel M., Abrami G., Pachzelt A., Mehler A. 2021. Multiple annotation for biodiversity: developing an annotation framework among biology, linguistics and text technology. Lang. Resour. Eval. 56:807–855. doi:10.1007/s10579-021-09553-5.

Mabee P.M., Balhoff J.P., Dahdul W.M., Lapp H., Mungall C.J., Vision T.J. 2020. A logical model of homology for comparative biology. Syst. Biol. 69(2):345–362. doi:10.1093/sysbio/syz067.

Matentzoglu N., Goutte-Gattat D., Tan S.Z.K., Balhoff J.P., Carbon S., Caron A.R., Duncan W.D., Flack J.E., Haendel M., Harris N.L., Hogan W.R., Hoyt C.T., Jackson R.C., Kim H., Kir H., Larralde M.,

McMurry J.A., Overton J.A., Peters B., Pilgrim C., Stefancsik R., Robb S.M., Toro S., Vasilevsky N.A., Walls R., Mungall C.J., Osumi-Sutherland D. 2022. Ontology Development Kit: a toolkit for building, maintaining and standardizing biomedical ontologies. Database 2022:baac087. doi:10.1093/database/baac087.
Midford P.E., Dececchi T.A., Balhoff J.P., Dahdul W.M., Ibrahim N., Lapp H., Lundberg J.G., Mabee P.M., Sereno P.C., Westerfield M., Vision T.J., Blackburn David C. 2013. The vertebrate taxonomy ontology: a framework for reasoning across model organism and species phenotypes. J. Biomed. Semant. 4(1):34. doi:10.1186/2041-1480-4-34.
Mikó I., Masner L., Ulmer J.M., Raymond M., Hobbie J., Tarasov S., Margaría C.B., Seltmann K.C., Talamas E.J. 2021. A semantically enriched taxonomic revision of *Gryonoides* Dodd, 1920 (Hymenoptera, Scelionidae), with a review of the hosts of Teleasinae. J. Hymenopt. Res 87:523–573. doi:10.3897/jhr.87.72931.
Minelli A, Fusco G. 2013. Homology. In: Kampourakis K., editor. The philosophy of biology: a companion for educators. History, philosophy and theory of the life sciences. Netherlands, Dordrecht: Springer, p. 289–322. doi:10.1007/978-94-007-6537-5_15.
Monteiro E.C., Tamaki F.K., Terra W.R., Ribeiro A.F. 2014. The digestive system of the "stick bug" *Cladomorphus phyllinus* (Phasmida, Phasmatidae): a morphological, physiological and biochemical analysis. Arthropod Struct. Dev. 43:123–134. doi:10.1016/j.asd.2013.11.005.
Mullins P., Kawada R., Balhoff J., Deans A. 2012. A revision of *Evaniscus* (Hymenoptera, Evaniidae) using ontology-based semantic phenotype annotation. ZooKeys 223:1–38. doi:10.3897/zookeys.223.3572.
Mungall C., Matentzoglu N., Balhoff J., Osumi-Sutherland D., Gaudet P., Tan S., Duncan D., Pilgrim C., Overton J.A., Hoyt C.T., Chan L., Harris N., Moxon S., Vasilevsky N., Toro S., Brush M., Touré V., Caron A., Goutte-Gattat D., Sinclair M., Bretaudeau A., Cain S., Haendel M., Zhang B., Hammock J., Laporte M.-A. 2023. oborel/obo-relations: v2023-02-22. GitHub Reposit. doi:10.5281/zenodo.593101.
Mungall C.J., Torniai C., Gkoutos G.V., Lewis S.E., Haendel M.A. 2012. Uberon, an integrative multi-species anatomy ontology. Genome Biol. 13:R5. doi:10.1186/gb-2012-13-1-r5.
Musen M.A.; Protégé Team. 2015. The Protégé project: a look back and a look forward. AI Matters 1:4–12. doi:10.1145/2757001.2757003.
