NLP cTakes Home

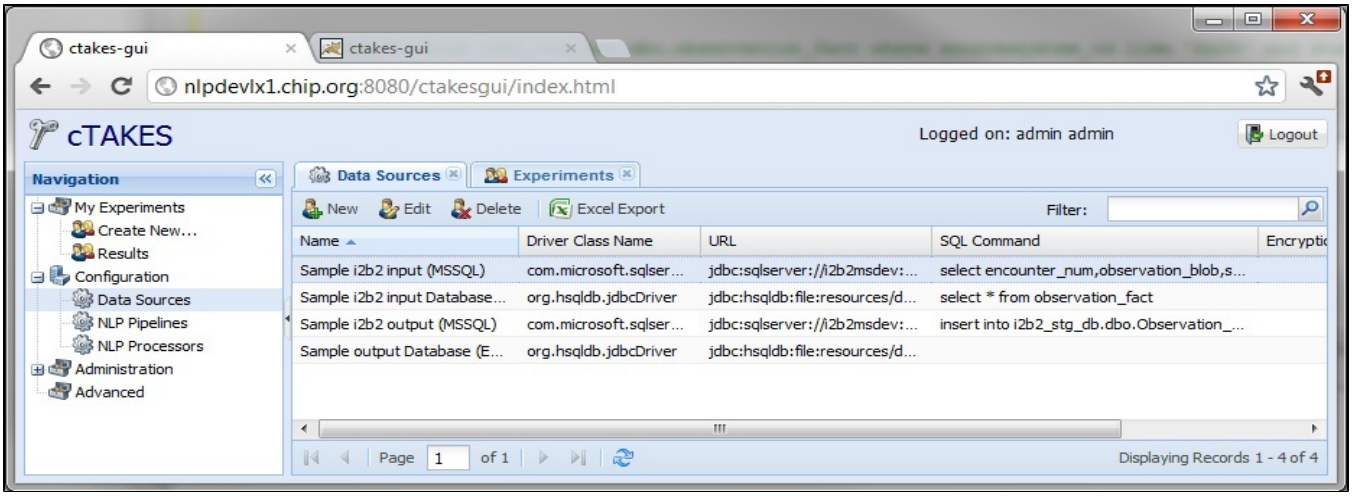
This is the home of the Related Project - NLP cTAKES Project space.

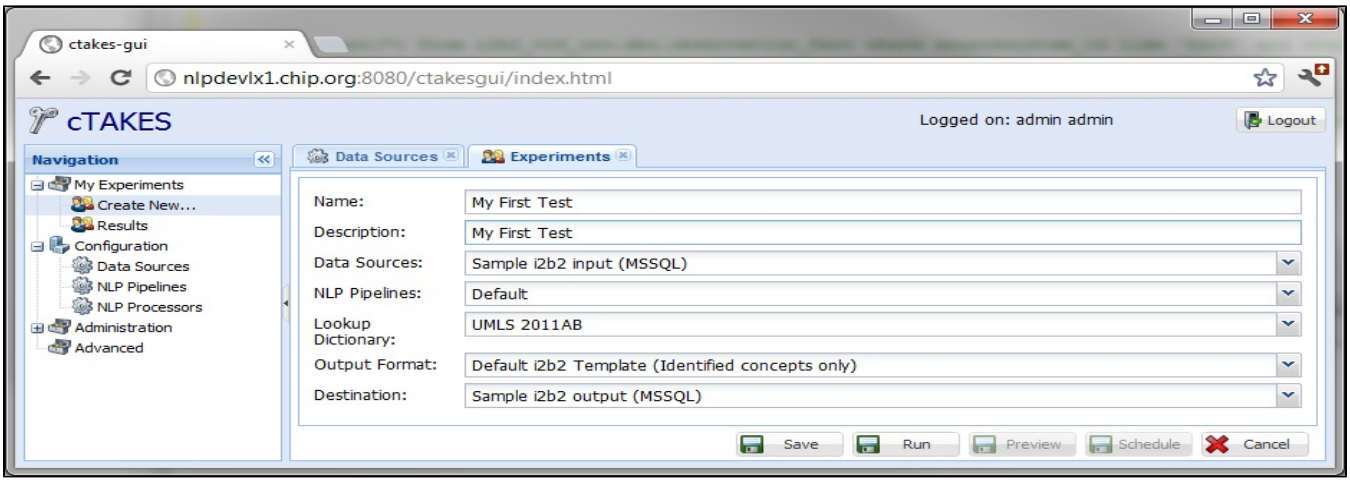
News: cTAKES is now a Top Level project on Apache Software Foundation !

There is a wealth of information within the plain text clinical narrative. The purpose of this cell is to harness the unstructured information by allowing i2b2 users to query and join that information with existing i2b2 concepts. Currently, the entire note is commonly stored as a single row in the observation_blob field in the observation_fact table in i2b2. One of NLP cTAKES' features is its capability to 'read' through and extract concepts from plain text notes and transform them into structured and normalized information. The purpose of this cell is to incorporate cTAKES and i2b2 by formatting the output of cTAKES into the i2b2 observation_fact table format (facts, concepts, modifiers, and values) which can then be easily queried by existing i2b2 interfaces.

