A generic organ based ontology system, applied to vertebrate heart anatomy, development and physiology

Laura M.F. Bertens†, Joris Slob†, Fons J. Verbeek*

Imaging & Bioinformatics, LIACS, Leiden University, The Netherlands
† Both authors contributed equally

Summary

We present a novel approach to modelling biological information using ontologies. The system interlinks three ontologies, comprising anatomical, developmental and taxonomical information, and includes instances of structures for different species. The framework is constructed for comparative analyses in the field of evolutionary development. We have applied the approach to the vertebrate heart and present four case studies of the functionality of the system, focusing on cross-species comparisons, developmental studies, physiological studies and 3D visualisation.

1   Introduction

Over the last few years ontologies have become increasingly important in biology as a formal basis for collecting and sharing knowledge. They are used in several fields, such as genetics, biochemistry and anatomy. A vast amount of literature can be found on the use of ontologies in biological sciences, ranging from overview articles like [1] and [2] to textbooks, such as [3],[4],[5]). Journals such as *Computers in biology and medicine* have devoted special issues to bio-ontologies (issues 7-8, vol. 36, 2006) and the well-established journal *Bioinformatics* dedicates a permanent section to databases and ontologies. A collaboration, the OBO foundry, has been set up to bring ontologies together in biological and biomedical fields ([1],[6]). This foundry contains a number of anatomy ontologies for different species, *e.g.* the Foundational Model of Anatomy (FMA, [7],[8]) for humans, the Mouse Gross Anatomy and Development ontology (EMAP, [9]), the Amphibian Gross Anatomy (AAO, [10]) and the Zebrafish Anatomy and Development ontology (ZFA, [11]). Each of these ontologies describes the anatomy of the particular species in great detail and as such is very valuable. However, since they all use different class hierarchies to describe the anatomy, comparable studies between species are often hampered. Therefore we propose an additional approach, which combines the anatomical knowledge of multiple species in one generic class hierarchy. As a case study we use the vertebrate heart. Put another way, the single species ontologies can be described as following a vertical approach (see Fig. 1), modelling the entire anatomy of one species. In contrast we use a horizontal approach, modelling the anatomy of one organ system, the heart, for a wide range of species, the subphylum of vertebrates.

The collected knowledge is structured in a system of three ontologies, one comprising all anatomical structures (referred to as the anatomy ontology), another providing developmental staging information (the development ontology) and thirdly the taxonomical species database of the National Center for Biotechnology Information (the NCBI, [12]). In addition to the ontologies a set of instances is used to capture the information for specific species. By founding the class hierarchy of our anatomy ontology on the basic biological distinctions (such as atrial versus ventricular) as opposed to species-specific distinctions (such as right

---

* To whom correspondence should be addressed. Email: fverbeek@liacs.nl

versus left heart chambers) we ensure that any vertebrate species can be added to the system without losing the consistency of the anatomical class hierarchy.
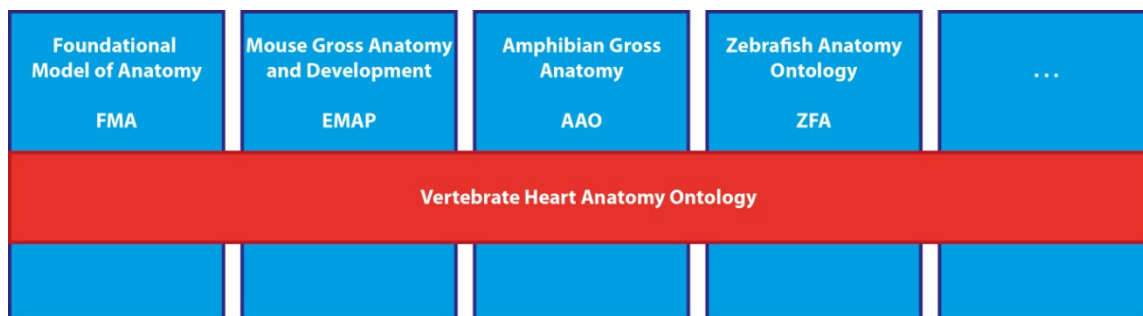

Figure 1: Illustration of the vertical approach, adopted by most anatomy ontologies in the OBO Foundry, and the horizontal approach presented here.

By using these three ontologies, a generic framework is created in which anatomical, physiological and developmental information for different species can be brought together, using instances for structures in different species. This makes the system very suitable for comparative studies in the field of evolutionary development.

In the remainder of the paper, we will provide relevant biological background information regarding anatomy, physiology and development, used to construct the knowledge base. Subsequently we will lay out the architecture of the system, discussing the different ontologies, properties and the use of instances. The ontologies are provided on our website; the reader can download and use the ontologies to include his or her own data or to study and query the content (using an ontology editor such as Protégé and/or SPARQL queries). Furthermore an interface is being developed (see section 4.4) which allows the user to query the system from our website, without the need for an ontology editor. The ontology system allows for different types of queries and we describe four of these as case studies in using the system. They concern cross-species comparisons, developmental questions, context-dependent queries about the physiology, using SWRL rules [13] and 3D visualisation of query results from interface queries. Future additions to the functionality of the system and extensions of the information contained in it will be discussed in the last chapter. Throughout the paper, whenever we speak of structures, we do so in a biological sense; the terms classes and instances refer to the elements in the ontology.

2   Biological background information

The vertebrate heart is a highly diverse structure and although the basic developmental plan is similar for the entire subphylum, the anatomy can vary quite extensively between species. This and the vast body of literature on the topic make it a suitable case study for our horizontal approach. Through our ontology we aim to provide a framework that can be used to bring together anatomical information for vertebrate species. For this reason three anatomically diverse species have been annotated first in the system: the zebrafish (*Danio rerio*), the European pond terrapin (*Emys orbicularis*) and the saltwater crocodile (*Crocodylus porosus*). Currently mammalian and avian species are being added to the ontology. Here we will briefly describe the cardiac anatomy and relevant physiology of these species, followed by a short description of developmental staging as used in biology.

