“Tree of Life” Teaching Tool

Madeline Yasner
University of Pennsylvania
DMD Senior Design
May 1, 2009
Advisor: Val Tannen

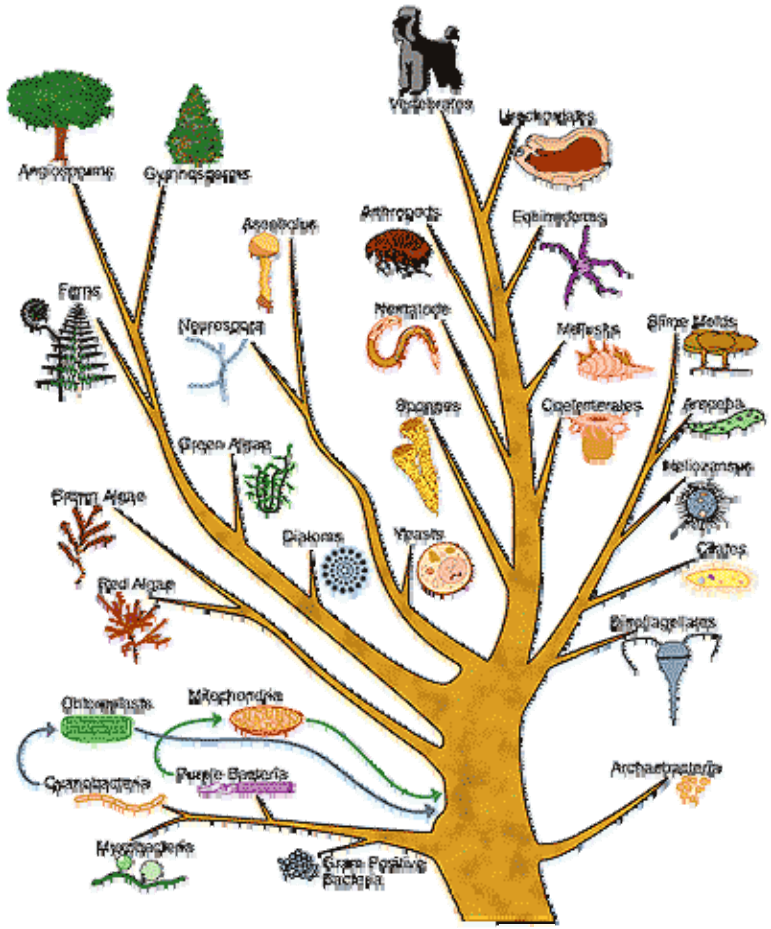
Presentation Outline

- Describing My Project
- Expanding PhyloWidget
  - Algorithms
  - Results and Achievements
  - Demo
- Wrap-Up
  - Future Directions
  - Tools/Languages
  - Thanks
- Questions

The Problem

Phylogenetics

the study of how organisms are related by evolution



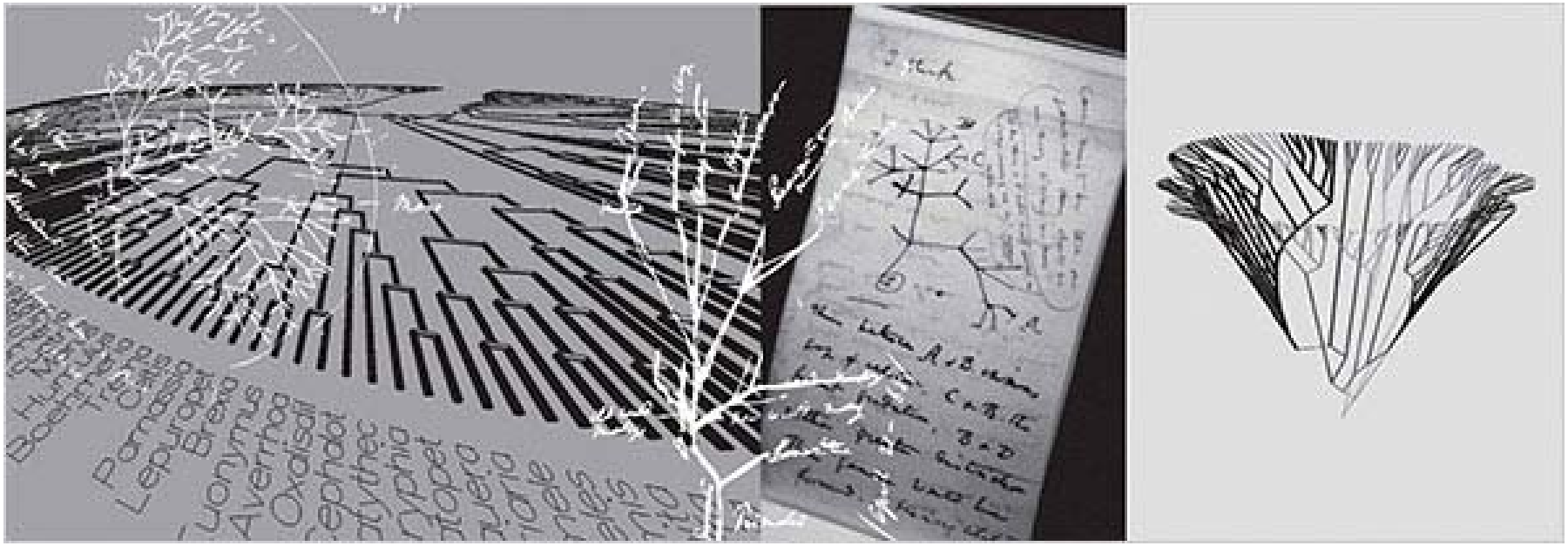
The Problem

Data is only as valuable as a user's ability to interact with it and learn from it.

My program allows students to use this great body of knowledge in a way that is better suited to their needs.

The Problem

Tree Visualization in The New York Times



Prior Work

Tree of Life Web Project

- Catalogue and classify data about *biodiversity* and *phylogenetics*

- Aid *learning* about and *appreciation* of biological diversity and the evolutionary Tree of Life

Prior Work

TreeBASE II

published phylogenetic research and the trees generated by those studies

Prior Work

Tamara Munzner's TreeJuxtaposer

- Scalability in Tree and Display Size
- Guaranteed Visibility of Landmark Nodes
- Limited Label Overlap

Prior Work

Tamara Munzner's TreeJuxtaposer

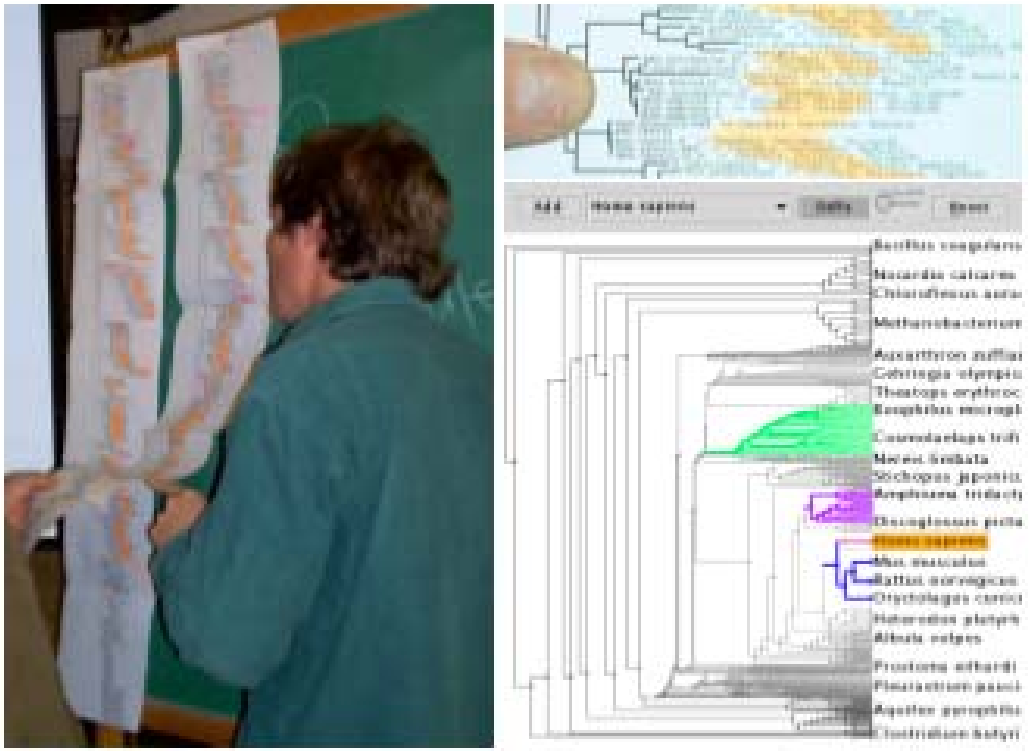


Figure 1: Left, Right Top: Biologists faced with inadequate tools for comparing large trees have fallen back on paper, tape, and highlighter pens. Right Bottom: TreeJuxtaposer is a scalable tool for interactive exploration and comparison of trees.

