



BIOINFORMATICS 2015 SPRING SEMINAR SERIES

Hosted by: Department of Computer and Information Sciences,
Department of Electrical and Computer Engineering &
Center for Bioinformatics and Computational Biology
<http://bioinformatics.udel.edu/seminars>

MONDAY, May 4, 2015

3:30pm

DBI Room 102

From Retrieval to Discovery in the Biosciences

Padmini Srinivasan, PhD

Professor, Computer Science Department

University of Iowa

<http://homepage.cs.uiowa.edu/~psriniva/newsite/index.html>

ABSTRACT: The study of automatic hypothesis discovery from text collections aka LBD, Literature Based Discovery, started as a cool and exciting innovation in the mid 80s. It has now yielded a variety of LBD algorithms and system implementations. Our contribution to this area was a prototype system built around a decade ago called Manjal, which supported open and closed discovery approaches of LBD and more. Manjal was used to successfully replicate a number of manually identified 'discoveries' and to suggest new hypotheses. We continued our LBD efforts in recent years by developing methods to harvest Twitter conversations for possible proto-hypotheses regarding drugs and diseases. We have learned key lessons from our LBD efforts. One is that it is important to provide a clear rationale for a suggested hypothesis. While an end user expects this, it is unfortunately not easy to accomplish. Another lesson is the difficulty of ensuring novelty and yet another is that the suggested hypothesis should not be absurd. Moreover, in our experience systems such as Manjal ignore the fact that the bioscience literature is used and of value for a wide range of purposes besides hypothesis generation and exploration. Set in this context, we present Ferret, our current prototype system for literature scanning and its underlying technology. Ferret is lightweight, flexible, species agnostic, and gene-centric. It resolves several of the problems encountered with Manjal and similar LBD systems by firmly placing the user in the driver seat. Several case studies have been conducted with Ferret showing its effectiveness.

BIO: Padmini Srinivasan is a Professor in Computer Science at the University of Iowa. She received her PhD from Syracuse University and has held sabbatical visiting appointments at Cornell University, the National Library of Medicine and the Indian Institute of Science. Her research interests are in text retrieval, text classification and text mining. She has explored problems in domains such as biomedicine, politics (using social media) and accounting (mining 10K filings). She was a member of the Biomedical Library and Informatics Review Committee (2009 to 2013) and also a member of the Board of Scientific Counselors, Lister Hill Research Center, NLM.