

“Tree of Life” Teaching Tool

CIS 499 SENIOR PROJECT DESIGN DOCUMENT

Madeline Yasner
Advisor: Val Tannen
University of Pennsylvania

PROJECT ABSTRACT

The Tree of Life Web Project is an endeavor to catalogue and classify data about “biodiversity and the evolutionary relationships of all organisms”¹. Originally created for biologists, the Tree of Life has become popular to educators, students, and anyone searching for information about a specific group of organisms or about the diversity of life. One of the goals of the project is to “aid learning about and appreciation of biological diversity and the evolutionary Tree of Life”². It is within the scope of this goal that my project resides.

I will create a software application to allow students to explore both simply and thoroughly the every-growing catalogue of all living things, compiled by the Tree of Life Web Project. This will involve the integration of phylogenetic (evolutionary) tree data with image data of living organisms, followed by the visualization of these two forms of data in helpful ways. Users will be able to correlate the relevant data from both sources and interact with the program’s visualization to edit or manipulate the data representation.

Project blog:

<http://adventuresofmaddy.blogspot.com>

1. INTRODUCTION

My project will use the information stored in the Tree of Life Web Project database, combined with image data for living organisms, and enable students to explore this vast amount of information in intuitive and constructive ways.

1.1. Significance of Problem or Production/Development Need

The incredible amount of data organized by the Tree of Life project is only as valuable as a user’s ability to interact with it and learn from it. My goal is to create a software tool

¹ “Goals of the Tree of Life,” http://tolweb.org/tree/home_pages/goals.html

² *ibid.*

that will allow students to make use of this great body of knowledge in a way that is most suited to their needs. The ability to individually explore this vast store of information will greatly benefit students' understanding and appreciation in hopes of stimulating their interest in evolution and the sciences in general. In addition, visualization of such data will be helpful in understanding all forms of modern biological research, not only for students but for anyone from amateurs to educators to the researchers themselves.

1.2. Technology

I am currently deciding to what extent I will incorporate existing code from the PhyloWidget application. I plan to code in Java so as to make my program as transportable as possible. One of the additional decisions I'll need to make involves selecting which libraries I'll use to create my interface and visualization components.

The program will be created using a modular approach, so that user interaction and visualization is completely separated from internal structures and data integration. The two halves will communicate via a controller.

One clear challenge for me will be the requisite amount of data integration that I will need to do in order to present fresh data from the ever-changing databases used to store the information compiled by the Tree of Life project.

1.3. Design Goals

Our project will work to find a comprehensive way to interact with the data including both visual and textual information in a way that encourages exploration in a variety of directions while maintaining a simple interface structure for ease of navigation.

1.3.1 Target Audience.

Students and teachers of biology at many levels should be able to interact with this program and gain new exposure to evolutionary data.

1.3.2 User goals and objectives

The user should be able to interact with the program to search for images and information about desired living things, with the goal of seeing the visualization of the relevant tree and interacting with it to obtain information or make more queries.

1.3.3 Project features and functionality

The program should provide access to:

- Pictures of as many of the living things as are available.
- Scientific as well as common names, in addition to a brief description.
- Diagrams of the selected organisms and their connections on the Tree.

Features will be tweaked and added to as time allows, once the basic program is functioning correctly.

2. Prior Work

My work on this project is largely guided by two previous programs for phylogenetic tree visualization. One is Tamara Munzner's TreeJuxtaposer. The other is Gregory Jordan and Bill Piel's PhyloWidget.

Tamara Munzner is an associate professor at the University of British Columbia, and she researches information visualization, with specific work on phylogenetic trees. Her TreeJuxtaposer program deals with both visualization and structural comparison of trees, although the scope of my project is limited to the visualization aspect. Of note are her accomplishments in the following areas:

- Scalability in Tree and Display Size
- Guaranteed Visibility of landmark nodes, regardless of user's navigation. While her goal in doing this is for the sake of comparison, it would be of additional value to my program, where students may want to focus on the relationships between specific organisms.
- Limited occlusion of other nodes due to labels

PhyloWidget is a program created for "viewing, editing, and publishing phylogenetic trees online"³. It was created by Gregory Jordan under the mentorship of Bill Piel, the Associate Director of Evolutionary Bioinformatics at Yale University's Peabody Museum of Natural History. It takes as input a tree formatted using the Newick tree format and generates an editable and customizable visualization of the tree.

3. PROJECT DEVELOPMENT APPROACH

3.1. Algorithm Details

This section will be developed further along in the semester, after more research is completed and design decisions are made.

3.2. Target Platforms

3.2.1 Hardware: independent

Because I will be programming in Java, I intend for my application to work on any machine that is capable of running Java-based programs.