Prior Work

Tamara Munzner's TreeJuxtaposer

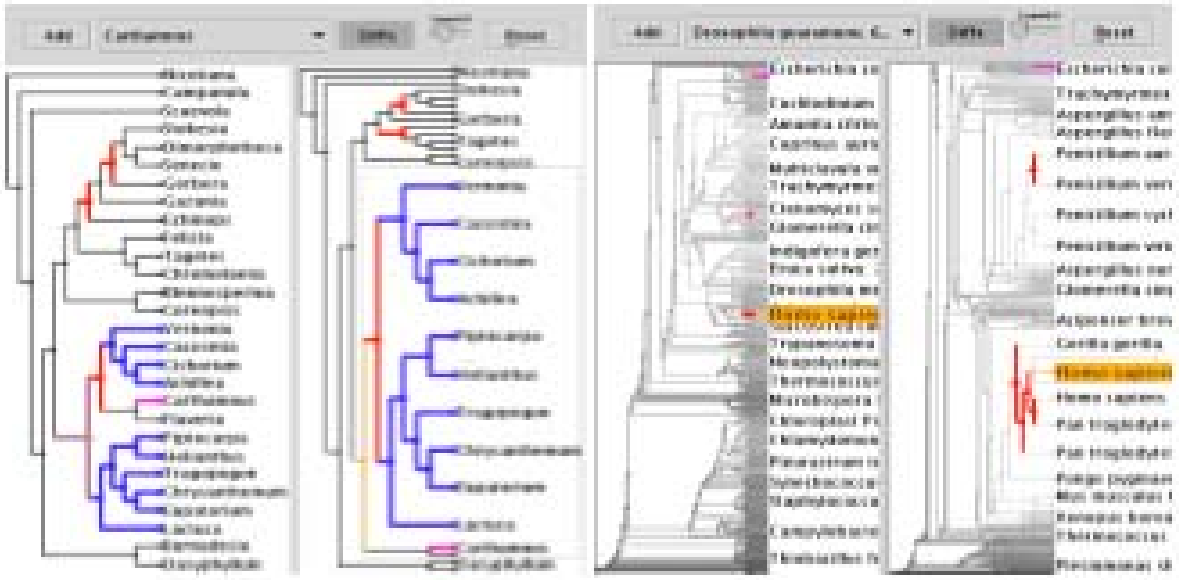
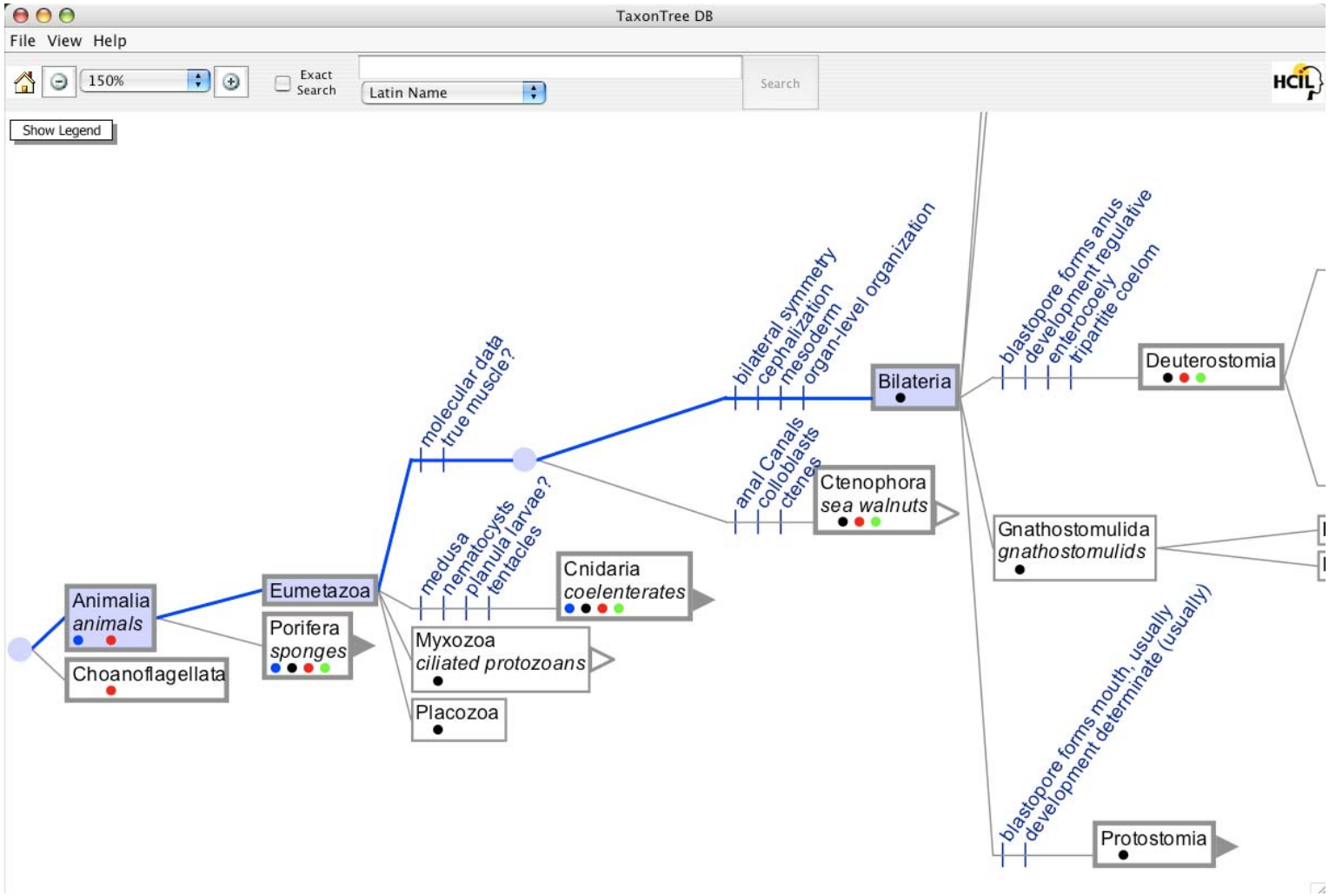


Figure 5: Tree comparison between two variants of a single phylogenetic reconstruction run, with the exact location of structural differences (in biological terms, nonmonophyletic clades) marked in red. The left tree is in the undistorted overview position, while parts of the right side have been expanded. **Left:** A small 55-node tree. **Right:** A larger 1600 node tree.