2.1   Anatomy

In all vertebrates the heart arises as a tube. During development two distinct structures arise in this tube, the atrial and the ventricular chambers. The basic heart can therefore be said to be

made up out of four elements: the inflow tract (where blood enters the atria), the atrial compartment, the ventricular compartment and the outflow tract (where blood leaves the ventricle). The classification in the ontology is built on this basic lay-out.
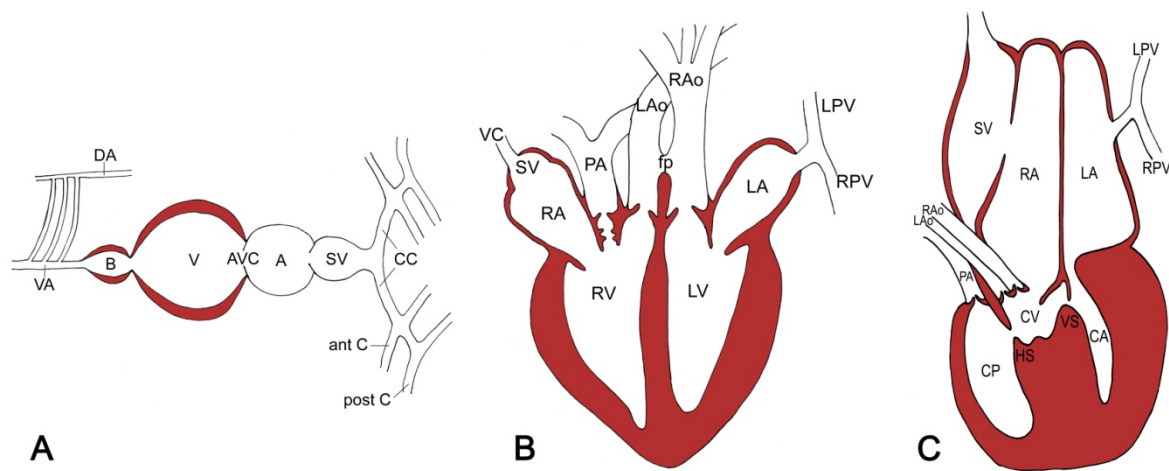

Figure 2: Schematic drawings of the hearts of (A) zebrafish, (B) crocodiles and (C) non-crocodilian reptiles. A, atrium; ant C, anterior cardinal; AVC, atrioventricular canal; B, bulbus; CA, cavum arteriosum; CC, common cardinal; DA, dorsal aorta; fp, foramen of panizza; LA, left atrium; LAo, left aorta; LPV, left pulmonary vein; LV, left ventricle; PA, pulmonary artery; post C, posterior cardinal; RA, right atrium; RAo, right aorta; RPV, right pulmonary vein; RV, right ventricle; SV, sinus venosus; V, ventricle; VA, ventral aorta; VC, vena cava.

In fish this anatomical plan stays essentially the same; the fish heart has one atrium and one ventricle, connected through the atrio-ventricular canal (the AV canal) and guarded by atrio-ventricular valves (see Fig. 2a). Blood enters the atrium through the sinus venosus and leaves the ventricle through the bulbus. In higher vertebrates the heart tube develops into a multi-chambered heart through looping and further dividing. In mammals and birds looping and septation lead to a four-chambered heart, with two atria and two ventricles. In reptiles the heart contains two atria and one or two ventricles; the crocodilian order has two completely separated ventricles, divided by an interventricular septum (resembling the mammalian heart but differing significantly in its physiology), whereas the other three reptilian orders have one ventricle, partially divided by two septa into three cavities (see Fig. 2b and c). As case studies for the generic use of our ontology we have chosen, in addition to the zebrafish, two reptile species representing these two different heart types: the *Crocodylus porosus* of the Crocodilian order and the turtle *Emys orbicularis* of the Testudine order. These two reptile species have very distinct heart structures, not found in other types of vertebrates. As shown in section 3 the classification of the ontology allows users to add these structures without losing consistency in the classification.

2.2   Physiology: shunting

In addition to the distinctive anatomy of these hearts the two reptilian species have also been chosen as case studies here for their specific physiology. In both the crocodile and the turtle heart bloodflow can be regulated for different physiological situations, according to pressure changes in the different heart compartments ([14],[15],[16],[17],[18],[19],[20],[21]). We will focus first on the crocodile as a case study for context-dependent querying (as explained in section 4.3); this animal is unique in having a completely separated ventricle while at the same time retaining the ability to adapt its bloodflow to differing circumstances through shunting ([16],[22],[14]]). In addition to the pulmonary artery, the right ventricle also opens

into a right aorta, which communicates with the left aorta through the foramen of Panizza ([17],[20]), as can be seen in Fig. 2b (structure *fp*). When pressures in the ventricles and aortae change, bloodflow can be redirected from the pulmonary to the systemic circuit, causing a pulmonary-to-systemic shunt (a right-to-left or R-L shunt) and thereby allowing blood from the right ventricle to bypass the lungs and be recirculated into the systemic circulation [16]. It is because of this ability to both separate bloodflows completely and establish shunts, that Axelsson has described the crocodile heart as probably the most functionally sophisticated found among vertebrates [14]; this makes it an excellent case study for the physiological functionality of our ontology.

Although we focus on the crocodile in presenting the dynamic blood pressure querying of our ontology, this functionality is not limited to this species and for this function again the structuring of information in the ontology is kept generic and is therefore applicable to all vertebrates in which shunting is present. The turtle has a very intricate shunting system, including both R-L and L-R shunts, for which information will be added to the ontology. Furthermore, shunting plays an important role during development in for instance the human embryo (between left and right atria through the foramen ovale and between pulmonary artery and aorta through the ductus arteriosus). Since the modeling of shunting in the ontology is generic, shunting via these embryological structures can also be modeled within this framework.

2.3   Development

We have chosen to include, in addition to anatomical structures present in the adult heart, developmental structures as well. These are often transient, arising during a particular developmental stage and disappearing (or merging with other structures) at a later stage. For many species, in particular model species such as zebrafish and mouse for which extensive knowledge is available, conventional staging systems exist; *e.g.* for zebrafish Kimmel *et al.* [23] have provided names and descriptions of all developmental stages, which are used by scientists all over the world to refer to zebrafish development. Similar stage descriptions have for instance been provided by Yntema [24] for reptiles, Theiler [25] for mouse, Carnegie [26] for humans and Hamburger and Hamilton [27] for chick. These staging systems are specific for particular species (or related groups of species) and we include in our development ontology (described in section 3.3) the staging systems for the main vertebrate model animals. In addition to these staging systems other features are often also used when discussing development; development can be measured in absolute time, *i.e.* hours post fertilization. Alternatively developmental landmarks can be used, structures that arise at a certain point during development and are used as indicators of developmental progress. For vertebrates the number of somites is often used to this end. As described in section 3.3 we use several of these features in our development ontology.

3   Architecture of the system

The system is build up out of four components: three ontologies and a set of instances. In this section we describe these components, starting with the ontology of anatomical structures, followed by a description of its properties and subsequently the ontology of developmental stages. For both ontologies we will explain the biological reasoning behind choosing the classes and properties. Lastly the use of instances is described, including the way these are linked to the first two ontologies as well as to the NCBI database of taxonomy.

