Development of the Amphibian Anatomical Ontology

Anne Maglia

Department of Biological Sciences

Missouri University of Science and Technology

Jennifer Leopold and Analía Pugener, Missouri S&T Susan Gauch, U. Arkansas

Community-identified need (ATOL)

- Anatomical ontology vital to amphibian research
 - Common gene expression and embryology models
- Three disparate lexicons
- Lack of terminological standardization limits understanding of phenotype evolution and integrative research on gene expression, embryology, and comparative anatomy

Challenge

Develop an amphibian anatomical ontology that accommodates the diversity of structures present in all amphibians; include definitions, literature references, coding of phylogenetic characters, and images and annotations

- ...and include the domain expert community
- ...and maintain interoperability with other ontologies
- ...and do it in a reasonable amount of time

AmphibAnat approach

Combine existing tools and methodologies with new approaches

- 1. Semi-automated construction
- 2. Ontology maintenance
- 3. Web-based community curation

1. Semi-automated construction

- Challenge requires automated techniques to reduce manual efforts and enhance existing, manually created ontologies by:
 - Enriching the ontology (adding terms to the AAO) and annotating (= character states)
 - Applying metrics to benchmark the semi-automated performance

Semi-automated approach

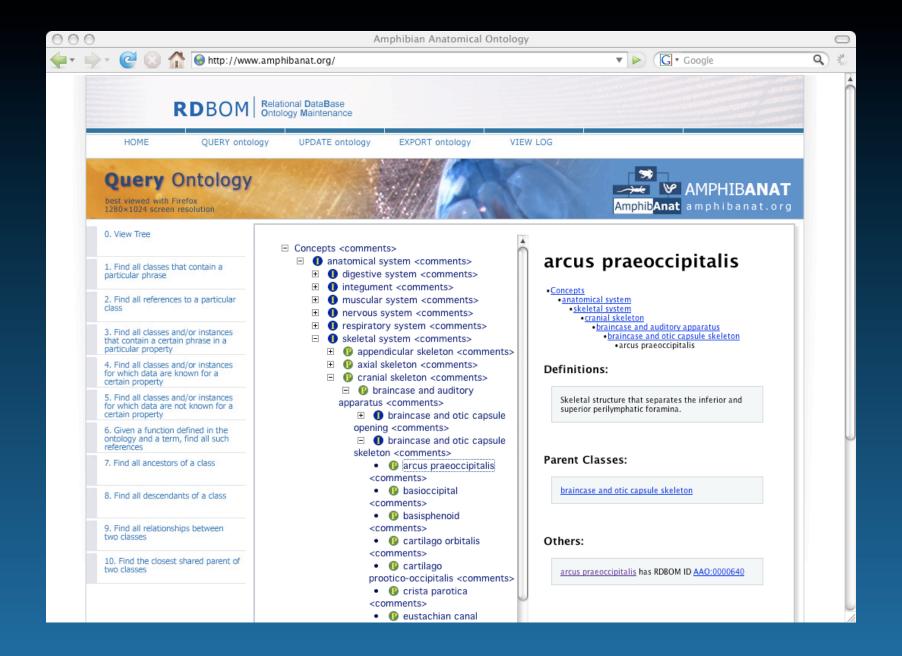
- Manually construct small subset(s)
 - Started with ZFIN/TAO
- Rigorous community evaluation, augmentation, and modification of subset ontology
- Seed IR software with skeleton of ontology
 - Use to find terms and annotations
- Benchmark, repeat until software is reasonably accurate, then let loose on entire ontology

How does the software work?