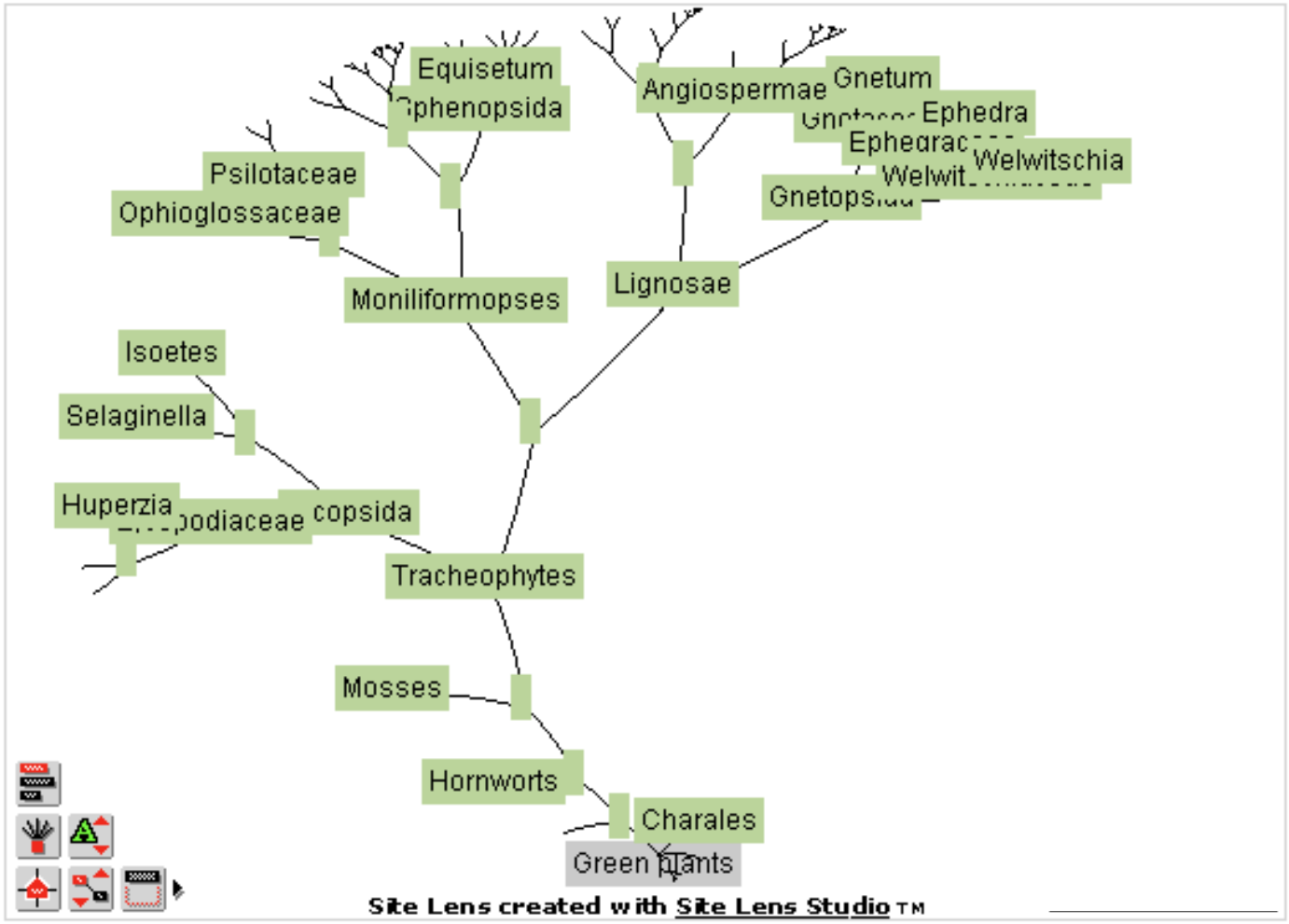
Prior Work

Denise Green and Rebecca Shapley



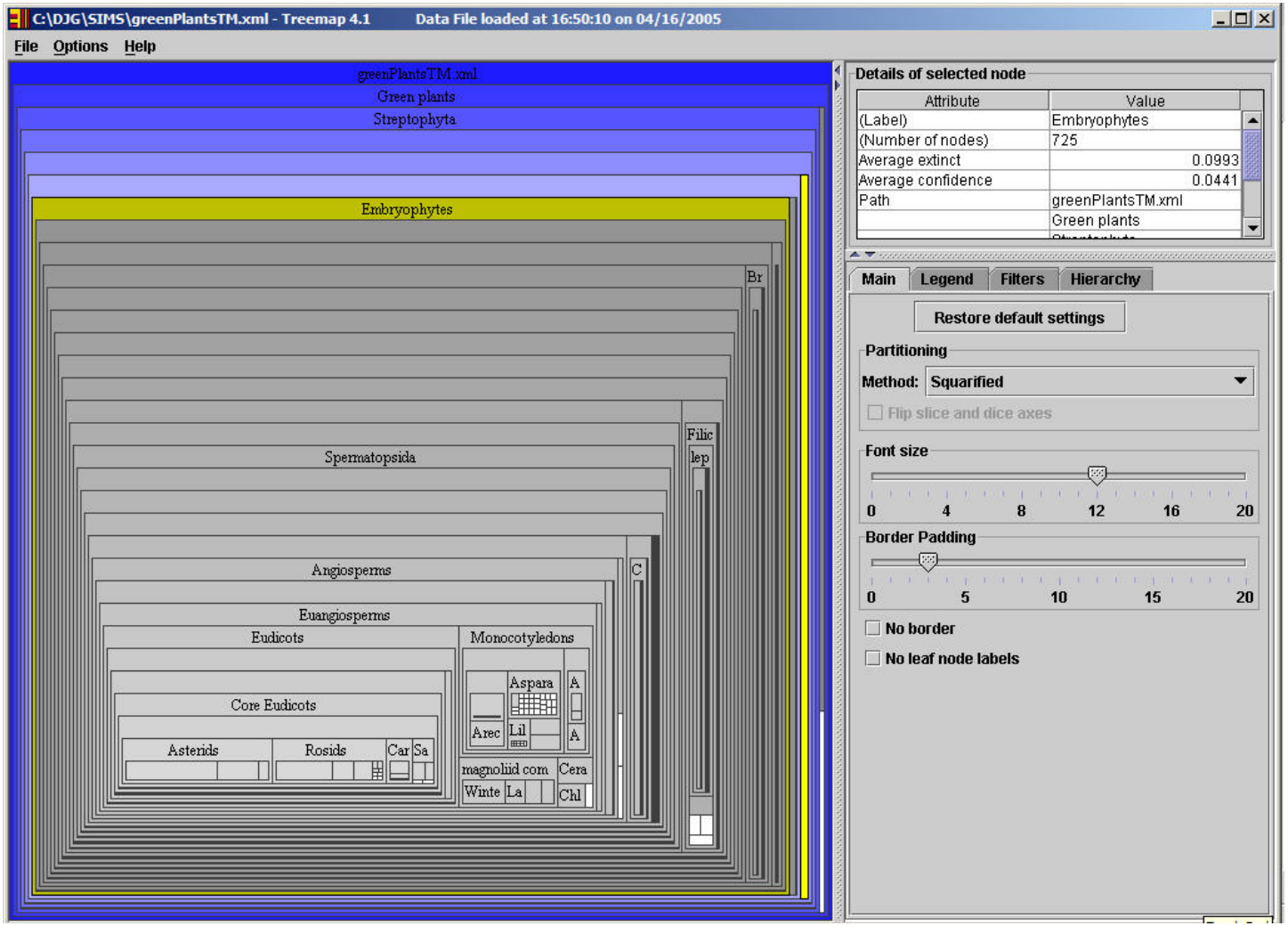
Prior Work

Denise Green and Rebecca Shapley



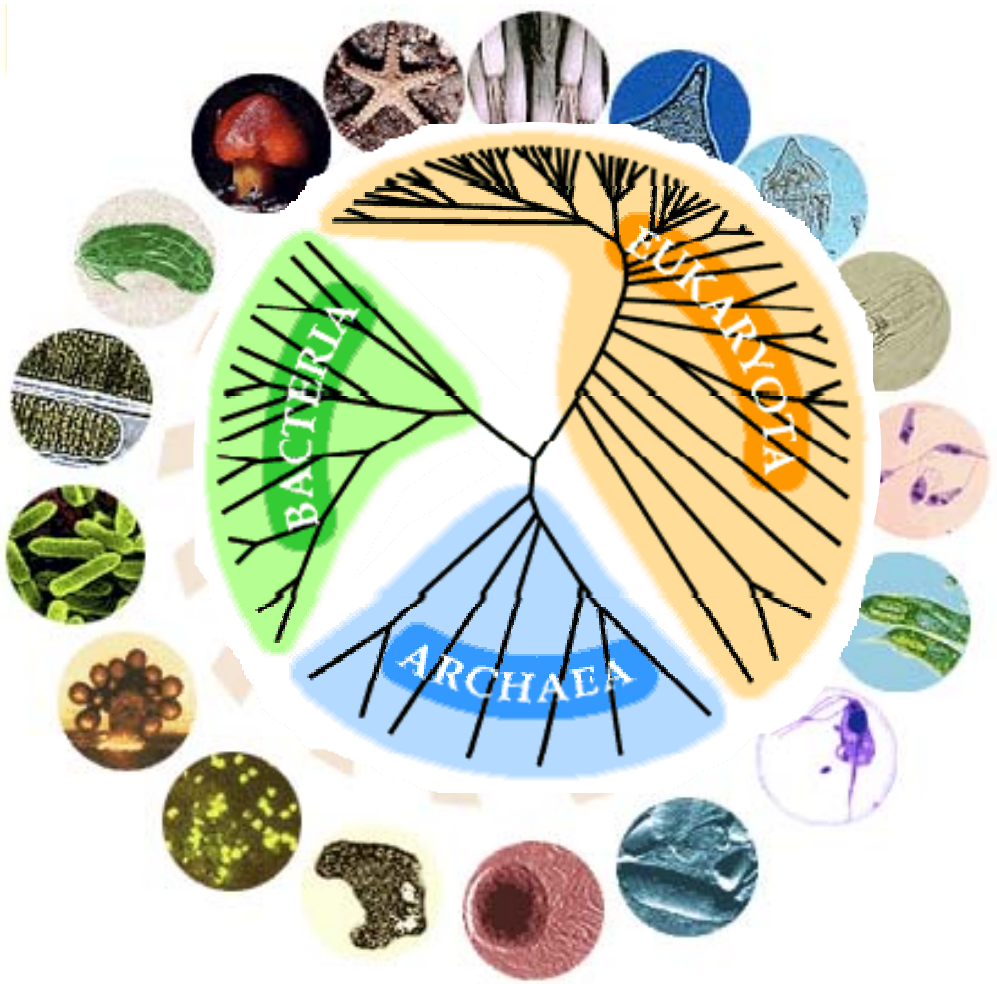
Prior Work

Denise Green and Rebecca Shapley



Prior Work

Denise Green and Rebecca Shapley



Prior Work

Denise Green and Rebecca Shapley

General Recommendations:

- Solve the “Overwhelm Factor”
- Simple, Teachable Tree Views
- Connect with the Familiar: History, Common Names, and Pictures

Prior Work

Denise Green and Rebecca Shapley

Relevant Recommendations

- *Images of organisms*
- *Common name search*
- Variable branch lengths
- Rank-free circular layout
- Animated changes
- Non-overlapping labels
- Resizeable window and text

Prior Work

Denise Green and Rebecca Shapley

A *fully integrated system** ... would be fantastic. Having a website that seamlessly relates [data from independent sources] would *save an unimaginable amount of teaching and preparation time*.

* Emphasis added

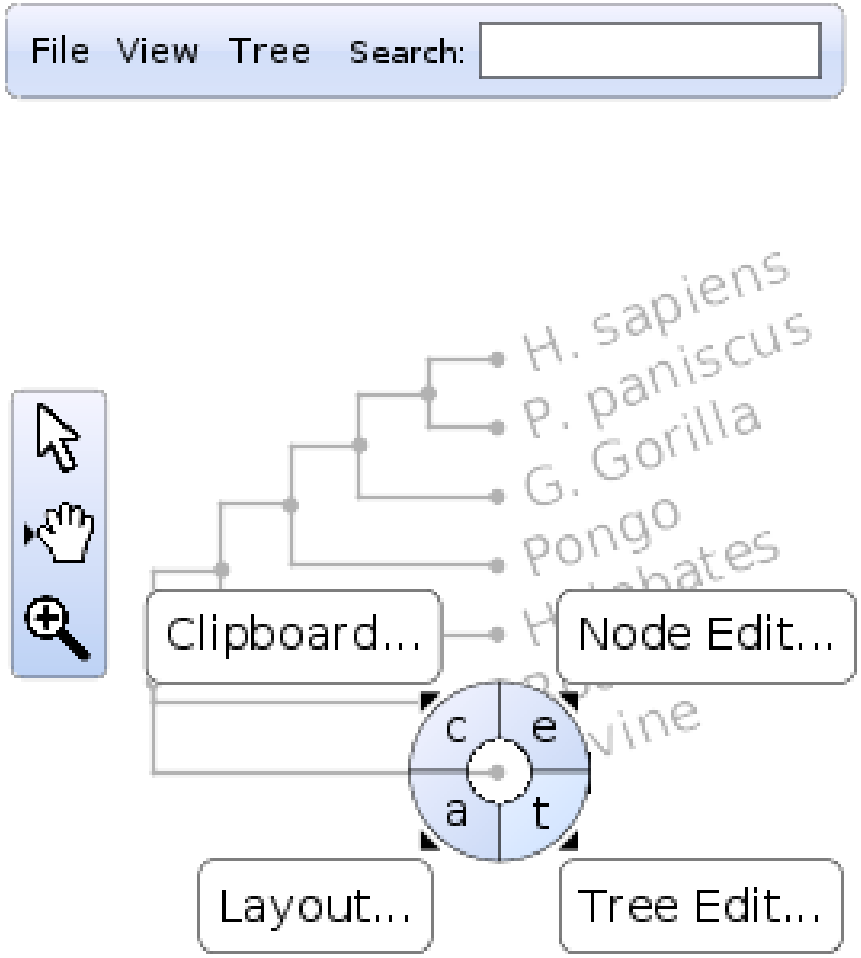
Prior Work

PhyloWidget

"*viewing*, *editing*, and *publishing*
phylogenetic trees online"