3.1   The anatomy ontology

We have strived to make the anatomy ontology generically applicable to all vertebrate species. On the highest level in the classification tree the main division is in the superclasses cavitated structures, embryological structures, heart valves, heart septa and tissue types (see Fig. 3, presenting a screen dump of the ontology in Protégé).
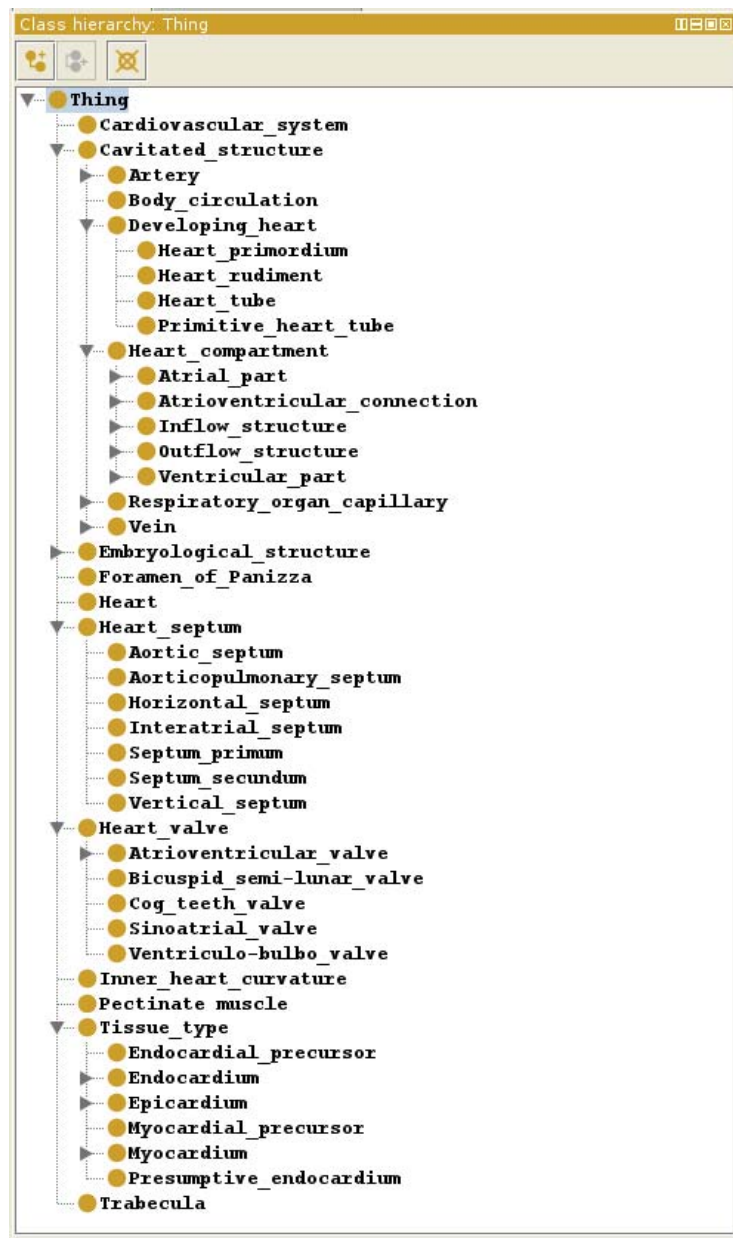

Figure 3: Screen dump of the partially unfolded anatomy ontology in Protégé, showing the main superclasses and several further subdivisions of these; arrows pointing downward indicate the subclasses of a class are shown, arrows pointing to the right indicate the subdivision is not shown.

The cavitated structures comprise all anatomical structures that contain a lumen, through which blood can flow, which means that bloodflow is located entirely in this superclass. The class of embryological structures comprises all structures that are present during development, but are no longer present in the adult animal. Since some of the embryological structures can also be categorised under one of the other superclasses (*e.g.* the common atrium is an embryological structure as well as an atrial structure), these classes have been assigned two superclasses and can be found under both. For classes which can have both embryological and

adult instances for different species (*e.g.* the sinus venosus, which is a transient structure in humans, but an adult structure in zebrafish), the class itself is not linked to the embryological superclass, but the specific instances are. Having a separate superclass for embryological structures enables the querying of only adult structures for a particular species, by specifically leaving out all structures categorised as embryological. Similarly, the adult structures can be disregarded, in case only structures from development are relevant to a query. Information about developmental staging is also captured in the development ontology (as described below), but by adding the superclass 'embryological structure' to the anatomy ontology we ensure that even when quantitative staging information is not known or included it is still possible to separate the structures qualitatively; furthermore, it means that this information is also available when only the anatomy ontology is used, without linking this to the development ontology.

Heart septa and valves both comprise anatomical structures in the heart which (partially) separate cavitated structures (heart compartments and blood vessels), thereby directing bloodflow. The final superclass in the first tree division is of tissue types and stands apart from the other structures in the sense that these classes are no specific anatomical structures in their own right, but indicate the cell types of which anatomical structures consist. As such they have been assigned their own superclass, next to the anatomical superclasses, allowing the user to link classes and instances to both the relevant anatomical structure and the tissue type.

On the next level in the hierarchy the majority of cavitated structures get divided into arteries, veins and heart compartments, thereby separating the heart proper from the connected blood vessels (which could be extended to include the entire circulation if so required). Subsequently the heart compartments are divided into inflow, atrial, ventricular, atrioventricular connecting and outflow structures. This division directly reflects the basic anatomical plan of the vertebrate heart and allows us to add structures of even highly diverse vertebrate species in a consistent system. For instance, even though not all vertebrates have the same number of atrial and ventricular chambers, all have at least one atrial and one ventricular structure. The three species currently annotated in the system illustrate this issue: while the zebrafish has only one atrium and one ventricle, the crocodile has two of both and the turtle has two atria and one ventricle, partially divided into three cavities. In this sense the ontology differs from the main ontologies in OBO, which focus on the entire anatomy of one species.

Note that general information, applicable to all vertebrates, is added in classes; species-specific information however, is added in instances, as descibed below. A clear illustration of the importance of this difference is given by the structures in the outflow tract: the outflow tract comprises structures as the bulbus, conus and truncus. These terms are used in diverse ways in different species and by using instances to add the species-specific information in our system we avoid confusion about the definitions; the classes in the anatomy ontology are general, which means we do not have to include different classes of for instance the bulbus for the different uses of the term in different animals. Each instance on the other hand is specific for one species and therefore only has one exact definition, preventing ambiguities. The same applies to the inflow structures, of which for instance the sinus venosus can be either an embryological structure (as in humans), or an adult structure (as in zebrafish).

Further subdivision of the hierarchy follows the same method: at each step down a level in a tree superclasses are chosen in such a way that each vertebrate anatomical class can be categorised in a consistent way. The ontologies (provided as OWL files) as well as additional information about the entire ontology system is available at http://bio-imaging.liacs.nl/ontologies/ontologies.html.

