Presenter: Nancy Baker
Date: March 22, 2005
Paper: Marc Weeber, et al. (2003). Generating Hypotheses by Discovering Implicit Associations in the Literature: A Case Report of a Search for New Potential Therapeutic Uses for Thalidomide. *Journal of the American Medical Informatics Association* 10(3): 252–259.

Summary of paper:

This paper discusses the use of a software system developed by the authors which automates the Swanson A-B-C paradigm. The text-based discovery system called DAD was created to help researchers to formulate hypotheses and do initial testing on them. The system uses title and abstract as text, not just title as Swanson did. The user is very involved along the whole process; he/she decides interestingness. A key innovation is that the system maps the text to UMLS terms. This gives them the ability to include compound terms, select only biomedical terms, and make use of the semantic types and hierarchies inherent in MESH terms.

In this paper they give a short description of the system, but the purpose of the paper is to show how the system was used in one case: looking for new uses for thalidomide. Thalidomide is known to have immunomodulatory action which could potential treat a number of diseases.

Using the sytem they were able to come up with a number of diseases that thalidomide might treat. The user was key to the whole process.

Discussion:

Much of our discussion centered around the relationship between the developer of such a system and a user of it. Who would use it? Are physicians and researchers asking for such systems? Would they if they knew about them? Should IS people build them and they try to get people to use them? Does such a system really have a place in drug development or are traditional approaches what people should spend their time on.

We also discussed at length how to validate such a system. Most groups use finding Reynauds/fish oil connection or migraine/magnesium as validation. Is this enough? Is validation different for this kind of system?