

# **Biofacets: A Facet-Based Integration System for Efficient Querying of Remote Scientific Databases**

**Date/Time:** Thursday, October 23, 2008, Noon-1:00pm

**Place:** SL 165 (723 West Michigan Street)

**Speaker:** Dr. Malika Mahoui, Assistant Professor, School of Informatics and Adjunct Professor, ECE

## **ABSTRACT**

Biological databases are an important component in bioinformatics studies. However with the large number of databases available and the wealth of data they host, the traditional techniques using keyword-based querying and browsing through the numerous results are not sufficient to leverage the value of the information they contain. Biofacets is an integration system for biological databases that provides for researchers a common interface for querying through multiple online databases with biological facets as a mechanism to restrict the search criteria and to browse and refine the results. Biological facets such as gene information are user-defined features that researchers can use to provide a multi-faceted description of database records. The challenge to utilize a facet-based system when compared to deploying facets in commercial online systems such as amazon.com, is that both the data and the metadata deployed by the integration system are only available at the time the query is submitted to the system. This presentation describes the architecture of the Biofacets system and the solution adopted to deal with on-the fly categorization of the queried data using the biological facets.

## **BIOGRAPHICAL SUMMARY**

Malika Mahoui is an Assistant Professor at the School of Informatics and Adjunct Professor with the Department of Electrical and Computer Engineering and the Department of Computer Science. Her research background covers the areas of data management and data/text mining. Recently she has been interested in applying her knowledge to the areas of bioinformatics and health informatics. She has been involved in data integration projects that aim at facilitating for researchers accessing multiple databases and/or manipulating multiple analytical tools to perform their studies. She has also been involved in text mining projects where the focus is to provide automation support to extract useful information from text; including scientific literature and medical reports.

\*\*\*\*\* Refreshments will be served at Noon \*\*\*\*\*