

# CDAO-STORE: A New Vision for Data Integration

Brandon Chisham, Trung Le, Enrico Pontelli,  
Tran Son, Ben Wright

IEvoBio 2010  
Portland, OR



# CDAO

## Comparative Data Analysis Ontology

provide semantics to the descriptions of data and transformations commonly found in the domain of phylogenetic inference.

enables the description of phylogenetic trees and associated character data matrices



# What We Did

## CDAO-Store

A repository providing a rich set of API's for querying phyloinformatics data.

## CDAO-Explorer

A visualization tool for viewing data sets stored in the repository.



# CDAO-Store Repository?

## What's in it?

TreeBASE dump dated January 2009

Also allows the importation of CDAO formatted files.

- To get your files into CDAO, we can translate NEXUS, PHYLIP, and MEGA into CDAO format.

Files can be exported in RDF/XML using CDAO terms



# Querying CDAO-Store

## PhyloWS

Retrieve data sets via name, tree identifier, taxon, or size.

Supports computing the minimum spanning clade or the nearest common ancestor of a set of taxa.

## Web-Based

Search for data sets by author or study

View data sets online by tree, taxon, algorithm, method, or size.

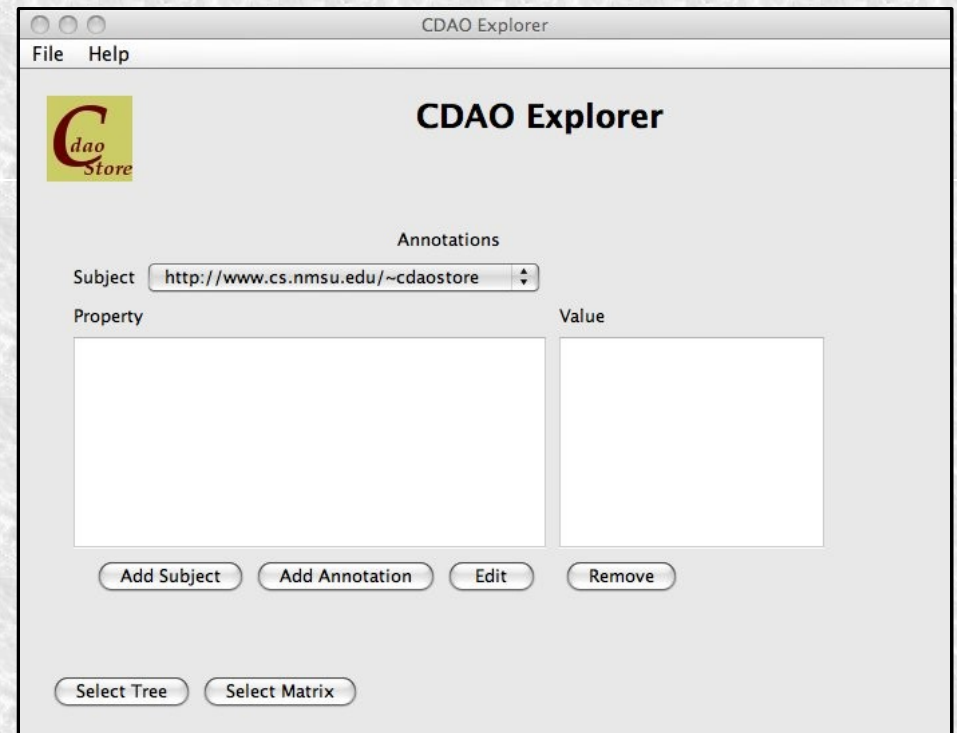


# Visualization

## CDAO-Explorer

Tree Viewer