- Identify relevant electronic data sources
 - Use topic-specific spider to generate queries for concepts in the ontology
 - Collect potentially relevant documents
- Text mine info from documents
 - Pattern-based extraction methods
 - Statistical natural language processing algorithms that identify and weight most important elements

2. Ontology maintenance

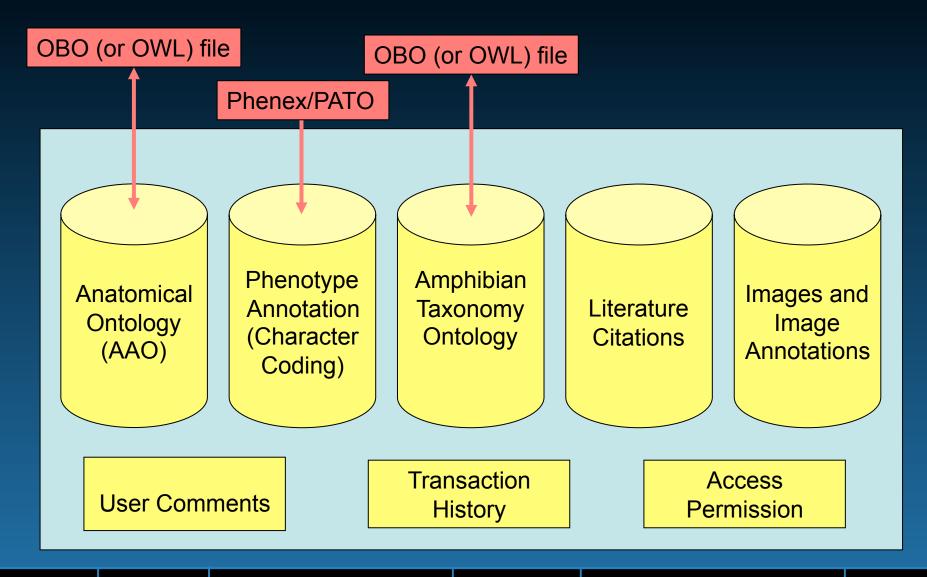
- Our needs required maintenance software that:
 - Allows web-based access and collaboration
 - Ensures consistency and concurrent, authorized access
 - Allows query and update of context and structure
 - Allows import/export of common exchange formats (e.g., OBO, OWL)
 - Allows node-based access and editing privileges



RDBOM ("red-bomb")

- Relational Database Ontology Maintenance
- Features:
 - Structural (e.g., ancestors/descendents) and content-based queries (e.g., all concepts with particular phrase)
 - Structural updating (e.g., move/insert/delete nodes)
 - Content-based updating (e.g., change the value)
 - Collaborative commenting
 - Security (node-based per user; multiple users)
 - Version control
 - Web-based access
 - Generic (any ontology)
 - Modularization of data

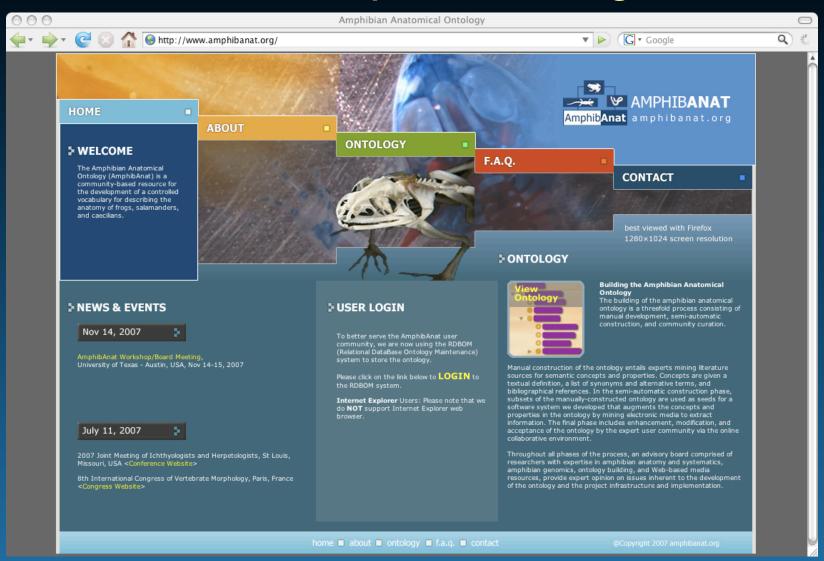
Modularization



3. Web-based community curation

- Challenge required us to develop a virtual organization of amphibian experts
 - Encourage user commenting by node
 - Allow "super-user" node-based curation
 - Teams and individuals take ownership of particular subsets for which they are the experts