Prior Work

PhyloWidget



Prior Work

PhyloWidget

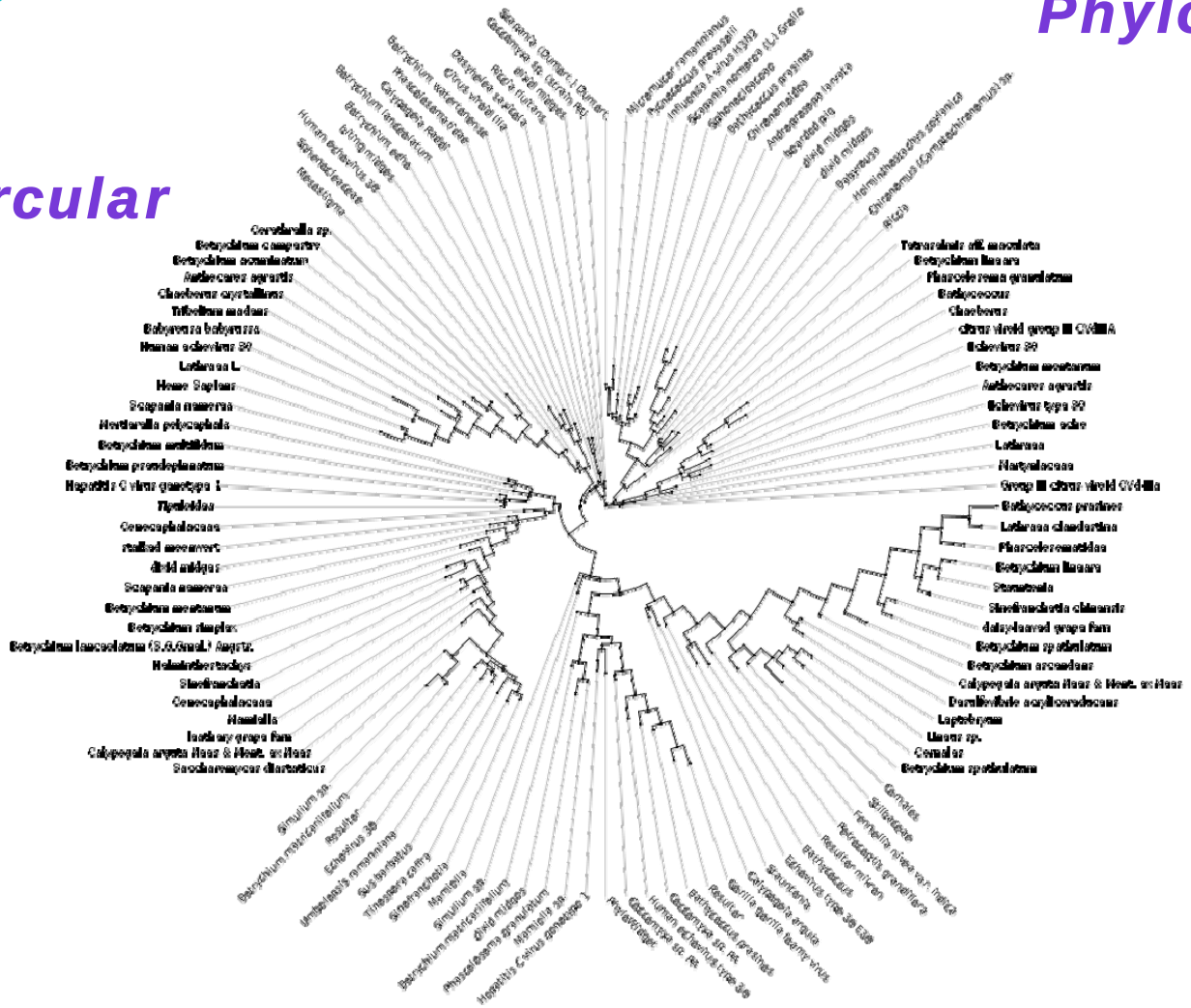
cladogram

diagonal

Prior Work

PhyloWidget

circular


Tetrarhanis aff. maculata
Botrycladus lineare
Pharcidoreuma granulatum
Botryococcus
Chaetovus
citrus viroid group III CVd-IIIA
Scleroviruz 99
Botrycladus acostaensis
Asdecoruz agrartis
Scleroviruz type 99
Botrycladus sobo
Lotarosa
Hartyalaceae
Group III citrus viroid CVd-IIIa
Botryococcus pratineri
Lotarosa clavidortina
Pharcoleoreuvidae
Botrycladus lineare
Stomatovela
Sinefranchesta chlamydis
daisy-leaved grape fern
Botrycladus sp. stanislaus
Botrycladus arcuadeus
Calypogeia arguta Haar & Hewt. ex Haar
Parmilevidale acyllocerodocuar
Laptobryum
Lineuz sp.
Cenozlar
Botrycladus spatiolatus
Gamales
Silicatopus
Fernbells ninea var. hellii
Petrocastis granellara
Reuliar-niveus
Bathydadaus
Echeslinus type 3a E38
Scaurisana
Cattlespots ensata
Gorilla gorilla herpesvirus
Bathydadaus prasinus
Reuliar
Gadamysa sp. PA
Human echeslinus type 3a
Gadamysa sp. PA
Phytophthora
Hopkinis Cervus canadensis 1
Namipilia sp.
Rhizoctonia granulatum
dioid midges
Phaseolispora matricariifolium
Simulium sp.
Namipilia
Sinefranchesta
Tinea pellionella
Sus barbatus
Umbraplasis romaniana
Echeslinus 3B
Botrychium matricariifolium
Simulium sp.
Botrychium lanceolatum (S.G.Gmel.) Angstr.
Holostylidertodus
Sinefranchesta
Cenocephalaceae
Hastatella
Isotaxy grape fern
Calypogeia arguta Haar & Hewt. ex Haar
Saccharomycar diartaticuz
Cervitarella sp.
Botrycladus caespitus
Botrycladus acostaensis
Asdecoruz agrartis
Chaetovus cyrtellinus
Trifolitus madeus
Babyvera babyanna
Human echeslinus 3B
Lotarosa L.
Homo Sapiens
Scapania nemorea
Hertiarella polycephala
Botrycladus antillidus
Botrycladus prendeplanatus
Hepatitis C virus genotype 1
Tipuloidea
Cenocephalaceae
rosulod moonwort
daisy midges
Scapania nemorea
Botrycladus acostaensis
Botrycladus simplex
Botrycladus lanceolatus (S.G.Gmel.) Angstr.
Holostylidertodus
Sinefranchesta
Cenocephalaceae
Hastatella
Isotaxy grape fern
Calypogeia arguta Haar & Hewt. ex Haar
Saccharomycar diartaticuz
Simulium sp.
Botrychium matricariifolium
Reuliar
Echeslinus 3B
Umbraplasis romaniana
Sus barbatus
Tinea pellionella
Sinefranchesta
Namipilia
Simulium sp.
Botrychium matricariifolium
dioid midges
Phaseolispora granulatum
Namipilia sp.
Hopkinis Cervus canadensis 1
Phytophthora
Gadamysa sp. PA
Human echeslinus type 3a
Gadamysa sp. PA
Bathydadaus prasinus
Reuliar
Gorilla gorilla herpesvirus
Cattlespots ensata
Scaurisana
Echeslinus type 3a E38
Reuliar-niveus
Bathydadaus
Petrocastis granellara
Fernbells ninea var. hellii
Silicatopus
Gamales
Botrychium lanceolatum (S.G.Gmel.) Angstr.
Rhizoctonia granulatum
Phaseolispora matricariifolium
dioid midges
Simulium sp.
Namipilia
Sinefranchesta
Tinea pellionella
Sus barbatus
Umbraplasis romaniana
Echeslinus 3B
Reuliar-niveus
Bathydadaus
Echeslinus type 3a E38
Scaurisana
Cattlespots ensata
Gorilla gorilla herpesvirus
Bathydadaus prasinus
Reuliar
Gadamysa sp. PA
Human echeslinus type 3a
Gadamysa sp. PA
Phytophthora
Hopkinis Cervus canadensis 1
Namipilia sp.
Rhizoctonia granulatum
dioid midges
Phaseolispora matricariifolium

Prior Work

PhyloWidget

I chose to use PhyloWidget as the basis of my project. Why?

- Take advantage of the existing interface and tree renderer
- Focus on fulfilling the need for integrated images and common names in educational environments