Besides being directly based on biological distinctions, the hierarchy in our anatomy ontology is also compatible with the class structuring of CARO [29]. CARO is an upper-ontology for anatomy terms. Top-level classes of our anatomy ontology can be placed under CARO classes; the heart is a subclass of the CARO class 'cavitated compound organ' and nearly all of our structures can be placed as sub-classes under the CARO class 'compound organ component'. This is a relevant feature for future integration and interoperability purposes.

3.2　　Properties used in the anatomy ontology

Apart from the *is_a* relationship with which the class hierarchy is built, information about the relationships between the different classes and instances is provided through a set of properties. These are either object properties, linking classes or instances together, or data properties, assigning pieces of data to classes or instances, *e.g.* integers, strings, booleans or floats. The properties specify spatial, developmental, taxonomical and physiological information.

To include information about the spatial relationship between structures we use the object properties *part_of* and *bounds*; in *part_of* a distinction is made between *constitutional_part_of* and *regional_part_of*, as in the FMA ([7],[8][30]). The property *constitutional_part_of* indicates that a structure is a fundamental element of a bigger structure; the parts are not just positioned in the bigger structure, but that structure consists of the parts (*e.g.* the atrial and ventricular parts of the heart). The property *regional_part_of*, on the other hand, signifies that a structure is positioned in another structure, *e.g.* the atrioventricular valves in the atrioventricular canal; without the valves the canal is still a valid structure. Furthermore, the property *bounds* denotes that a structure is a boundary of another structure; this is particularly relevant for valves and septa.

Physiological information concerning bloodflow is provided through a set of properties. Some of these are object properties, specifying which structures receive blood from which other structures. This can change for different physiological situations by means of bloodshunts, as described above, and we use a hierarchy of several properties to enable modelling each of these situations. In addition to the object properties we use a set of data properties, with which the common blood pressure values can be stored, for each of the cavitated structures and again for different physiological states. The implementation of these properties and their functionality, based mainly on inheritance rules, will be discussed in section 4.3.

Taxonomical information is provided using the property *structure_from_species*; this property links instances of anatomical structures to species codes in the taxonomical database of the NCBI [12]. We chose to make use of this comprehensive, well established database, as opposed to including classes for the species in our system in the anatomy ontology itself; this way we incorporate the knowledge gathered in the NCBI database, optimizing the functionality of our system for comparative studies in the field of evolutionary development. Since the NCBI codes are widely accepted, using these makes our system easily accessible and amenable to interoperability approaches; *i.e.* no knowledge is needed of the naming convention in the anatomy ontology itself, instances of particular species can be found using the property *structure_from_species* with the corresponding NCBI code. Furthermore, by using this external database we maintain a division between the generic knowledge in the anatomy ontology and the species specific information, preserving the generic nature of the anatomy framework.

Two types of developmental information are included through object properties; the property *develops_from* links an anatomical structure to the embryological structure that it developed from. Secondly, information about the developmental stage of structures is included using the

properties *end_stage* and *start_stage*. The uses of these are explained in the next paragraph, which describes the structure of the development ontology.

3.3 The development ontology and its properties

In several anatomy ontologies, such as the ontology of Human Developmental Anatomy, the EHDA [31], the hierarchy starts with a division in developmental stages. The anatomical structures that are present at that time point are added to the stage structure, using a *part_of* relation. Although this gives a good overview of the entire anatomy for each stage, it means that the same biological structure exists in the ontology as many times as there are stages at which it is present. Not only is this potentially confusing, it also increases the size of the ontology dramatically and forces the developer to add all relevant properties for each of the versions of the biological structure.

An alternative solution, adopted here, is to use developmental stages and singular versions of the anatomical structures and to use properties denoting the start stage and end stage for each of the anatomical structures. This way, querying for a particular developmental stage remains possible, without having to include a separate version of a biological structure for each of the stages it is present in. Some ontologies, such as the ZFA, include the developmental stages in the anatomy ontology itself. We have opted to build a separate ontology containing the developmental stages. Whereas the anatomical classes are generically applicable, the instances and developmental stage classes are specific to individual species. By setting these apart from the anatomy ontology we separate general information from species-specific information, leaving the framework generically applicable. Furthermore the development ontology serves a purpose in its own right by bringing together and relating staging information for diverse species (as described below) and can be used independently from the anatomy ontology. Lastly, the division allows the freedom to use a different external ontology for the developmental stages, if so required.

Instances of classes in the anatomy ontology can be linked to developmental stages in the external ontology by means of object properties in the anatomy ontology, *start_stage* and *end_stage*. Since staging information is not always exactly known for all structures the start and end stage properties both have two additional properties (*_at_earliest* and *_at_latest*) serving as a *terminus post quem* and *terminus ante quem*, which enables inclusion even of inexact staging information. Limiting contributors to the ontology system to providing only exact staging properties results in loss of information, since some structures are only known to appear or disappear after or before a certain stage. Using these additional properties developmental ranges for the presence of structures can be obtained.

As discussed above, each species or group of related species has its own staging system. We include in the development ontology staging systems for the major vertebrate model species. In some cases information is available to map these systems on each other [32] and in those cases we have done so with the property *corresponds_to*; note that this property can also be used when only partial mapping information (only for certain stages) is available. However, since developmental processes differ between species (*i.e.* in some species some structures may arise earlier while others arise later) it is often difficult to exactly relate the stages from different staging systems to eachother. We therefore also include other features to enable comparison of development between species. The absolute developmental time for the start of a developmental stage, in hours post fertilization, is included using *starts_at_hour*, as well as a relative measure of development in percentages (normalizing time from fertilization to adulthood), using *percentage_of_development* [33]. Furthermore the number of somites is provided, which is a well established means of assessing vertebrate development, using *somite_count*. In other words we include, in addition to the detailed species specific staging,

an absolute time measure, a relative time measure and a time-independent measure using landmarks.

3.4 Instances for specific species

We have designed the anatomy ontology as a generic framework, meant to be applicable to all vertebrates as an aid for annotation. The classes therefore do not contain species-specific information. However, we also set out to collect and include information for a number of vertebrate species, either because of their uses as a model organisms or because of their particular heart anatomy. In doing so we aim to present an ontology system with a predefined set of data, which can be used for comparative studies with data on other species or developmental stages. The anatomical data for the specific species are added to the ontology using instances, known in Protégé as individuals. In other anatomy ontologies dealing with one particular species, the classes are by default species-specific and instances are not used to differentiate between species.