Matrix Viewer



# Tree Viewer

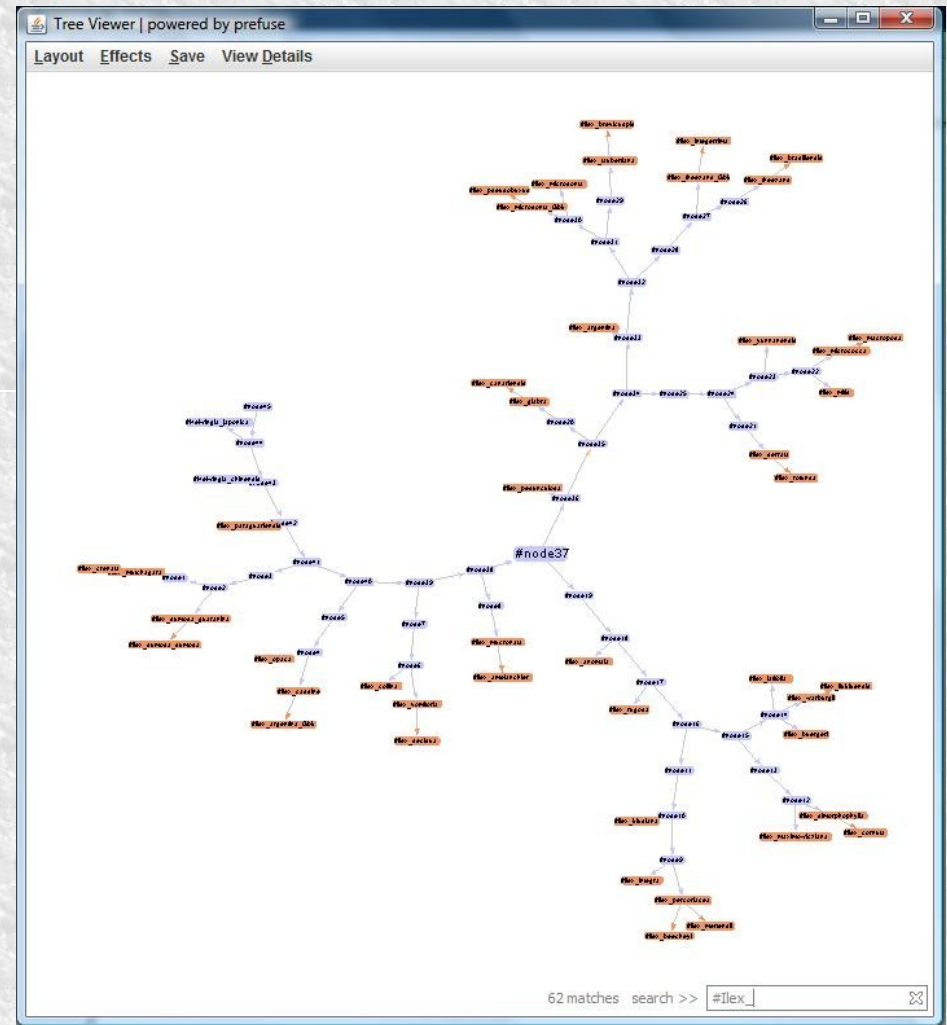
Uses the Prefuse framework

2 Layouts, “Force Layout” and “Node Layout”

Can search by node/edge name

View details of nodes or edges

Can save as jpg or png



# Matrix Viewer

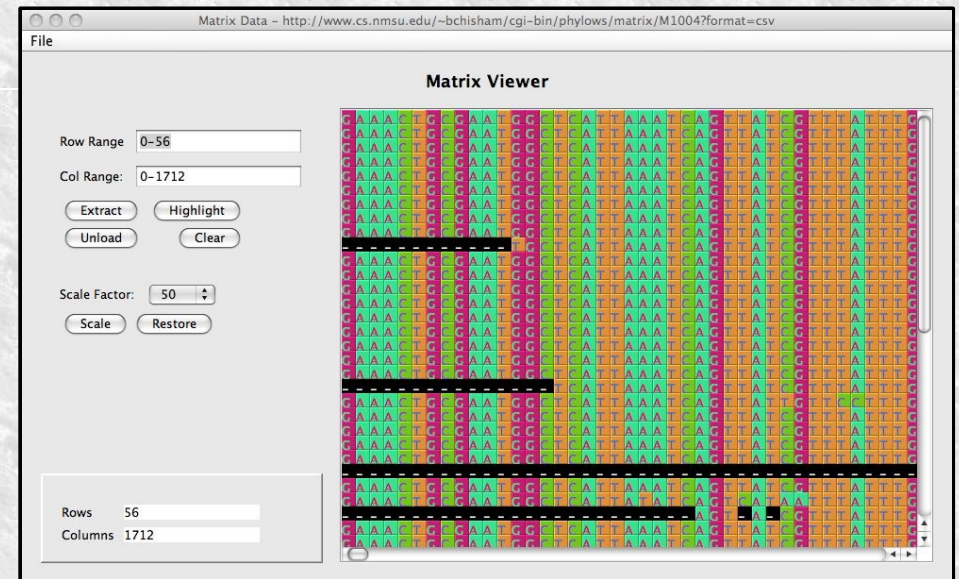
Custom built

Color-coded cells

Extract or 'crop' parts of the Matrix for closer views

Zoom in and out of the matrix

Annotation support in development.





# Conclusion

The CDAO-store tool set provides a robust foundation for a semantically aware, phylogeny resource

The CDAO-Explorer portion of the store has achieved a good base-line functionality and provides a set of useful features to advance the current state of visualization of large data sets in this field.



# Future

Annotations / MIAPA / OBI

User-defined SPARQL Queries

Better Tree / Matrix integration

Ambiguous Name Resolution (at taxon, tree, and study levels)

Integrating other stores besides TreeBASE



# Questions?

Find us at:

<http://www.cs.nmsu.edu/~cdaostore>

<http://cdaotools.sourceforge.net>

<http://www.twitter.com/cdaotools>

Funding for this project provided by:

NSF CREST grant HRD-0420407

NSF IGERT grant DGE-0504304

Additional Support provided by:

NESCent

NMSU

CDAO Development Team

