The provision of health services involves a wide range of stakeholders; including patients, direct health care providers, researchers, managed care organizations, and third-party payers. These stakeholders often have considerable differences in objectives, concerns, priorities and constraints, making data and knowledge management in health care organizations a challenging endeavor. The Data and Knowledge Management in Health Care Minitrack focuses on the use of knowledge management technologies for clinical, managerial, and population based activities. In recent years, cost pressures have again been affecting the health sector. The affordability and availability of health care has drawn the attention of candidates, political pundits, and working Americans. The adoption of information technology is viewed as one of the primary mechanisms for efficiency gains. Along with cost concerns, the quality of care is another area in which the application of information technology holds great promise. Initiatives from hospital and physician report cards to continuous quality improvement programs rely on increasingly sophisticated database infrastructures.

This year, the eight accepted papers (from thirteen submissions) deal with a cross section of data and knowledge management topics. The first paper, “Projecting Computational Sense of Self: A Study of Transition in a Chronic Illness Online Community” by McArthur, Bruza, Warren, and Kralik explores the automated categorization of e-mail utterances in an online discussion group. The analysis of communication patterns is used to identify “transitions” as people learn to deal with chronic illnesses and the day-to-day impact on their lives.

The second paper “Linking Tacit Knowledge in the Pediatric Pain e-Mail Archives and Explicit Knowledge in PubMed” by Chen, Shepherd, and Abidi investigates information retrieval strategies using the tacit knowledge from ongoing e-mail discussions to identify reference works from the PubMed archives. This paper builds on previous work that focused on clustering e-mail message threads. However, the current work makes the connection to the PubMed literature collection through the MeSH controlled medical vocabulary.

The third paper, “Linguistic and Cultural Differences in Information Categorization and their Impact on Website Use” focuses on issues of website design and localization for content delivery to different international audiences.

The fourth paper (and second session) deals with pattern discovery in large-scale databases. The paper “Knowledge Extraction from Prostate Cancer Data” by Delen and Patil investigates the applicability of commercially available data mining techniques for predicting prostate cancer survivability.

The fifth paper by Michalowski, Wilk, Thijssen, and Li uses Bayesian belief networks to model clinical pathways and predict length of stay for radical prostatectomy patients.

The sixth paper, “Uncovering the Patterns in Pathology Ordering by Australian General Practitioners: A Data Mining Perspective” by Zhuang, Churilov, and Sikaris continues the focus on knowledge discovery in databases. In this case, clusters of different pathology or medical test patterns are discovered and used to categorize patient groups.

The seventh paper “Feature Selection for Predicting Surgical Outcomes” by Tremblay, Berndt, and Studnicki looks at the attribute selection process prior to modeling. An attribute framework is proposed and used for predicting length of stay in the realm of digestive surgery.