Whenever an instance of a particular class is created, the instance is given the same name as the class, preceded by a two letter code; this code indicates the species to which the instance belongs, *e.g.* DR for a structure of the species *Danio rerio* and EO for *Emys orbicularis*. In this way the class 'Sinus_venosus' in the anatomy ontology can possess several instances for different species, such as 'EO_Sinus_venosus' and 'DR_Sinus_venosus'. These codes are only used to provide unique names for our individuals and are not used semantically. As mentioned before the instances are linked to their corresponding species numbers in the NCBI taxonomical database using properties in the anatomy ontology, allowing a semantic distinction between species.

The hierarchy in the anatomy ontology together with the instances represents a gradual transition from general to specific knowledge. We add information on the most general level possible, *i.e.* if for a certain anatomical structure information is known to be valid for all vertebrates possessing that structure, this information is added at the class level; this comprises object properties (for instance the right and left atrium always develop from a common atrium), synonyms and comments, which we use to provide anatomical definitions. Through inheritance rules this information is also automatically attached to the instances. Information that does not hold for all possible instances of a class is added only on instance level. This notion of hierarchy corresponds with the semantics of the OWL SubClassOf relation [34]. The inheritance rules also similarly apply to properties; *e.g.* every individual that is linked to another individual using the *constitutional_part_of* relation, is also automatically connected with the *part_of* relation. The notion of property hierarchy corresponds with the semantics of the OWL SubObjectPropertyOf relation [35]. To model the context-dependent behavior of bloodflow we also use multiple properties in a hierarchy reflecting the complex physiological behavior, as discussed below.

4 Case studies

The broad setup of the system allows for various types of queries. We describe four querying uses here. For each of the uses an example question is formulated and the corresponding SPARQL query is given, along with the result set as produced by the Pellet Reasoner [36]. The ontologies are provided as OWL files on our website, http://bio-imaging.liacs.nl/ontologies/ontologies.html, and can be downloaded there. In addition to this we also aim at facilitating the use of the system by presenting an interface on our website with which the user can formulate queries in a webform, which subsequently get translated into SPARQL queries. The design and construction of the interface are discussed in full in section

4.4. The most recent stable version of this interface, along with additional information on its use, can also be found on our website. Although currently not all queries described here are feasible using the interface, the full functionality of the system can be employed by performing SPARQL queries directly on the OWL files.

4.1   Cross-species comparisons

The main purpose of the horizontal approach is to allow for comparative studies between species. Using the NCBI taxonomy database we can ask questions including taxonomical information. An example of such a query and the way in which it is handled by the system is shown below:

"Find all heart septa present in all Sauropsida species included in the ontology". This informal query gets translated into a SPARQL query, shown below with the result set:

```
SELECT ?x
WHERE {
    ?x rdf:type anatomy:Heart_septum.
    ?x anatomy:structure_from_species ?y.
    ?y rdfs:subClassOf ncbi:NCBITaxon_8457.
}
```

Query Results (10 answers):
x
===========================
EO_Aorticopulmonary_septum
EO_Interatrial_septum
EO_Aortic_septum
EO_Horizontal_septum
EO_Septum_primum
EO_Vertical_septum
CP_Aorticopulmonary_septum
CP_Interatrial_septum
CP_Aortic_septum
CP_Interventricular_septum

This result set allows one to study differences in particular structures between species of the same phylogenetic group. Using the information contained in the NCBI database, more complex queries are also feasible; questions can for instance be postulated for all species exluding particular taxonomical groups. Also questions regarding evolutionary development could be formulated, such as " Do the similarities in cardiac anatomy of the mouse, the chick and the zebrafish reflect their taxonomic relationships?". To this end taxonomic distances could be compared to graph distances of the particular RDF subgraphs of different species.

4.2   Developmental studies

As a proof of concept all relevant developmental structures of the zebrafish have been included in the instance set (based on literature; [23],[37]) and linked to stages in the development ontology. This information was obtained from Kimmel *et al.* [23]. Here we describe an example query, concerning the development of the zebrafish, followed by the result set:

"Find all heart compartments in the zebrafish that develop before Kimmel stage prim-15"

```
SELECT ?x
WHERE {
    ?x rdf:type anatomy:Heart_compartment.
    ?x anatomy:structure_from_species ncbi:NCBITaxon_7955.
    ?x anatomy:start_stage ?y.
    ?y development:order_number ?a.
    development:Kimmel_prim_15 development:order_number ?b.
    FILTER ( ?a < ?b )
}
```

Query Results (6 answers):

x

==

DR_Presumptive_atrium_primitive_heart_tube
DR_Presumptive_bulbus_arteriosus
DR_Presumptive_atrium_heart_tube
DR_Presumptive_ventricle_heart_tube
DR_Presumptive_sinus_venosus
DR_Presumptive_ventricle_primitive_heart_tube

Whereas this query results in a list of all structures that arise before prim-15 (regardless of whether they disappear again and if so, when), similar queries can be formulated for ranges of developmental stages as well, specifying start and end stages within which structures arise and/or disappear, or studying the structures that are present at one particular developmental stage.

Furthermore information about taxonomy and development can be included in queries, providing answers to questions such as: "Does the anterior heart field arise first, in a relative sense (by normalising development time), in mouse or chick?". For this the percentage of development and the somite number can be used.

4.3   Context-dependent queries for physiology

Our ontology has been designed to deal with the physiology of bloodflow in different circumstances, which might lead to shunting in particular organisms or during particular stages of development. To model this changing bloodflow, we make use of property semantics in OWL and SWRL logic. In our design qualitative information can be included on bloodflow between cavitated structures using two properties: *can_ receive_blood_from* and *always_receives_blood_from*. Quantitative information about the blood pressure values of the cavitated structures for different shunting situations can be set using the properties *systolic_pressure_no_shunt*, *systolic_pressure_LR_shunt* and *systolic_pressure_RL_shunt*. To distinguish between different shunting situation the boolean properties *RL_shunt* and *LR_shunt* can be set to *true* or *false*. Below we describe how the system uses these properties to model the bloodflow for a specified shunting situation, followed by an example query.