Natale D.A., Arighi C.N., Blake J.A., Bona J., Chen C., Chen S.-C., Christie K.R., Cowart J., D'Eustachio P., Diehl A.D., Drabkin H.J., Duncan W.D., Huang H., Ren J., Ross K., Ruttenberg A., Shamovsky V., Smith B., Wang Q., Zhang J., El-Sayed A., Wu C.H. 2017. Protein Ontology (PRO): enhancing and scaling up the representation of protein entities. Nucleic Acids Res. 45:D339–D346. doi:10.1093/nar/gkw1075.
Osumi-Sutherland D., Marygold S.J., Millburn G.H., McQuilton P.A., Ponting L., Stefancsik R., Falls K., Brown N.H., Gkoutos G.V. 2013. The *Drosophila* phenotype ontology. Biomed. Semant. 4(1):30. doi:10.1186/2041-1480-4-30.
Page R.D.M. 2013. BioNames: linking taxonomy, texts, and trees. PeerJ 1:e190. doi:10.7717/peerj.190.
Page R.E., Amdam G.V. 2007. The making of a social insect: developmental architectures of social design. Bioessays 29:334–343. doi:10.1002/bies.20549.
Penev L., Lyal C., Weitzman A., Morse D., King D., Sautter G., Georgiev T., Morris R., Catapano T., Agosti D. 2011. XML schemas and mark-up practices of taxonomic literature. ZooKeys 150:89–116. doi:10.3897/zookeys.150.2213.
Pustejovsky J, Stubbs A. 2012. Natural language annotation for machine learning: a guide to corpus-building for applications. O'Reilly Media Inc., p. 342 https://www.google.com/books/edition/Natural_Language_Annotation_for_Machine/A57TS7fs8MUC?hl=en&gbpv=0.
Ramírez M.J., Michalik P. 2019. The spider anatomy ontology (SPD)—a versatile tool to link anatomy with cross-disciplinary data. Diversity 11(202):2021–2029. doi:10.3390/d11100202.
Richards A.G., Richards P.A. 1979. The cuticular protuberances of insects. Int. J. Insect Morph. 8:143–157. doi:10.1016/0020-7322(79)90013-8.
Sautter G., Böhm K., Agosti D. 2007. Semi-automated xml markup of biosystematic legacy literature with the GoldenGATE editor. Pac. Symp. Biocomput. 12:391–402. doi:10.1142/9789812772435_0037.
Smith B., Ashburner M., Rosse C., Bard J., Bug W., Ceusters W., Goldberg L.J., Eilbeck K., Ireland A., Mungall C.J., et al. 2007. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat. Biotech. 25(11):1251–1255. doi:10.1038/nbt1346.
Smith B., Ceusters W., Klagges B., Köhler J., Kumar A., Lomax J., Mungall C., Neuhaus F., Rector A.L., Rosse C. 2005. Relations in biomedical ontologies. Genome Biol. 6:R46. doi:10.1186/gb-2005-6-5-r46.
Snodgrass R.E. 1951. Anatomy and morphology. J. New York Entomol. Soc. 59:71–73. https://www.jstor.org/stable/25005398.
Snodgrass R.E. 1963. A contribution toward an encyclopedia of insect anatomy. Smithsonian Misc. Collect. 146:1–48. https://www.biodiversitylibrary.org/page/30259679.
Snodgrass R.E. 1993. Principles of Insect Morphology. Cornell University Press, p. 768 https://www.jstor.org/stable/10.7591/j.ctv1nhm1j.
Sommer R.J. 2020. Phenotypic plasticity: from theory and genetics to current and future challenges. Genetics 215(1):1–13. doi:10.1534/genetics.120.303163.
Spear A.D., Ceusters W., Smith B. 2016. Functions in basic formal ontology. Appl. Ontol. 11(2):103–128. doi:10.3233/AO-160164.
Stork N.E. 2018. How many species of insects and other terrestrial arthropods are there on Earth? Annu. Rev. Entomol. 63(1):31–45. doi:10.1146/annurev-ento-020117-043348.
Stucky B., Balhoff J., Barve N., Barve V., Brenskelle L., Brush M., Dahlem G., Gilbert J., Kawahara A., Keller O., et al. 2019. Developing a vocabulary and ontology for modeling insect natural history data: example data, use cases, and competency questions. BDJ 7:e33303. doi:10.3897/BDJ.7.e33303.