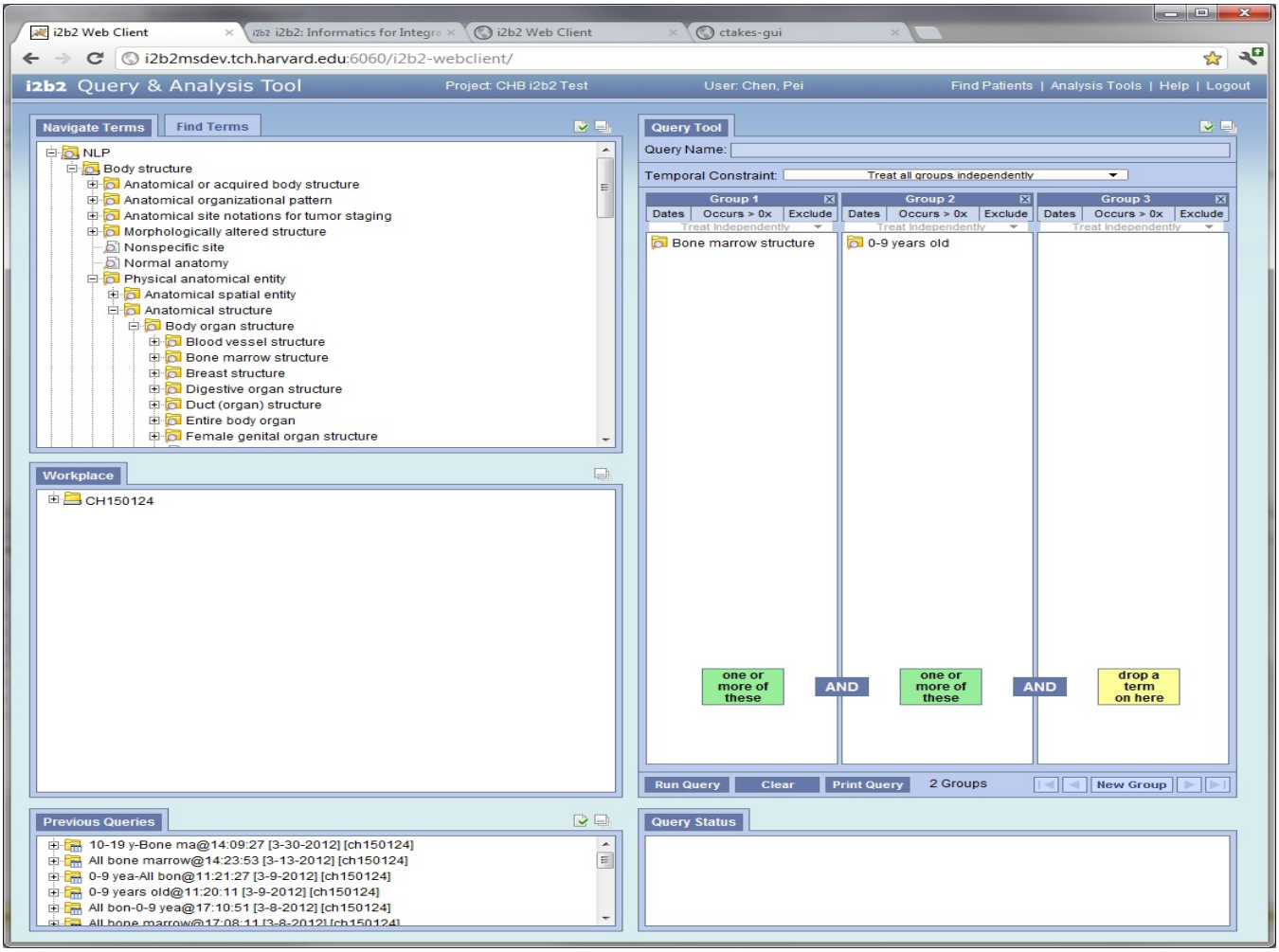
There will be 2 main components:

- An administrative tool (cTAKES GUI) that will allow users to specify the input DataSource of the note(s), the output of the notes(s), and the NLP pipeline to be used. The cTAKES GUI will be designed to be a web interface in a self contained package using Jetty- just run startup.sh or startup.bat. (However, it can also be packaged a war file to be easily deployed to other standard servlet containers such as Tomcat). The configuration information will be stored and could be reused for future experiments. The output in this case will be specially formatted in the observation_fact format.





- An interface for users to query the extracted data. We plan to reuse the existing i2b2 web client tool by adding an 'NLP' ontology which contains all of the concepts that could be used to filter and joined with other ontologies such as demographics or codified data.



Documentation

Quick Start Guide:

In addition to deploying the cTAKES GUI as a war file to a managed servlet container such as JBoss or Tomcat, cTAKES provides a standalone web server version running Jetty out of the box.

1. Java JDK (not JRE) 1.6+ is required. Feel free to download and install the Java JDK from: http://www.oracle.com/technetwork/java/javase/downloads/index.html
2. Download the cTAKES GUI code (currently in beta) from https://ohnlp.svn.sourceforge.net/svnroot/ohnlp/branches/cTAKES-GUI-0.0.1/ctakes-gui-0.0.1.zip
3. Unzip the contents to a local directory
4. If required, modify the startup.bat or startup.sh with the path to your JAVA_HOME environment variable. Also available as an environment variable is the JETTY_PORT (defaults to 9998).
5. Open a command line window/terminal and execute startup.sh or startup.bat
6. Wait for the message "Jetty running on http://localhost:9998/index.jsp"
7. Open your favorite web browser and point the url to: http://localhost:9998/index.jsp
8. cTAKES GUI Quick Reference Guide.pdf

*__Design and Architecture__
*__Installation Guide__

Contact Information

Pei J Chen
Children's Hospital Boston
E-mail: chenpei[at]apache.org
Guergana K Savova, PhD
Children's Hospital Boston
E-mail: guergana.savova[at]childrens.harvard.edu
Jonathan Bickel, MD
Children's Hospital Boston
E-mail: jonathan.bickel[at]childrens.harvard.edu