POSITION PAPER

Sameer Antani, PhD
santani@mail.nih.gov
http://archive.nlm.nih.gov/staff/antani.php

Biographical Sketch: Dr. Sameer Antani is a Staff Scientist with the Lister Hill National Center for Biomedical Communications an intramural R&D division of the National Library of Medicine (NLM) at the National Institutes of Health (NIH). His research interests are in image and text data management for large biomedical and multimedia archives. His recent research includes content-based indexing, and retrieval of biomedical images (CBIR), combining image and text retrieval for evidence-based medicine, topics in advanced multimodal medical document retrieval, and next-generation interactive (multimedia rich) documents. He earned his Ph.D. degree in Computer Science and Engineering from the Pennsylvania State University, in 2001. Dr. Antani is a member of the IEEE, the IEEE Computer Society, and SPIE. He serves on the steering committee for IEEE Symposium for Computer Based Medical Systems (CBMS).

Interest in Visualization of Electronic Health Records: As increasing amount of health data becomes electronic it is important for clinical and biomedical researchers, health practitioners, and consumers to have access to tools that can help improve access and sharing of medical information. Recent developments in creation of medical record warehouses through Google Health and Microsoft Health Vault are also indicative of an increased interest in allowing transparency in sharing and archiving health information and in general enabling the lay public in better understanding and managing their health (and health records.) In contrast to these consumer efforts, hospital systems have maintained health records as transactions of a particular patient as his/her case moves through various centers during his treatment. During this the image data are stored in a PACS or RIS and the health data are stored in a HIS. Current plans for EHRs or PHRs do not model interactions with images. It would be a valuable for the patient to be able to carry (or access, or share) their hospital health record including images making it important to develop tools that support this activity. These could include database technologies that support large heterogeneous text data as well as interactive visualization, search, and retrieval tools that take advantage of the abundant text data, image data, other multimodal information, and ancillary information. Such tools would be of great interest to medical practitioners, biomedical researchers, and consumers alike.

Area of Work: I am working on correlating image and text features for improved information retrieval and question answering for evidence-based medicine. Work in this area has concentrated largely on full-text biomedical articles; but we do not exclude EHRs from playing a role in this in the future. The project has offshotts in in providing information at the point of care for nurses while we expand into the use of PACS/DICOM storage for images. Also, NLM is interested in PHRs and is interested in developing repositories that can potentially combine many different health record data (varying sizes, sources, and topics) for studying data mining, de-identification, PHR, and other related issues. I am interested in solving problems in use of image data that may accompany such clinical information and working on targeting these different, but related, modalities.