**Lin, J., DiCuccio, M., Grigoryan, V., & Wilbur, J. (2008). Navigating information spaces: A case study of related article search in PubMed. *Information Processing and Management,* 44, 1771-1783.**

Keywords: Information retrieval systems, information space, and browsing

The authors conduct a case study on MEDLINE (the authoritative repository of abstracts from the medical and biomedical primary literature maintained by the US National Library of Medicine (NLM)). In this study, they examine one conception of an information space characterized by similarity links to focus on the effectiveness of related article search.

The authors set the stage for the study by stating that documents are entrenched in an interconnected network, and explain that information seeking: (a) navigates this network, and (b) provides a model to guide the design of information retrieval systems. However, they believe that retrieval systems are inadequate in assisting users navigate the

information space defined by their search results.

They explain that MEDLINE contains over 17 million records, including

bibliographic information, abstract text and links to full text. They elucidate that PubMed’s NLMs public gateway to MEDLINE provides a search feature designed to assist users browse the literature. They explain further that the current PubMed interface display links to five related articles. They posit that these articles are in turn connected to others

via the same type of links. Consequently, these connections define a vast document network in which the nodes represent MEDLINE citations, and the links represents content-similarity. They elucidate that each inherent call of the related article search provides the user with articles similar in content. Additionally, clicking on the related links

moves the users through this environment. The authors highlight several theories to support the ideas put forward such as: (a) the theory of effective group navigation, (b) the cluster hypothesis, and (c) the information foraging theory.

They focus on the foraging theory, which hypothesizes that when feasible, natural information systems evolve toward states that maximize gains of information per unit cost. They further explain that one assumption of the information foraging theory is the tendency for relevant information to cluster together. Hence, a user is faced with the choice of exploiting the cluster, or searching for the next cluster. The authors state that PubMed’s related article search provides cues to help users make decisions about their search behavior. They state that users might decide to follow related article links, or seek out entirely different locations.

The paper addresses issues such as: (a) the topological features of related document networks, (b) the relevance of topological features to related document network, and (c) the usefulness of navigating information spaces via content-similarity links. The authors use a simple approach to construct a related document network for each topic

in the TREC 2005 genomics. They highlight the results from the analysis of these networks and use effective visualization tools such as tables and histograms to assist in this regard.

In their conclusion, the authors state that they were able to: (a) identify document clusters, (b) gather requisite documents through browsing, and (c) perform searches that take advantage of related article links. Future research could look at a comparative study between two different approaches.