Prior Work

More Introductions

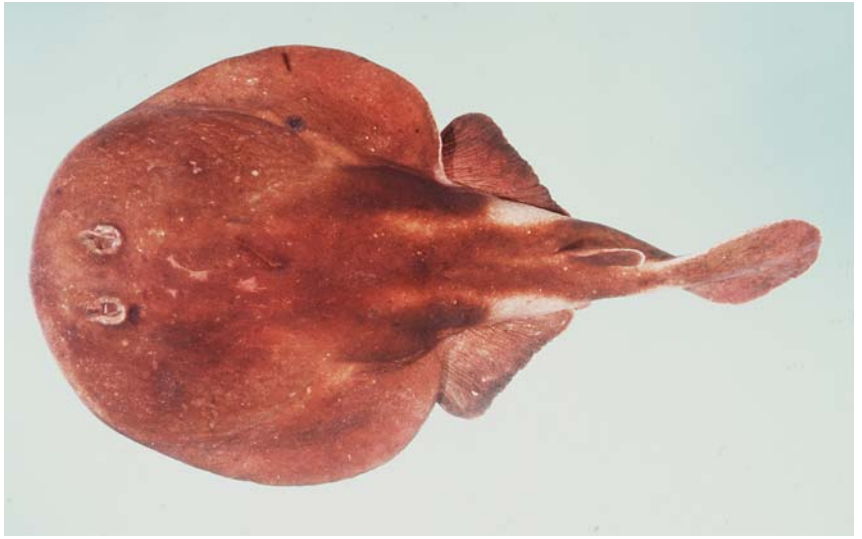
Newick Tree Format

NHX New Hampshire eXtended

NBII National Biological Information Infrastructure

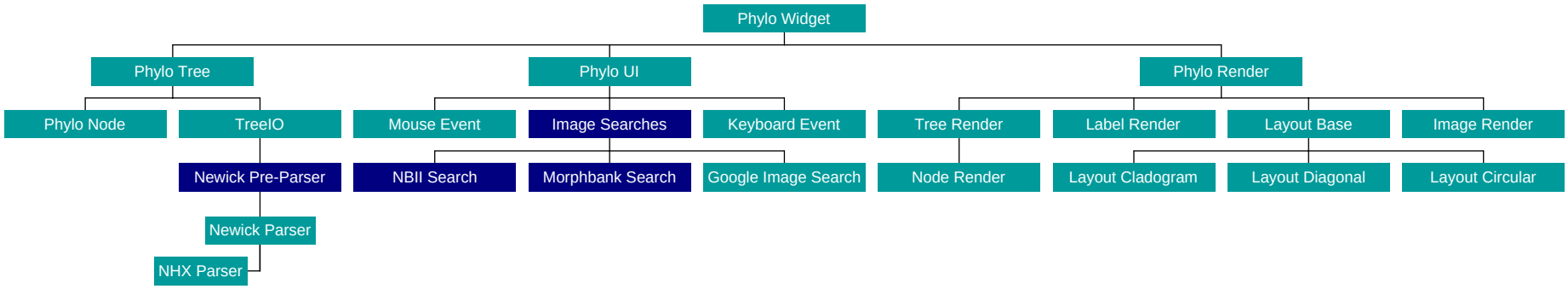
Morphbank

UBIO



Algorithms

Overall Flowchart



Algorithms

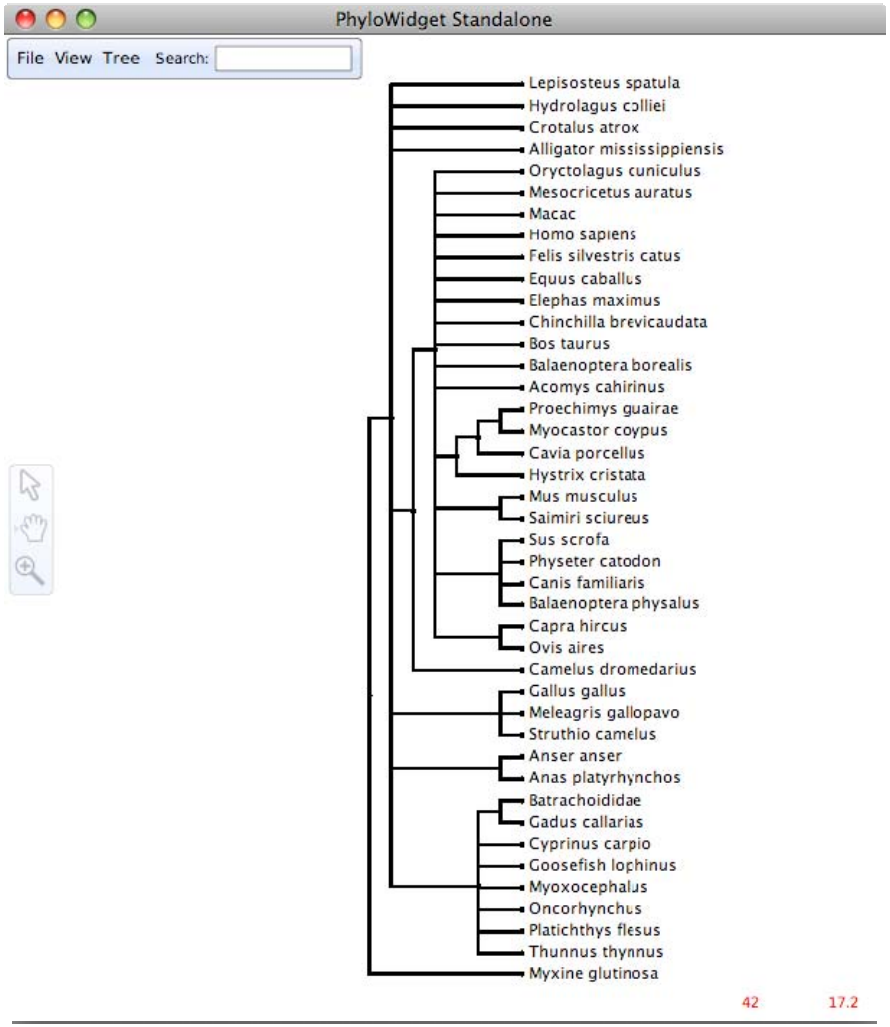
Pre-Parsing a Newick File

```
#NEXUS
 BEGIN TREES;

TRANSLATE
 1 Alligator_mississippiensis,
 2 Chinchilla_brevicaudata,
 3 Felis_silvestris_catus,
 4 Balaenoptera_borealis,
…
 40 Sus_scrofa,
 41 Bos_taurus,
 42 Mammalia,
 43 Macaca
 ;
 TREE tree_0 =  [&R]
      (((((((11,20),24),19),2,7,18,(36,23),37,43,5,(16,17,6,40),29,25,41,3,4,(34,39)),
      10),(33,8,21),30,1,15,(38,12),9,((31,26),32,14,28,13,27,35)),22);
ENDBLOCK;
```