Tajiri R. 2017. Cuticle itself as a central and dynamic player in shaping cuticle. Curr. Opin. Insect Sci. 19:30–35. doi:10.1016/j.cois.2016.10.009.
Tarasov S. 2019. Integration of anatomy ontologies and evo-devo using structured Markov Models suggests a new framework for modeling discrete phenotypic traits. Syst. Biol. 68(5):698–716. doi:10.1093/sysbio/syz005.
Topalis P., Tzavlaki C., Vestaki K., Dialynas E., Sonenshine D.E., Butler R., Bruggner R.V., Stinson E.O., Collins F.H., Louis C. 2008. Anatomical ontologies of mosquitoes and ticks, and their web browsers in VectorBase. Insect. Mol. Biol. 17(1):87–89. doi:10.1111/j.1365-2583.2008.00781.x.
Tuxen SL. 1970. Taxonomist's glossary of genitalia in insects (2nd ed.), Copenhagen: Ejnar Munksgaard, p. 359. http://antbase.org/ants/publications/21060/21060.pdf.
Vogt L. 2008. Learning from Linnaeus: towards developing the foundation for a general structure concept for morphology. Zootaxa 1950:123–152. doi:10.11646/zootaxa.1950.1.11.
Vogt L. 2017. Assessing similarity: on homology, characters and the need for a semantic approach to non-evolutionary comparative homology. Cladistics 33(5):513–539. doi:10.1111/cla.12179.
Vogt L. 2019. Levels and building blocks—toward a domain granularity framework for the life sciences. J. Biomed. Semant. 10:4. doi:10.1186/s13326-019-0196-2.
Vogt L., Bartolomaeus T., Giribet G. 2010. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. Cladistics 26(3):301–325. doi:10.1111/j.1096-0031.2009.00286.x.
Wagner GP. 2014. Homology, genes, and evolutionary innovation. NJ, U.S.: Princeton , p. 496 https://press.princeton.edu/books/hardcover/9780691156460/homology-genes-and-evolutionary-innovation.
Wang M., Li Z., Zhou Z. 2017. Insight into the growth pattern and bone fusion of basal birds from an Early Cretaceous enantiornithine bird. P. Natl. Acad. Sci. U.S.A. 114(43):11470–11475. doi:10.1073/pnas.1707237114.
Washington N.L., Haendel M.A., Mungall C.J., Ashburner M., Westerfield M., Lewis S.E. 2009. Linking human diseases to animal models using ontology-based phenotype annotation. PLoS Biol. 7:e1000247. doi:10.1371/journal.pbio.1000247.
Wirkner C.S., Göpel T., Runge J., Keiler J., Klussmann-Fricke B.-J., Huckstorf K., Scholz S., Mikó I., Yoder M J., Richter S. 2017.

The first organ-based ontology for arthropods (Ontology of Arthropod Circulatory Systems—OArCS) and its integration into a novel formalization scheme for morphological descriptions. Syst. Biol. 66(5):754–768. doi:10.1093/sysbio/syw108.
Yoder M.J., Mikó I., Seltmann K.C., Bertone M.A., Deans A.R. 2010. A gross anatomy ontology for Hymenoptera. PLoS One 5:e15991. doi:10.1371/journal.pone.0015991.
Yoshizawa K. 2005. Morphology of Psocomorpha (Psocodea: 'Psocoptera'). Insecta matsumurana 62:1–44. http://hdl.handle.net/2115/10524.
Young B.A. 1993. On the necessity of an archetypal concept in morphology: with special reference to the concepts of "structure" and "homology". Biol. Philos. 8(2):225–248. doi:10.1007/BF00850483.
Zartman J.J., Shvartsman S.Y. 2010. Unit operations of tissue development: epithelial folding. Annu. Rev. Chem. Biomol. 1(1):231–246. doi:10.1146/annurev-chembioeng-073009-100919.
Zhao X., Gou X., Qin Z., Li D., Wang Y., Ma E., Li S., Zhang J. 2017. Identification and expression of cuticular protein genes based on *Locusta migratoria* transcriptome. Sci. Rep.-U.K. 7:45462. doi:10.1038/srep45462.