www.amphibanat.org



Query Ontology

Dest moved with Firefus 1280 v 5024 screen resolution



- 1. View Tree
- Find all classes that contain a particular phrase
- Find all references to a particular class.
- Find all classes and/or instances that contain a specific phrase/property combination
- Given a class and property, find all linked classes/instances
- 6. Find all ancestors of a class.
- 7. Find all descendants of a class
- Find all relationships between two classes
- Find the closest shared ancestor of two classes

- □ Concepts <comments>
 - 0 amphibian anatomical entity <comments>
 - In anatomical space <comments>
 - ① anatomical structure <comments>
 - - anatomical cluster < comments>
 - anatomical system <comments>
 - ⊕ circulatory system <comments>
 - digestive system
 <comments>
 - Integument < comments>
 - muscular system
 - <comments>

 - <comments>

 - <comments>

 - <comments>
 - appendicular skeleton
 comments>
 - □ 0 cranium <comments>
 - ⊕ braincase and auditory apparatus
 <comments>
 - □ braincase and otic capsule opening
 - basicranial fontanelle
 - <comments>

 O fenestra ovalis.
 - <comments>
 - ☐ foramen

orbital region

- Concepts
 - amphibian anatomical entity
 anatomical structure
 - ·anatomical group
 - anatomical system
 skeletal system
 - •cranium •orbital region

Definitions:

An anatomical cluster that is part of the cranium and structurally supports the eye.

Parent Classes:

cranium

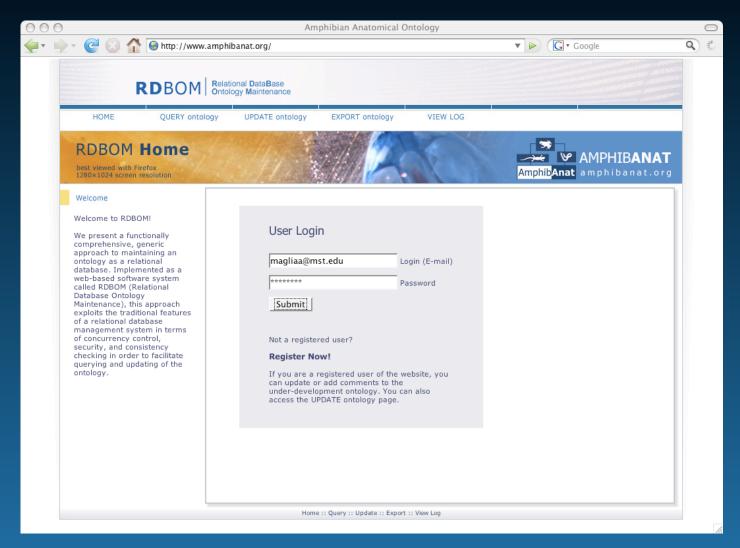
has_RDBOM_ID:

orbital region has_RDBOM_ID AAO:0010220

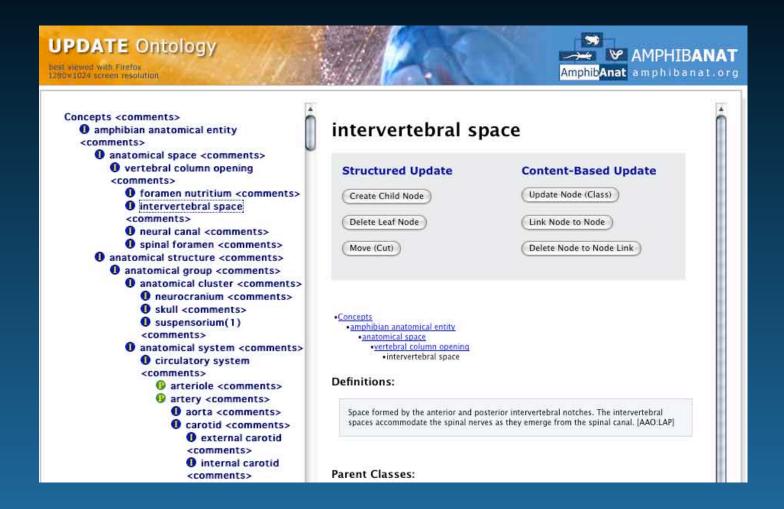
Commenting tool

Hi Dr. Anne "basale com		come to the	View / Add Comments: "basale commune" page. You can view or add comments at this	page for
If you want	to edit submi	itted comme	nts for "basale commune", please go to EDIT MY COMMENTS page.	
* Maximum	input 2000 c	haracters.		
Submit Co	mments			
Node Name	Firstname	Lastname	Comments	DateTime
basale commune	David	Wake	The basale commune is a synapomorphy of living caudata and may ultimately to be a synapomorphy of a deeper node. It evidently evolved from a phylogenetic amalgamation of original distal tarsals 1 and 2. However, there is scant to no evidence that there are two centers during ontogeny, so it is advised to use the term basale commune instead of distal tarsal $1 + 2$.	11/9/2008 9:34:32 AM

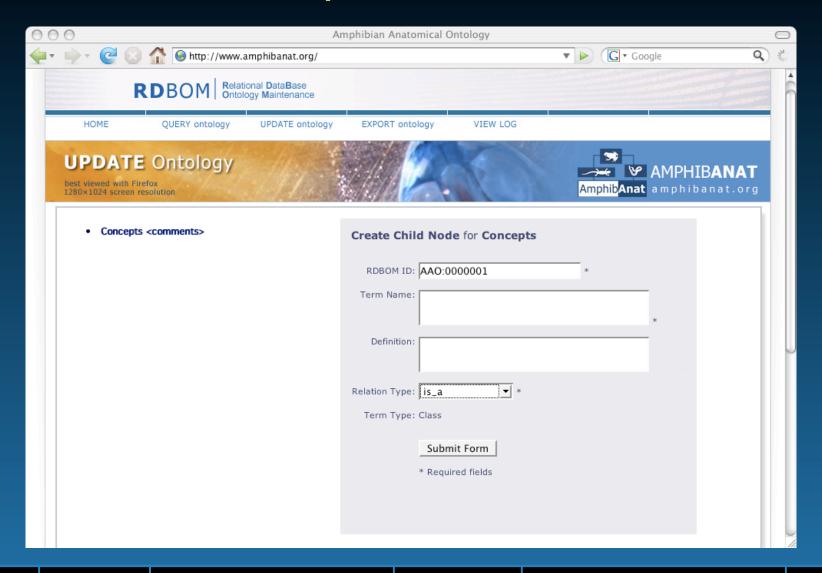
Privilege assignment



"Super-user" functions



Simple interface



Web-based community curation

- Facilitates inclusion of area expertise
- Provides forum for argumentation
 - E.g., definitions, homology, preferred terminology
- Allows community ownership (and thus guaranteed use)
- Value-added: unites diverse groups toward common goal

Summary

- Biodiverse group with disparate lexicons required combination of existing tools and novel approaches
- Semi-automated construction should help populate ontology faster, with less manual effort
- Community-based curation allows user ownership, thus mining expert knowledge and facilitating use
- Database maintenance system allows biodiversity data and functionality while allowing interoperability with other groups (e.g., OBO Foundry)

Where are we now?

- Two user workshops
 - 40 domain experts trained in system
- Community helping to build subset ontologies
- Text classification, evaluation, and training tools in final testing phase
- Automated annotation tools in first phase of testing/benchmarking







Acknowledgements

- NSF DBI 0640053
- Barbara Fears, Bonnie Beasley, Alton Coalter, Leong Lee, Sarah Havens
- AmphibiaTree (NSF ATOL)
- AmphibAnat Advisory Board, workshop participants, and evaluators/contributors