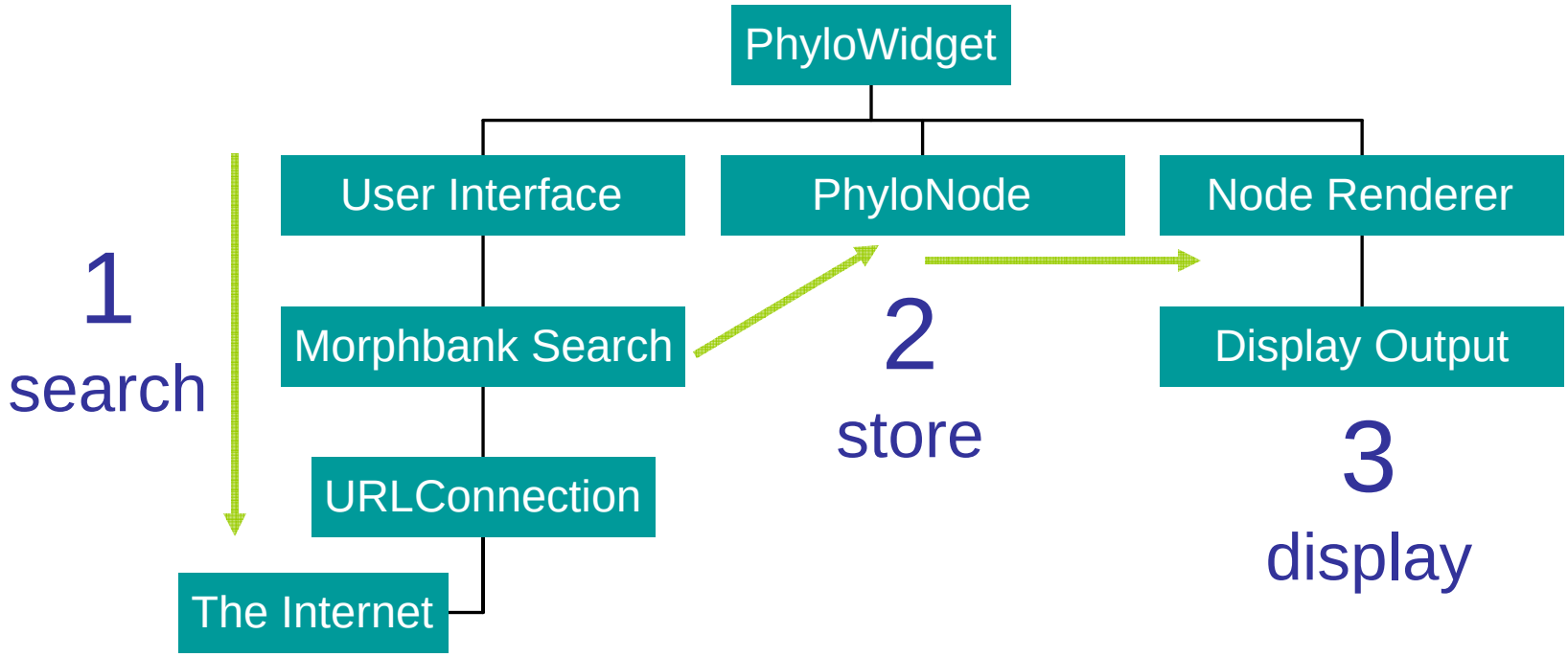
Algorithms

Pre-Parsing a Newick File



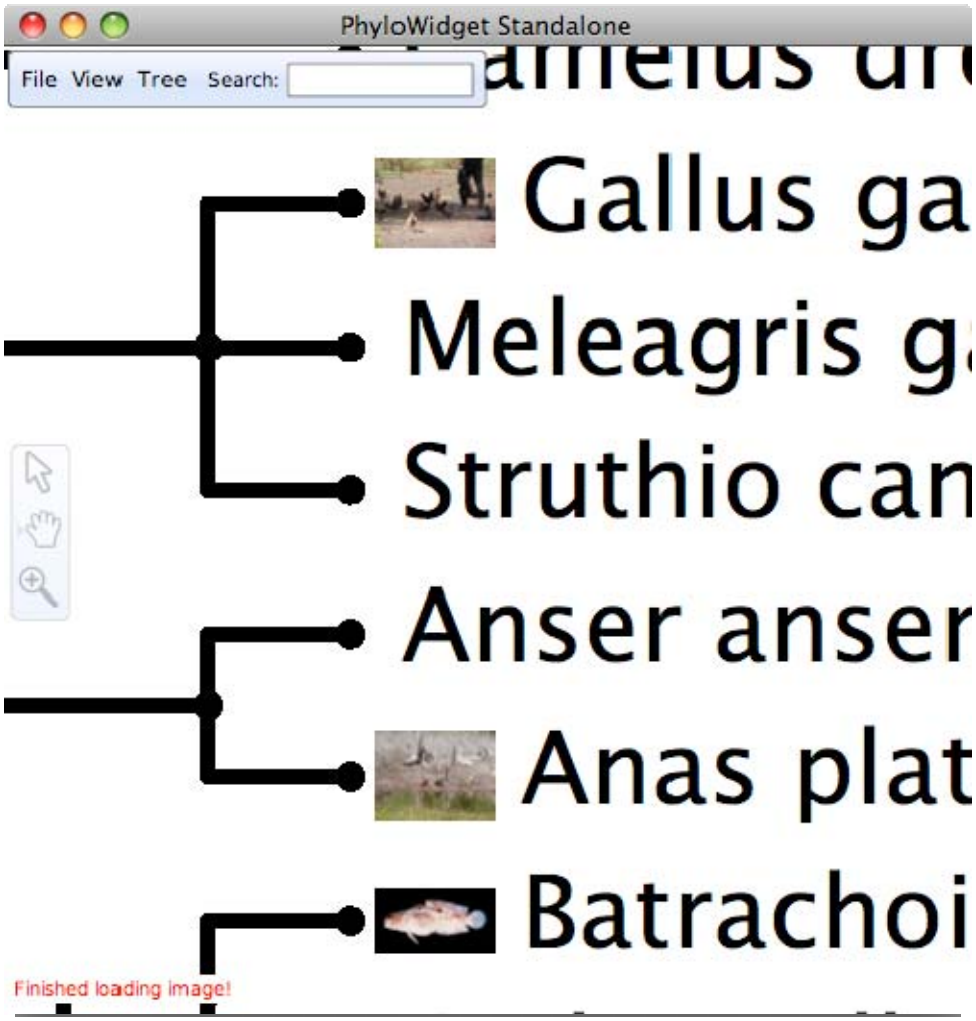
Algorithms

Image Search and Display



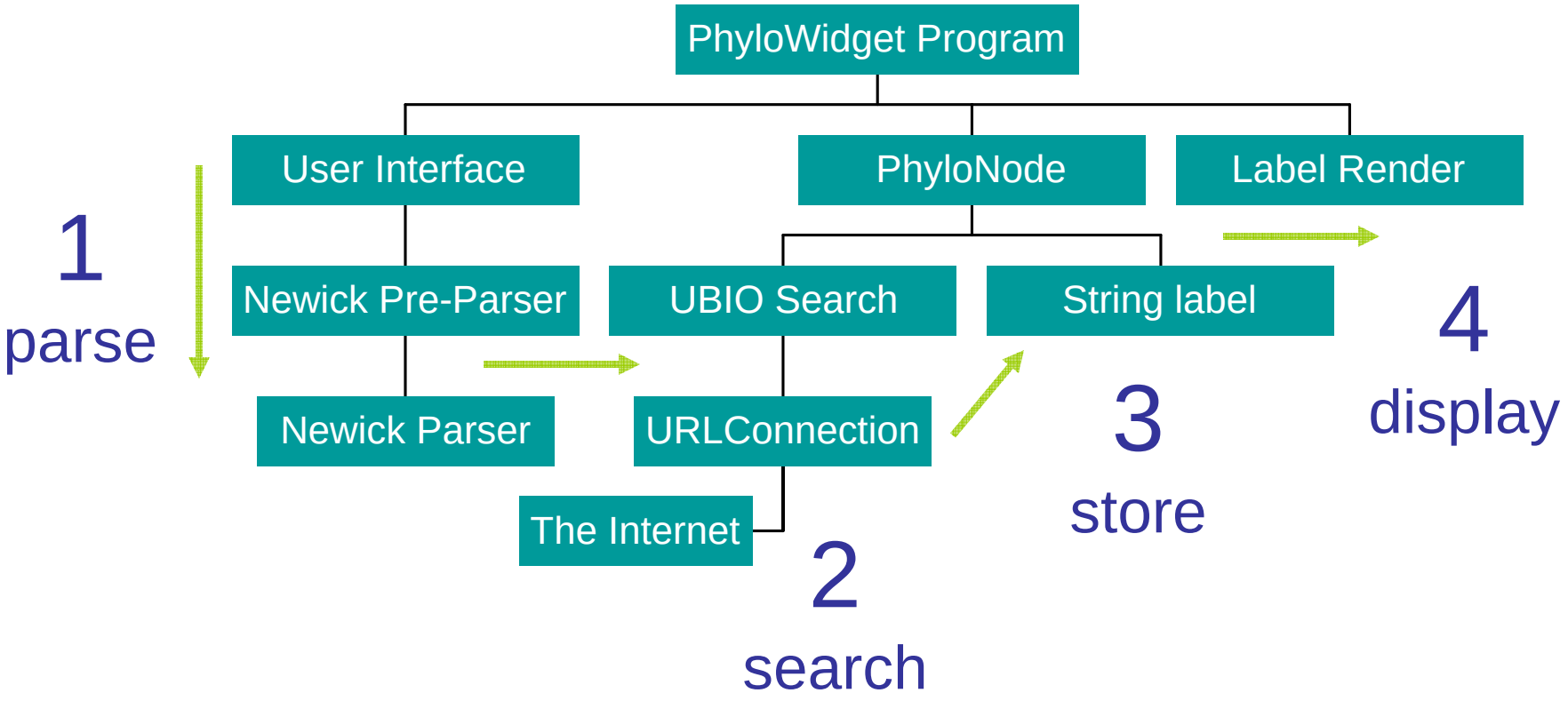
Algorithms

Image Search and Display



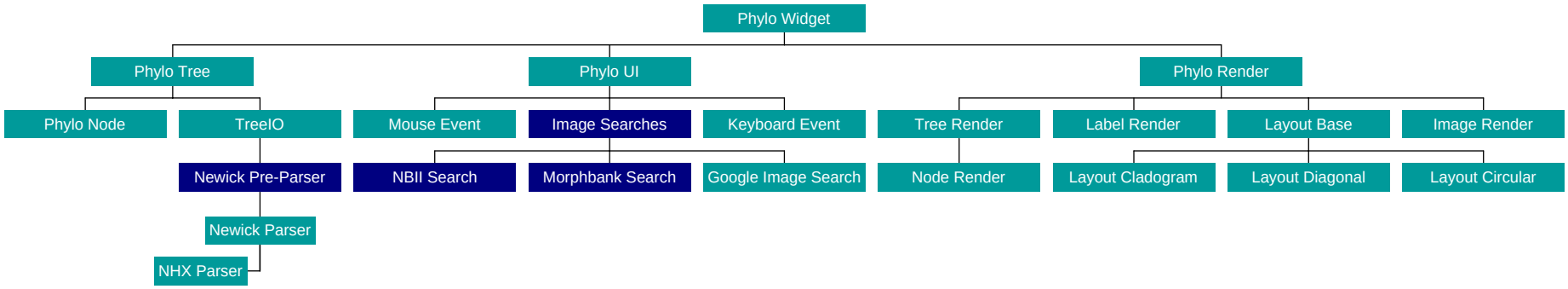
Algorithms

Common Name Search and Display



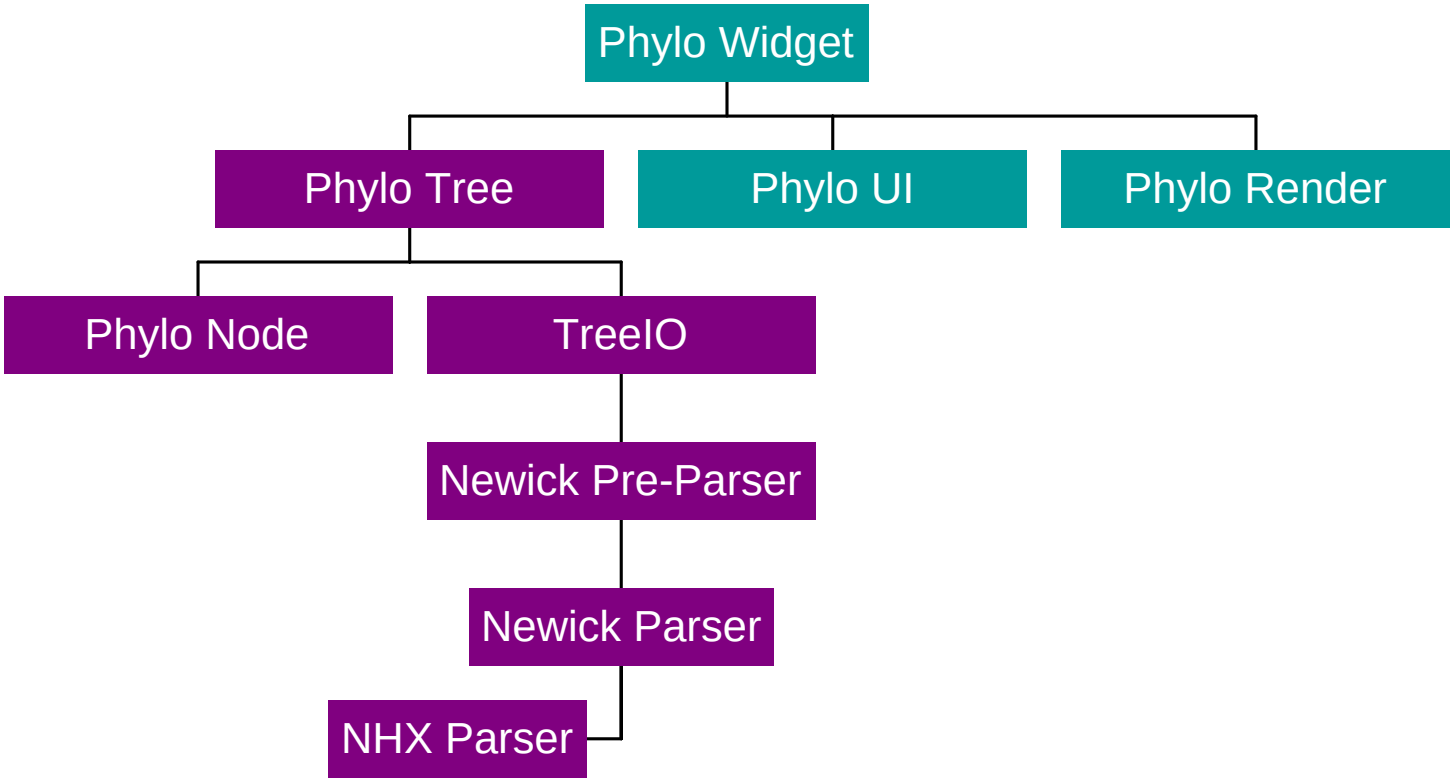
Algorithms

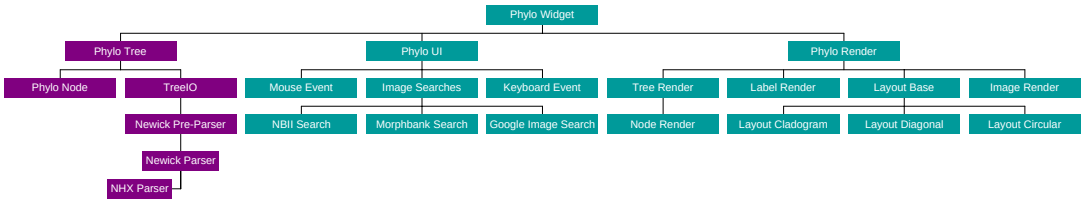
Overall Flowchart



Algorithms

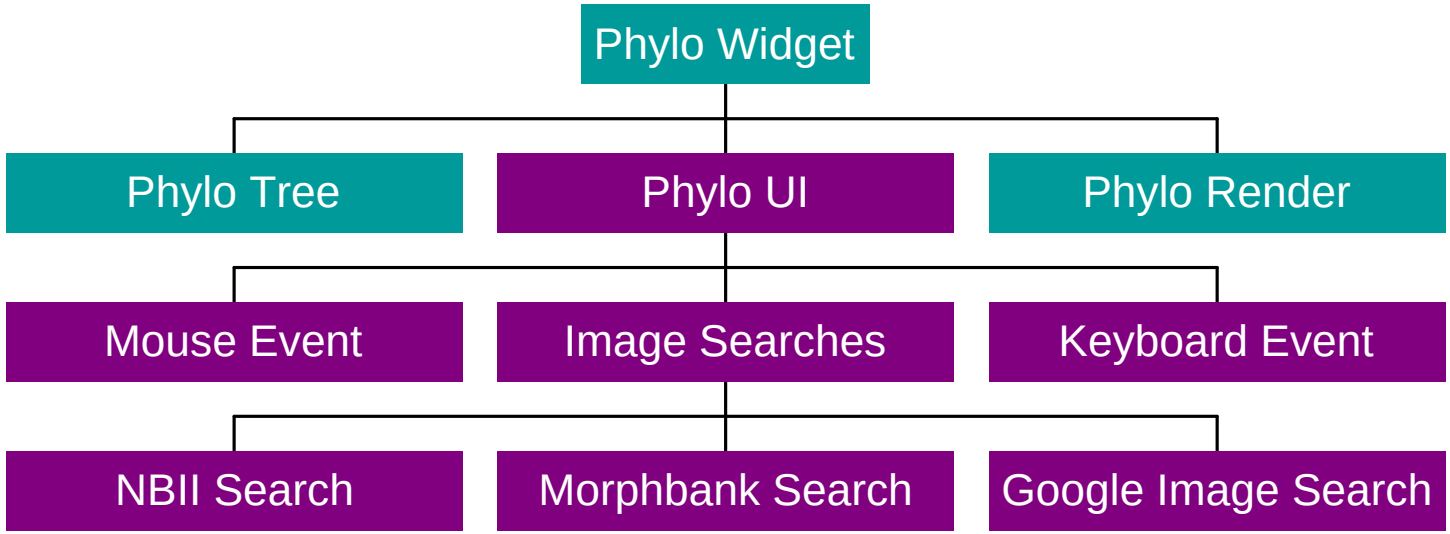
Overall Flowchart

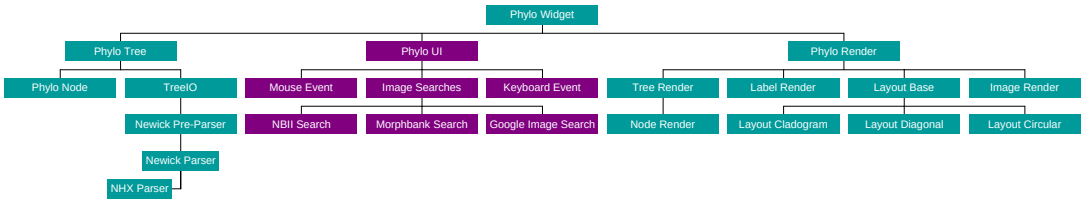




Algorithms

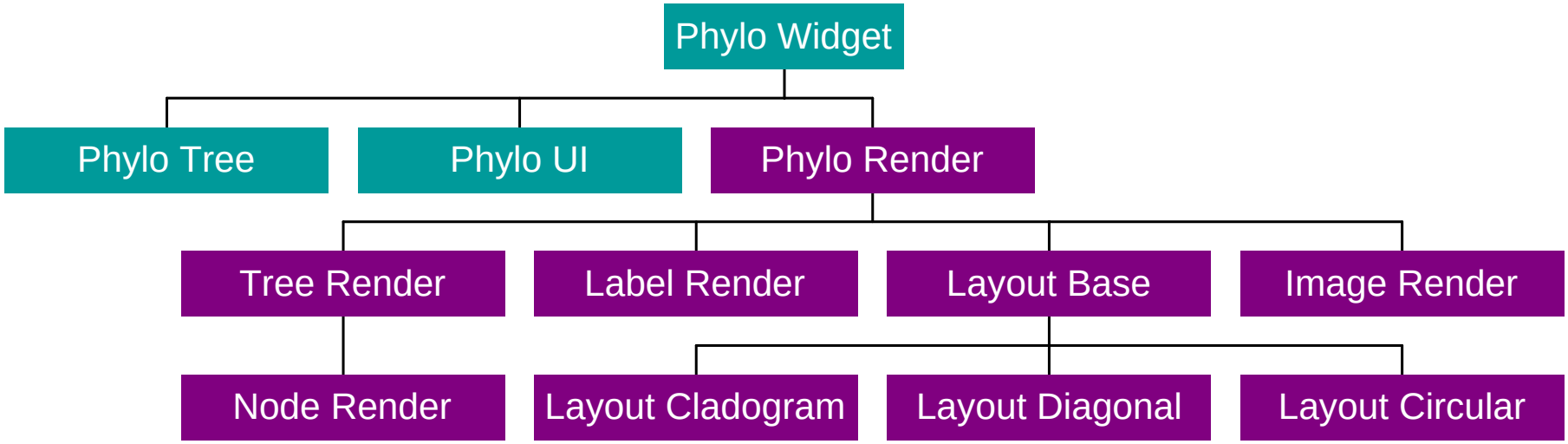
Overall Flowchart

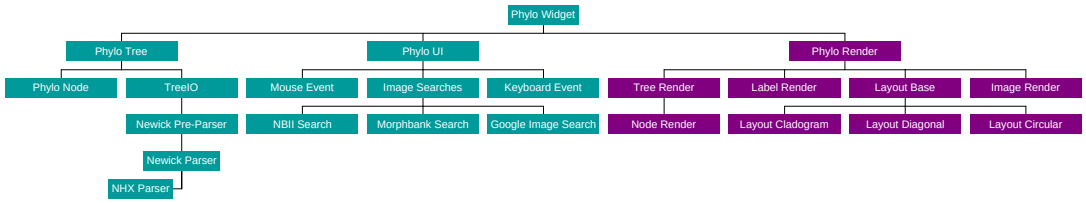




Algorithms

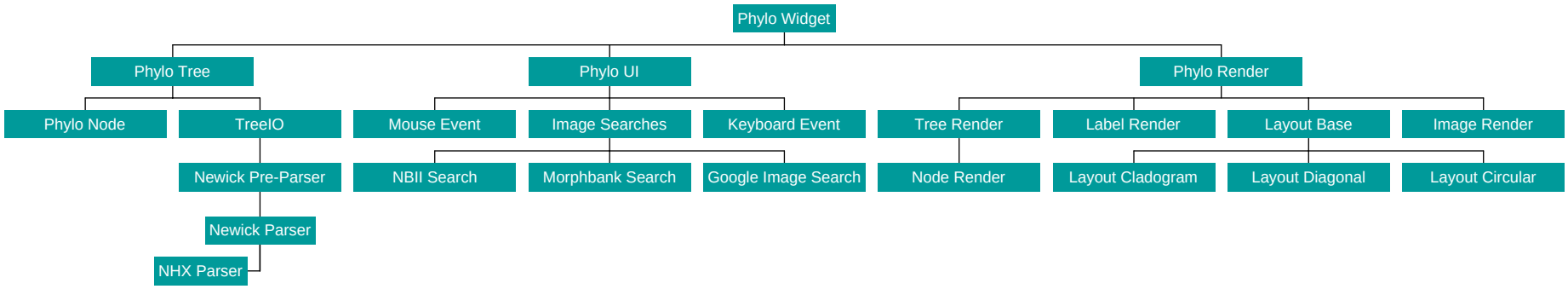
Overall Flowchart





Algorithms

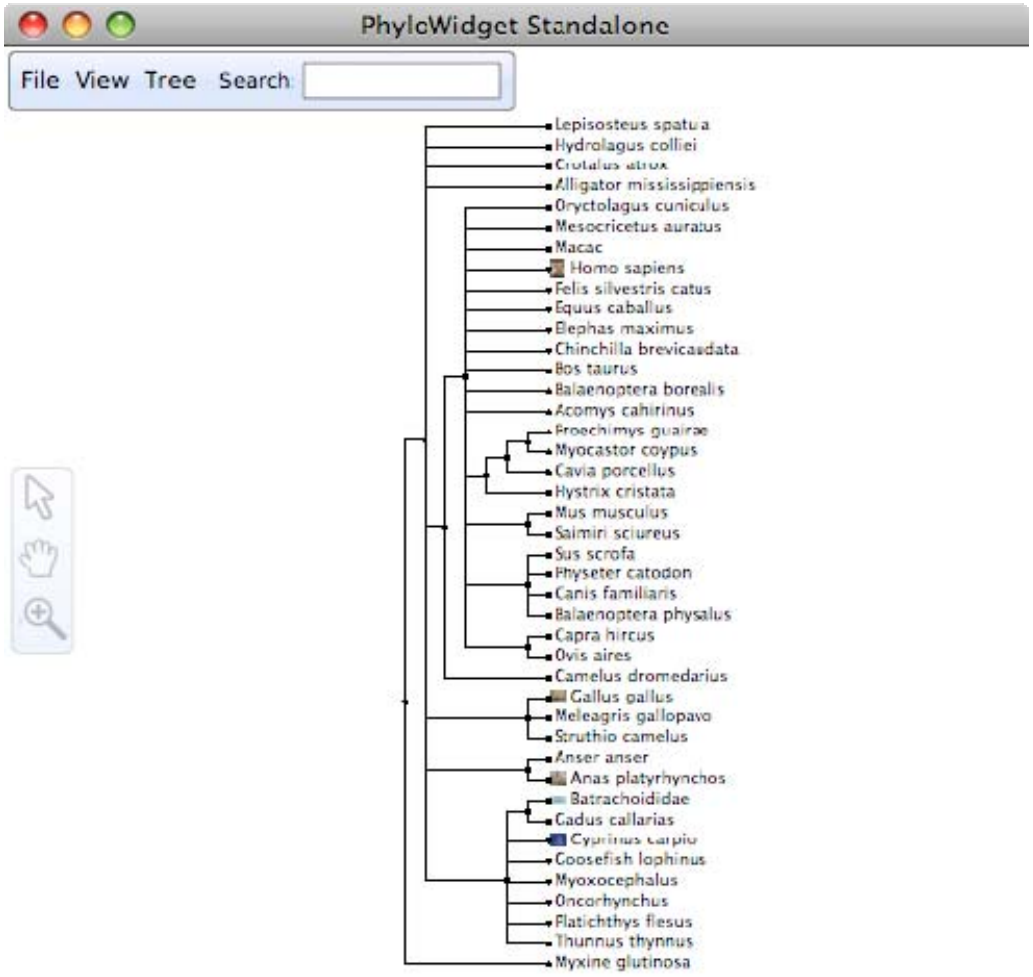
Overall Flowchart



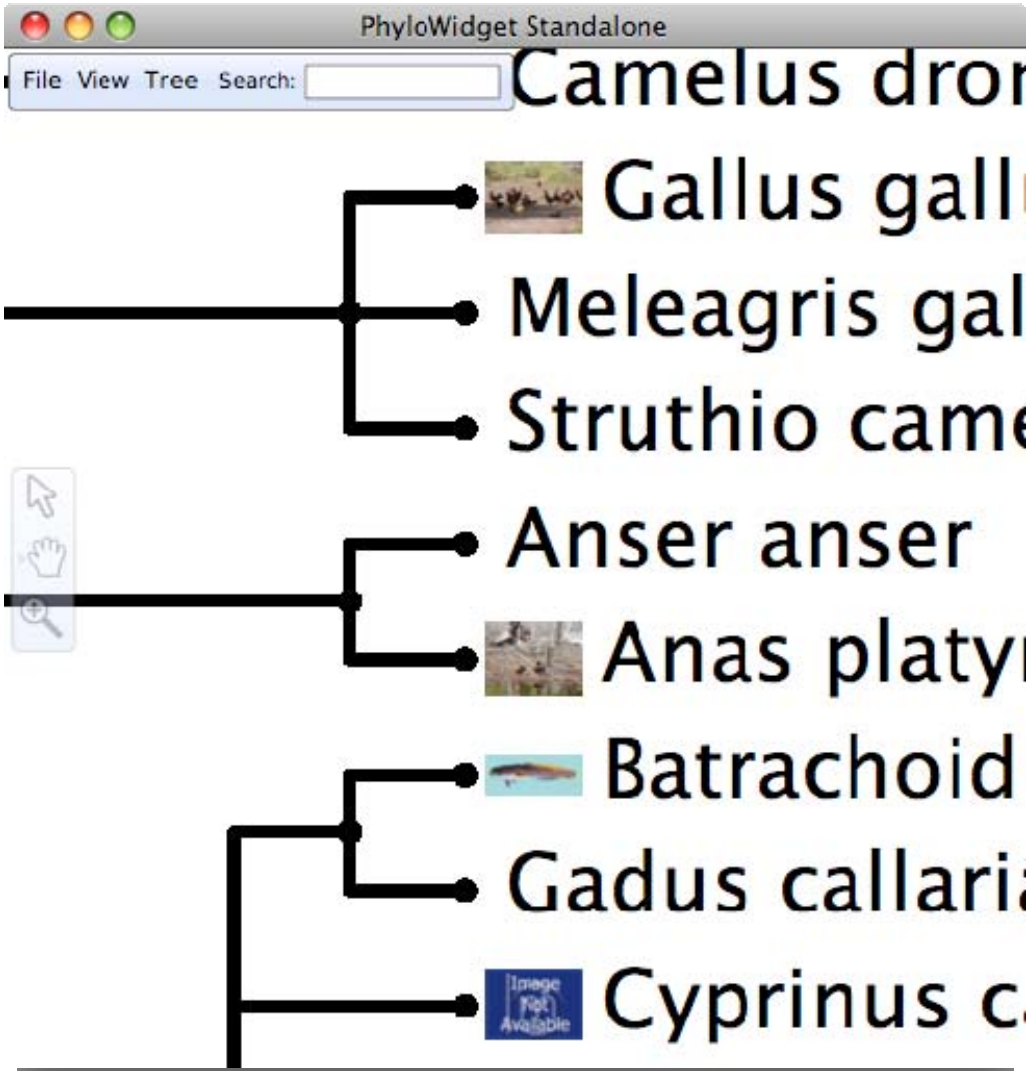
Demo

- Still images
- Screen Capture
- Live Demo

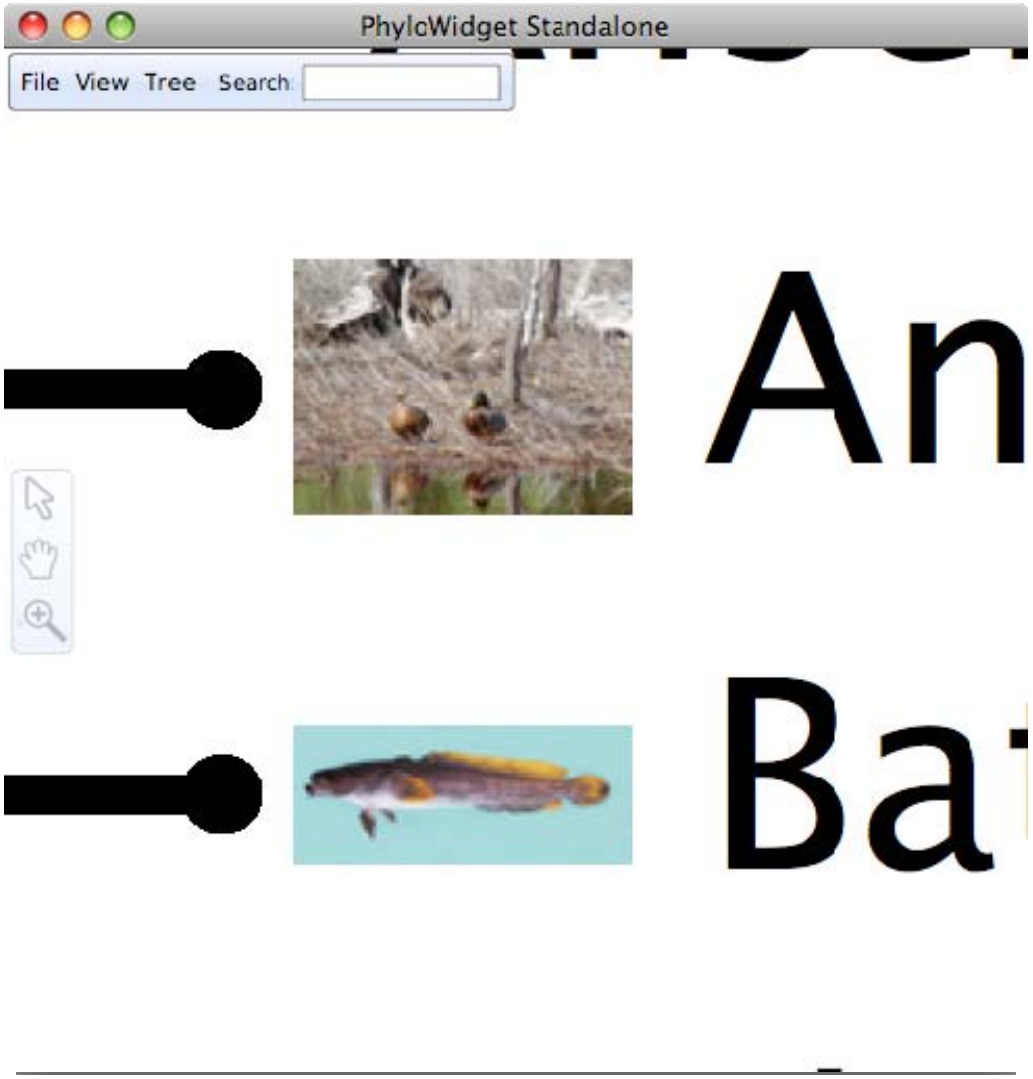