3.2.2 Software

Java, using Eclipse

UI program and other libraries to be determined at a future date

³ <http://www.phylowidget.org/>

4. WORK PLAN

4.1.1. Project Milestone Report (Alpha Version)

- Completed Lit review and updated Design Doc
- Finalized decisions about programming language/platform, UI library
- Software design plan
- UI functionality plan

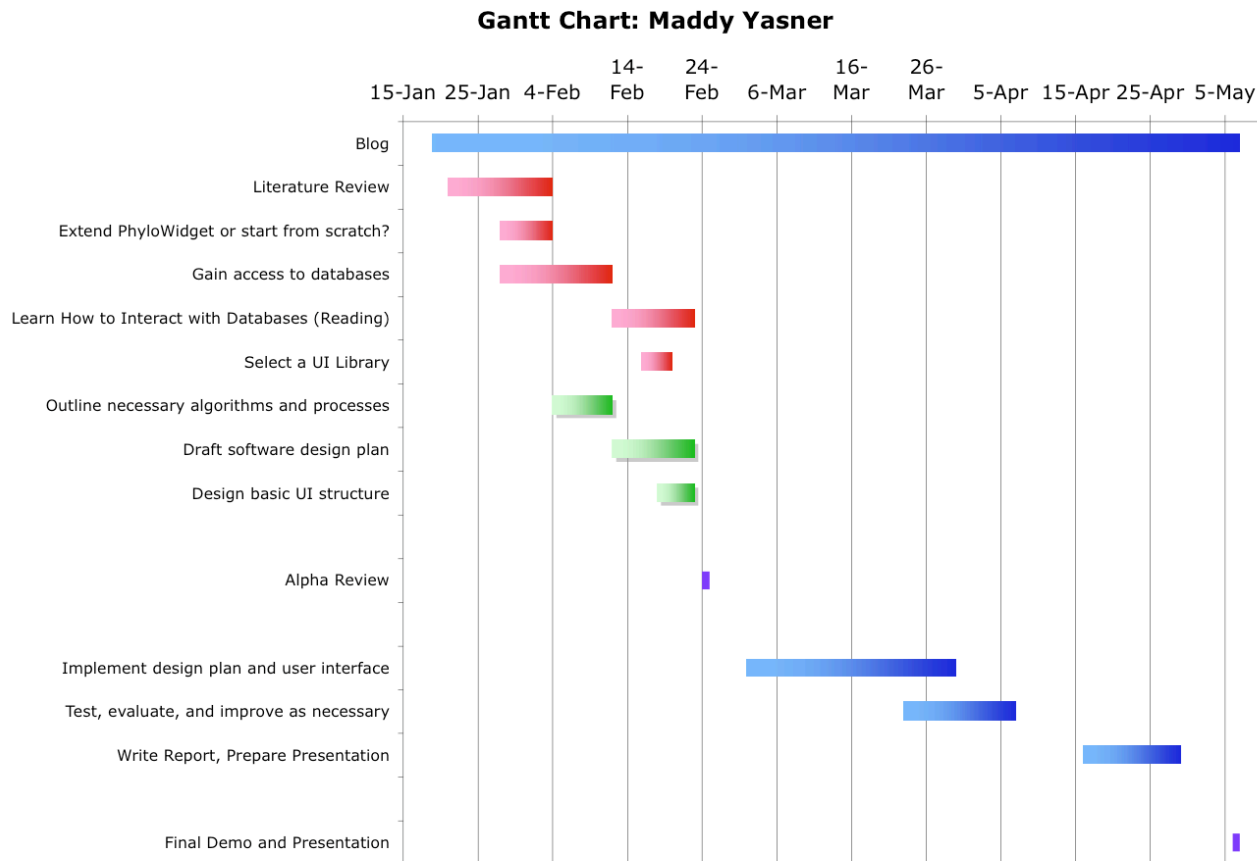
4.1.2. Project Final Deliverables

Workable software application with complete data integration and visualization components as well as functional internal model and controller.

4.1.3 Project timeline.

- Complete Literature Review
 - Choose whether to extend PhyloWidget or start from scratch
 - Outline necessary algorithms and processes
 - Draft software design plan
 - Gain access to databases
 - Select a UI Library
 - Design basic UI structure
-
- Alpha Review
-
- Implement design plan and user interface
 - Test, evaluate, and improve as necessary

4.1.4 Gantt Chart



Continuous activities:

- Progress blog updates
- Continued familiarization with Eclipse, PhyloWidget code, Processing (Graphics Library), databases, etc.

5. REFERENCES

Jordan, Gregory E. and William H. Piel, "PhyloWidget: web-based visualizations for the tree of life," *Bioinformatics* 2008 24: 1641-1642.

Maddison, D. R. and K.-S. Schulz (eds.), *The Tree of Life Web Project*, 2007, Accessed on Feb 1, 2009 at <http://tolweb.org>

Munzner, Tamara, et. al. "Tree Juxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility," *Proc. SIGGRAPH 2003*, published as *ACM Transactions on Graphics* 22(3), pages 453—462. Accessed on Jan 30, 2009 at <http://www.cs.ubc.ca/labs/imager/tr/2003/tj/tj.camready.pdf>.