

NIH Summer Research Program 2007

POSTER TITLE: Gathering, Formatting and Using Biomedical Data from Disparate Sources to Reveal Medical Knowledge

MENTOR: Jim DeLeo, Chief, Scientific Computing Section, DCRI, NIHCC

NIH AFFILIATION: Scientific Computing Section, DCRI, NIHCC (2007 Summer Intern)

NON-NIH AFFILIATION: University of Maryland, Baltimore County (Undergraduate)

THE IDEA IN BRIEF:

My poster considers the construction of a small-scale NIH Biomedical Knowledge Discovery System. Our goal is to make it simple to use and reusable. The task is to merge a principal investigator's various disparate data sources into a single modern database and to provide a suite of application programs to visualize and analyze it from the computer scientist's perspective, with the hope that new medical knowledge may be revealed.

ABSTRACT:

A detailed systems overview of a newly designed NIH Clinical Center Medical Knowledge Discovery System is presented with emphasis on the achievement of merging data from the Clinical Research Information System (CRIS) and other disparate data sources into a single modern uniform database. Data sets, application programs, graphical user interface, and web-server interactions are described. The possibility of gaining new medical knowledge with this system is discussed and possibilities for leveraging and reusing this system for similar applications are suggested.

PURPOSE:

The ultimate goal of this project is to develop a basic medical knowledge discovery system that is capable of merging disparate but related data sets, graphically displaying selected data to the user, and performing statistical analysis and data mining on the newly created data set. The end result is a useful system that has the potential to reveal new medical knowledge and, in addition, can be easily reused and adapted to a new data set and principal investigator.