On the most stringent level, bloodflow between certain structures is known to be the same for different physiological states and these structures are connected using *always_receives_blood_from*; on the most relaxed level, when bloodflow between two structures only occurs during a specific physiological state and no exact information can be given for all physiological states at once, the structures are connected using *can_receive_blood_from*. Naturally, due to inheritance rules, structures connected by *always_receives_blood_from* are also automatically connected by *can_receive_blood_from* for all states. A third property, *currently_receives_blood_from* is used dynamically, *i.e.* it is

not used to add static, generally valid information about structures, but can be changed by the system, depending on the shunting situation. Whenever a physiological situation is specified this property is set accordingly, as described below. The hierarchy between the preceding object properties is dictated by the implications each property has. Since we want to model the following logic:

x *always_receives_blood_from* y => x *currently_receives_blood_from* y
x *currently_receives_blood from* y => x *can_receive_blood_from* y

we place *can_receive_blood_from* on the most general level in the hierarchy and *always_receives_blood_from* on the most specific level. SWRL rules test the occurrence of particular physiological states and make sure that the correct structures are connected for that particular situation by the property *currently_receives_blood_from*. This testing of the physiological situation is done by checking the boolean properties *RL_shunt* and *LR_shunt*; the common physiological state (the default situation) is considered to have no shunting, meaning both boolean properties are set to *false*. When we want to study a situation in which shunting occurs we can set the corresponding property (for right-to-left or left-to-right shunting) to *true* in the query. Note that both properties can be set to *true*, which represents a biologically valid bidirectional intracardiac shunt [20]. We are currently working to implement modelling the blood pressure for this situation; the additional properties and rules are not presented here. A high level overview of the implementation of the system is given below:

1) Both shunting properties are set to the situation under study, either *true* in case of shunting or *false* (the default state); this is done for the class Cardiovascular_system.

2) Every class that is part of the class Cardiovascular_system 'inherits' the shunting properties and its values.

3) For each class the blood pressure value for that particular state is set in the property *current_systolic_pressure*. This value is assigned by the system on the basis of the shunting and blood pressure properties.

4) When the blood pressure values for all relevant classes have been set for that particular situation, the following rule is used to determine bloodflow: if two structures (x,y) are related with a *can_receive_blood_from* property and x has a higher *current_systolic_pressure* than y, then y *currently_receives_blood_from* x.

All bloodflow, shunting and blood pressure properties are set manually, except for the properties *currently_receives_blood_from* and *current_systolic_pressure,* which are never used during the addition of new species, but are dynamically assigned by the reasoner using the following SWRL rules:

LR_shunt(?x,false) $\wedge$ RL_shunt(?x,false) $\wedge$ systolic_pressure_no_shunt(?x,?p) $\Rightarrow$
current_systolic_pressure(?x,?p).

LR_shunt(?x,true) $\wedge$ RL_shunt(?x,false) $\wedge$ systolic_pressure_LR_shunt(?x,?p) $\Rightarrow$
current_systolic_pressure(?x,?p).

LR_shunt(?x,false) $\wedge$ RL_shunt(?x,true) $\wedge$
systolic_pressure_RL_shunt(?x,?p) $\Rightarrow$ current_systolic_pressure(?x,?p).

can_receive_blood_from(?x,?y) $\wedge$ current_systolic_pressure(?x,?p) $\wedge$
current_systolic_pressure(?y,?q) $\wedge$ (?p<?q) $\Rightarrow$ currently_receives_blood_from(?x,?y)

Using this system queries can be formulated regarding the bloodflow during the normal physiological state as well as during periods of shunting. An example of a possible biological question is given below, with the SPARQL query and the resulting blood pressure values (with 'a' denoting the right ventricle and 'b' the left ventricle):

" What are the systolic pressures for right and left ventricle without shunting and for right-to-left shunting?"

```
SELECT ?a ?b
WHERE {
    anatomy:CP_Right_ventricle anatomy:current_systolic_pressure ?a.
    anatomy:CP_Left_ventricle anatomy:current_systolic_pressure ?b.
}
```

No shunt:
Query Results (1 answers):
a | b
=======
22 | 67

Right-to-left shunt:
Query Results (1 answers):
a | b
=======
47 | 50

4.4   Querying using the interface and 3D visualisation of query results

Querying the system will result in textual output, listing the classes and/or instances corresponding to the query. As a particular feature of the interface we have extended the querying results with the possibility of visualisation of the textual results, using 3D reconstructions previously made. Currently we have 3D reconstructions of the adult zebrafish heart and of several developmental stages of the turtle heart. These 3D reconstructions have been created using our inhouse software TDR-3Dbase [17] and can be viewed online using our TDR-3D viewer ([38],[17]).

The interface for querying our ontology system is currently under development. However the intended design and construction have been outlined and are described here. The interface consists of a web form, divided over four tabs, enabling the user to formulate a query. No hierarchy of priority exists between the tabs and the user is free to chose which filters he or she would want to use. While the form is being filled in by the user, the choices the user can make are automatically presented by looking up the remaining possibilities given the user's previous entries. The first tab lets the user select structures of interest. In addition to particular structures we allow groups of structures to be chosen as well, *e.g.* selecting all cavitated structures. A second tab lets the user specify the species he is interested in. Only species that have instances in our ontology can be selected. In another tab the staging system corresponding to the specified species is selected (*e.g. Kimmel* for zebrafish or *Theiler* for mouse). Here the user will be presented with several options for setting the staging restrictions; different staging methods can be used (for instance the stage names for a particular species, the absolute time of development in hours or merely the distinction between adult and embryological structures) and the user can indicate staging with an exact stage or a range of stages. Lastly there is a fourth tab which lets the user filter the information using the properties linked to structures and instances. At every step in the form, a SPARQL

query is produced, creating a subgraph of all the solutions that still match the restrictions given by the user in the form. When the user submits the query form, a list of all the individuals in our ontology is returned. If we have a 3D reconstruction available for this species at this stage, we provide a link to the TDR-3D viewer, with the correct model and structures already selected in the parameters.

The implementation of the form is done using AJAX calls to Java servlets that look up the relevant information from the ontology using the JENA library and SPARQL template queries. While this form is currently focused on 3D visualisations, we aim to extend its functionality to allow more complicated queries, encomporating more detailed information provided in the anatomy and development ontologies, combined with the NCBI database.

The following example illustrates the operation of an interface query. A biologist has a request: "Show all endocardial cushions from the species *Emys orbicularis* that are positioned in a compartment of the heart." This request is filled out in a web form and is internally translated into a query, shown below with result set:

```
SELECT ?x WHERE {
    ?x anatomy:structure_from_species ncbi:NCBITaxon_82168,
    ?x rdf:type anatomy:Endocardial_cushion,
    ?x anatomy:regional_part_of anatomy:Heart_compartment.
}
```

Query Results (4 answers):

```
x
==============================================
EO_Atrioventricular_cushion
EO_Distal_cushion_outflow_tract
EO_Proximal_cushion_outflow_tract
EO_Messenchymal_cap
```

This result set is subsequently visualised in the viewer applet, a screen dump of which is shown in Fig. 4. The viewer renders scenes on the client-side using contour information from TDR-3Dbase. In earlier work by Wong *et al.* [39] scenes were generated at the server side. Our design is focused more on the ability to serve multiple clients at the same time with minimal bandwidth.

This architecture has two seperate components for the heart terms and their 3D models. In a paper by Köhn *et al.* [40] a system is described that combines the two types of information into a single ontology. Because we incorporate multiple species this approach is not feasible for our system.