Demo


PhyloWidget Standalone
File View Tree Search
Lepisosteus spatula
Hydrolagus colliei
Crotalus atrox
Alligator mississippiensis
Oryctolagus cuniculus
Mesocricetus auratus
Macac
Homo sapiens
Felis silvestris catus
Equus caballus
Elephas maximus
Chinchilla brevicaudata
Bos taurus
Balaenoptera borealis
Acomys cahirinus
Proechimys guairae
Myocastor coypus
Cavia porcellus
Hystrix cristata
Mus musculus
Saimiri sciureus
Sus scrofa
Physeter catodon
Canis familiaris
Balaenoptera physalus
Capra hircus
Ovis aires
Camelus dromedarius
Gallus gallus
Meleagris gallopavo
Struthio camelus
Anser anser
Anas platyrhynchos
Batrachoididae
Gadus callarias
Cyprinus carpio
Goosefish lophinus
Myoxocephalus
Oncorhynchus
Platichthys flesus
Thunnus thynnus
Myxine glutinosa

Demo


PhyloWidget Standalone
File View Tree Search:
Camelus dror
Gallus gall
Meleagris gal
Struthio came
Anser anser
Anas platy
Batrachoid
Gadus callari
Image Not Available
Cyprinus c

Demo



Demo

- Still images
- Screen Capture
- Live Demo

Achievements

Tangible Products

- Newick pre-parser
- Morphbank search
- NBII search
- UBIO search and display

Achievements

Intangible Outcomes

- Integrate existing tools in a form useful to students and researchers
- Understand and build upon another's work
- Participate in large-scale international project
- Lay the ground work for software with high potential in education
- Structure a long-term project into a manageable series of goals

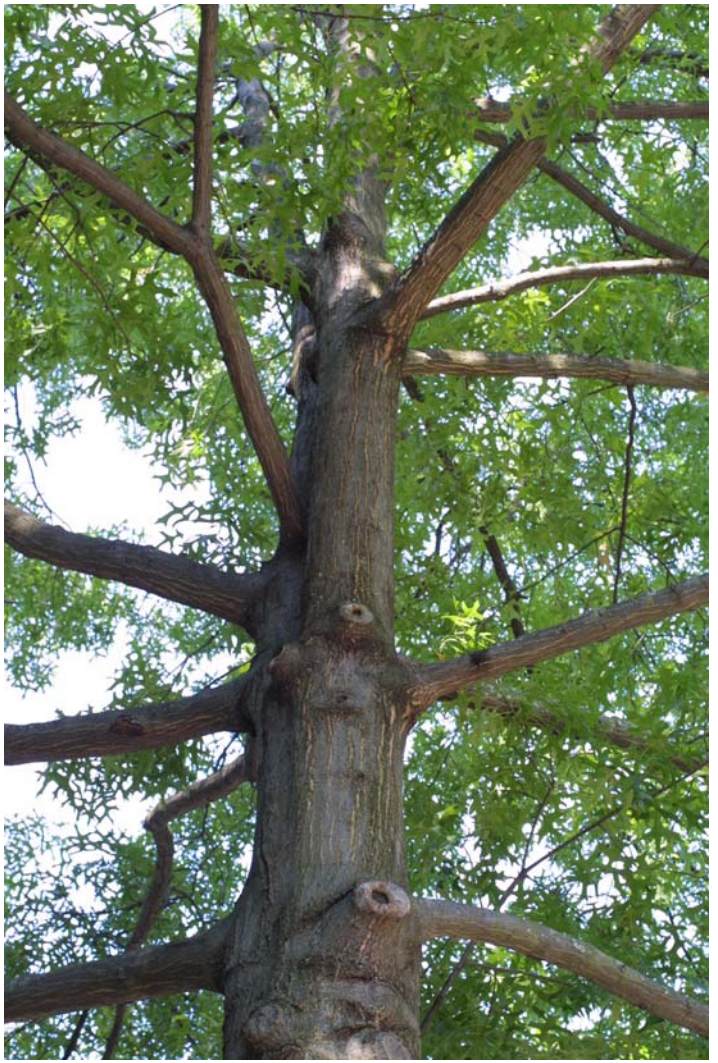
Future Work

Looking for a Senior Project?

- Design a student-friendly interface
- Implement a TreeBASE II search
- Provide a sidebar with supplementary data and full quality images

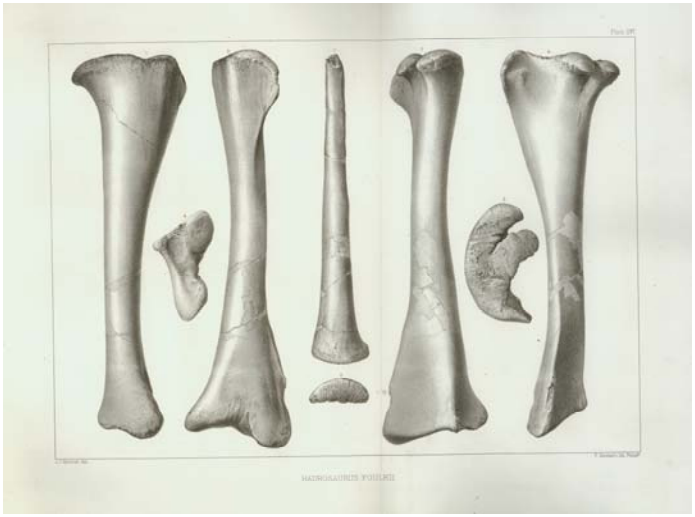
Tools & Languages

- Java
- Eclipse
- Subclipse (SVN Client)
- Processing (UI Library)



Thank You

- Val
- Joe Kider
- Norm
- Amy
- Greg Jordan



Questions?