5   Conclusions

Over the last few decades several ontologies have been created comprising anatomical information. We add to this rapidly growing field by presenting a novel approach centered around organ systems; we have elaborated this approach for the vertebrate heart. As opposed to 'vertical' modelling, focusing on the entire anatomy of one species, we propose a 'horizontal' way of modeling, looking at one organ for a wide range of species simultaneously. This is the key feature of the presented approach. The system we constructed serves a double purpose: the anatomy ontology can be used to annotate information for additional vertebrate species, allowing researchers to supplement the system with information about their own species. Furthermore, the instance information that is currently presented for different model animals forms a knowledge base which can be queried to further our

knowledge in the developmental, anatomical and physiological aspects of the vertebrate heart. The system is provided on our website and can be used either by directly using the OWL files (to complement or study the information) or by formulating queries using the interface.
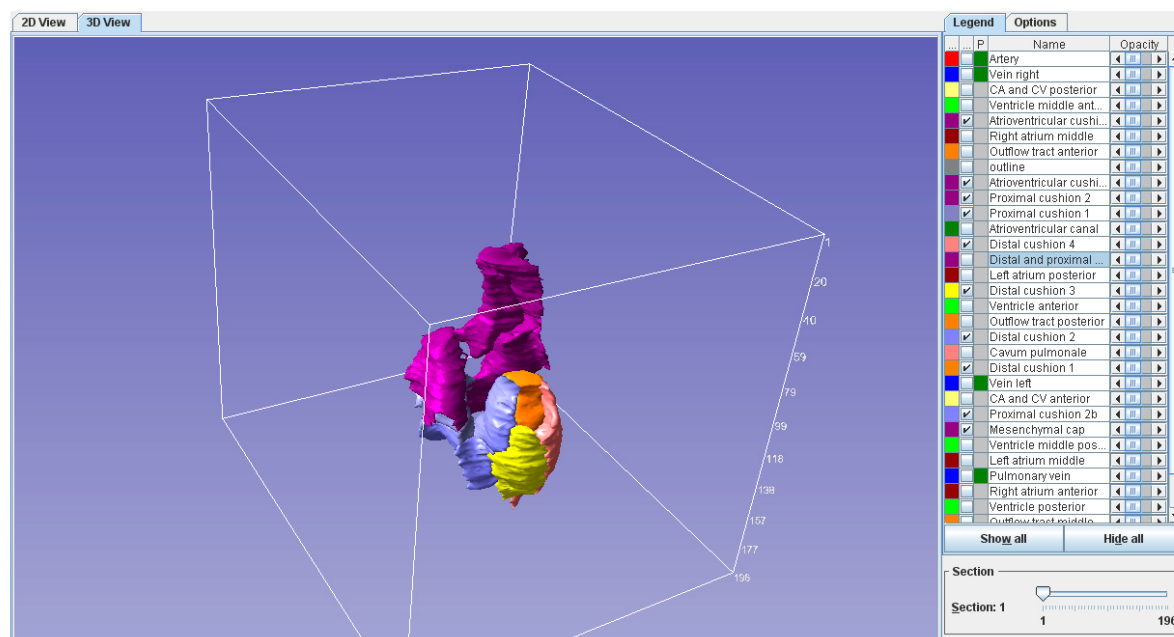

Figure 4: Screen dump of the TDR-3D viewer, in which a set of structures is visualised for the turtle species *Emys orbicularis*, corresponding to the query described in the text.

Here we have presented case studies of the vertebrate heart, but the architecture of the system makes it applicable to a wide range of organs and anatomical systems. Additionally, using NCBI accession numbers instead of locally defined name spaces to distinguish species, opens up the system to semantic interoperability. Using this generic approach we have build an open and easily extendible system.

6   Future work

We are currently incorporating instance information for avian (chick) and mammal (mouse) hearts and adding relevant anatomical classes to the anatomy ontology accordingly. For all species in the ontology developmental information is collected and linked to the development ontology. Furthermore future work will be directed to extending the functionality by providing more querying options; we strive for interoperability with ontologies relating to tissue types phenotypes and diseases. Lastly, the user interface will be amended to allow the user more freedom in formulating complex queries in the field of evolutionary development, anatomy and physiology. New instances of the interface will be made suitable for online use. To this end usibility studies will be performed.

References

[1]   O. Bodenreider and R. Stevens. Bio-ontologies: current trends and future directions. Briefings in bioinformatics, 7(3): 256-274, 2006.

[2]   R. McEntire. Ontologies in the life sciences. The knowledge engineering review, 17(1): 77-80, 2002.

[3]   K. Baclawski and T. Niu. Ontologies for bioinformatics. Cambridge, Massachusetts: MIT Press, 2006

[4]   A. Burger, D. Davidson, R. Baldock (eds.). Anatomy ontologies for bioinformatics. London: Springer, 2008.

[5]   S. Staab and R. Studer (eds.). Handbook on ontologies. Berlin: Springer, 2004.

[6]   B. Smith, M. Ashburner, C. Rosse, C. Bard, W. Bug, W. Ceusters, L.J. Goldberg, K. Eilbeck, A. Ireland, C.J. Mungall, The OBI Consortium, N. Leontis, P. Rocca-Serra, A. Ruttenberg, S. Sansone, R.H. Scheuermann, N. Shah, P.L. Whetzel and S. Lewis. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, Nature Biotechnology, 25: 1251 – 1255, 2007.

[7]   C. Rosse and J.L.V. Mejino Jr. A reference guide for biomedical informatics: the Foundational Model of Anatomy. Journal of biomedical informatics, 36: 478-500, 2003.

[8]   C. Rosse and J.L.V. Mejino Jr. The Foundational Model of Anatomy ontology. In: A. Burger, D. Davidson and R. Baldock (eds.). Anatomy ontologies for bioinformatics: principles and practice. New York: Springer, 2008.

[9]   R. Baldock, J. Bard, A. Burger, N. Burton, J. Christiansen, G. Feng, B. Hill, D. Houghton, M. Kaufman, J. Rao, J. Sharpe, A. Ross, P. Stevenson, S. Venkataraman, A. Waterhouse, Y. Yang and D. Davidson. EMAP and EMAGE: A Framework for Understanding Spatially Organised Data. Neuroinformatics 1: 309-325, 2003.

[10]   A.M. Maglia, J.L. Leopold, L.A. Pugener and S. Gauch. An anatomical ontology for amphibians. Proceedings Pacific Symposium on Biocomputing, 12: 367-378, 2007.

[11]   Y. Bradford, T. Conlin, N. Dunn, D. Fashena, K. Frazer, D.G. Howe, J. Knight, P. Mani, R. Martin, S.A. Moxon, H. Paddock, C. Pich, S. Ramachandran, B.J. Ruef, L. Ruzicka, H. Bauer Schaper, K. Schaper, X. Shao, A. Singer, J. Sprague, B. Sprunger, C. Van Slyke and M. Westerfield. ZFIN: enhancements and updates to the zebrafish model organism database. Nucleic acids research, 39(suppl 1): 822-829, 2011.

[12]   E.W. Sayers, T. Barrett, D.A. Benson, S.H. Bryant, K. Canese, V. Chetvernin, D.M. Church, M. DiCuccio, R. Edgar, S. Federhen, M. Feolo, L.Y. Geer, W. Helmberg, Y. Kapustin, D. Landsman, D.J. Lipman, T.L. Madden, D.R. Maglott, V. Miller, I. Mizrachi, J. Ostell, K.D. Pruitt, G.D. Schuler, E. Sequeira, S.T. Sherry, M. Shumway, K. Sirotkin, A. Souvorov, G. Starchenko, T.A. Tatusova, L. Wagner, E. Yaschenko and J. Ye. Database resources of the National Center for Biotechnology Information. Nucleic acids research, 37: 5-15, 2008.

[13]   SWRL: a Semantic Web Rule Language combining OWL and RuleML (accessed on 29-04-2011): http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/

[14]   M. Axelsson, C.E. Franklin, C.O. Löfman, S. Nilsson and G.C. Grigg. Dynamic anatomical study of cardiac shunting in crocodiles using high-resolution angioscopy. Journal of experimental biology, 199:359-365, 1996.

[15]   M. Axelsson, C.E. Franklin, R. Fritsche, G.C. Grigg and S. Nilsson. The sub-pulmonary conus and the arterial anastomosis as important sites of cardiovascular regulation in the crocodile *Crocodylus porosus*. Journal of experimental biology, 200: 807-814, 1997.

[16]   M. Axelsson. The crocodilian heart; more controlled than we thought? Experimental physiology, 86.6: 785-789, 2001.

[17]   L.M.F Bertens, M.K. Richardson and F.J. Verbeek. Analysis of cardiac development in the turtle *Emys orbicularis* (Testudines: Emidydae) using 3-D computer modeling from histological sections. Anatomical record, 293: 1101-1114, 2010.

[18]   G.C. Grigg and K. Johansen. Cardiovascular dynamics in *Crocodylus porosus* breathing air and during voluntary aerobic dives. Journal of comparative physiology, 157: 382-392, 1987.

[19]   G. Grigg. Central cardiovascular anatomy and function in Crocodilia. In: S.C. Wood, R.E. Weber, A.R. Hargens, R.W. Millard (eds.). Lung biology in health and disease, Volume 56. New York: Marcel Dekkers, 1991.

[20]   K. Johansen and W.W. Burggren. Cardiovascular function in lower vertebrates. In: G. Bourne (ed.). Hearts and heart-like organs. New York, Academic press, 1980.

[21]   J.W. Hicks and E. Krosniunas. Physiological states and intracardiac shunting in non-crocodilian reptiles. Experimental biology online, 1(3), 1996.

[22]   W.W. Burggren. Form and function in Reptilian circulations. American zoology, 27: 5-19, 1987.

[23]   C.B. Kimmel, W.W. Ballard, S.R. Kimmel, B. Ullmann and T.F. Schilling. Stages of embryonic development of the zebrafish. Developmental dynamics, 203: 253-310, 1995.

[24]   C.L. Yntema. A series of stages in the embryonic development of *Chelydra serpentina*. Journal of morphology, 125: 219-252, 1968.

[25]   K. Theiler. The house mouse, atlas of embryonic development. New York, Springer-Verlag, 1989.

[26]   R. O'Rahilly and F. Müller. Developmental stages in human embryos. Carnegie Institution of Washington Publication 637, 1987.

[27]   V. Hamburger and H.L. Hamilton. A series of normal stages in the development of the chick embryo. Juornal of Morphology, 88(1): 49-92, 1951.

[28]   H. Knublauch, R.W. Fergerson, N.F. Noy and M. Musen. The Protégé OWL Plugin: an open development environment for semantic web applications. Lecture notes in computer science, 3298: 229–243, 2004.

[29]   M.A. Haendel, F. Neuhaus, D. Osumi-Sutherland, P.M. Mabee, J.L.V. Mejino Jr., C.J. Mungall and B. Smith. CARO – the Common Anatomy Reference Ontology. In: A. Burger, D. Davidson and R. Baldock (eds.). Anatomy ontologies for bioinformatics: principles and practice. New York: Springer, 2008.

[30]   J.L.V. Mejino Jr., A.V. Agoncillo, K.L. Rickard and C. Rosse. Representing complexity in part-whole relationships within the Foundational Model of Anatomy. Proceedings AMIA 2003 symposium: 450-454, 2003.

[31]   A. Hunter, M.H. Kaufman, A. McKay, R. Baldock, M.W. Simmen and J.B.L. Bard. An ontology of human developmental anatomy. Journal of anatomy, 203(4): 347-355, 2003.

[32]   B.F. Schneider and S. Norton. Equivalent ages in rat, mouse and chick embryos. Teratology 19: 273-278, 1979.

[33]   R. Bathoorn, M. Welten, M. Richardson, A Siebes and F.J. Verbeek. Frequent episode mining to support pattern analysis in developmental biology. Lecture notes in computer science, 6282: 253-263, 2010.

[34]   OWL 2 Web Ontology Language Primer, class hierarchies (accessed on 29-04-2011): http://www.w3.org/TR/2009/REC-owl2-primer-20091027/#Class_Hierarchies

[35]   OWL 2 Web Ontology Language Primer, property hierarchies (accessed on 29-04-2011):
http://www.w3.org/TR/2009/REC-owl2-primer-20091027/#Property_Hierarchies

[36]   E. Sirin, B. Parsia, B.C. Grau, A. Kalyanpur and Y. Katz. Pellet: a practical OWL-DL reasoner. Journal of Web Semantics, 5(2), 2007.

[37]   N. Hu, D. Sedmera, H.J. Yost and E.B. Clark. Structure and function of the developing zebrafish heart. Anatomical record, 260: 148-157, 2000.

[38]   D. Potikanond, M. Belmamoune, M.C.M. Welten, F.J. Verbeek. Visual integration of 3D digital atlas and 3D patterns of gene expression in zebrafish. International Symposium on Integrative Bioinformatics, Wageningen, 2011.

[39]   B.A. Wong, C. Rosse and J.F. Brinkley. Semi-automatic scene generation using the Digital Anatomist Foundational Model. Proceedings American medical informatics association fall symposium: 637-41, 1999.

[40]   S. Köhn, R.H. van Lengen, G. Reis, M. Bertram and H. Hagen. VES: Virtual Echocardiography System. Proceedings IASTED Visualization, imaging and image processing (VIIP): 465-471, 2004.