

## **A Collaborative Platform for Sharing EHRs of Cancer Patients**

Praveen R. Rao<sup>1</sup>, Stanley A. Edlavitch<sup>2</sup>, Jeffrey L. Hackman<sup>3</sup>, Timothy P. Hickman<sup>2</sup>,  
Douglas S. McNair<sup>4</sup>, and Deepthi S. Rao<sup>5</sup>

<sup>1</sup> School of Computing and Engg., University of Missouri-Kansas City, Kansas City <sup>2</sup> School of Medicine, University of Missouri-Kansas City, Kansas City <sup>3</sup> Truman Medical Centers, Kansas City <sup>4</sup> Cerner Corporation, Kansas City <sup>5</sup> Argentine Family Health, Kansas City

Today the nation faces one of the toughest challenges in health care due to high operating costs. By 2017, it is projected that the nation would spend \$4.3 trillion for health care. One way to lower health care costs and provide faster and effective care to patients is by using Information Technology (IT). In a recent report by Stead and Lin (2009), “data sharing and collaboration” and “large scale management of health care data” have been identified as key IT challenges to advance the nation’s health care system.

Vast amounts of information (e.g., electronic health records, drug data, and data from clinical diagnosis) remain largely untapped due to the lack of suitable IT solutions. Moreover, the data sources are heterogeneous and evolve over time making their mediation and interoperability quite difficult. Although a federated database model is commonly adopted in today’s data integration systems (e.g., National Cancer Institute’s caGrid), the design requires the creation of a mediated schema and semantic mappings (for processing queries), which can be cumbersome as the local schemas become more complex and the number of data sources increases. Furthermore, the mediation process requires sufficient domain knowledge for each source. Another issue is that the presence of centralized components in current systems will lead to scalability bottlenecks and single points of failure.

We are investigating a new framework for sharing EHRs called a Collaborative Data Network (CDN). As a proof-of-concept, we focus on cancer cases, which is the second most common cause of death in the US. A CDN leverages two successful technologies, namely, (a) peer-to-peer (P2P) computing and (b) a combination of XML/JSON standards. The marriage of these two technologies provides numerous benefits for managing EHRs such as scalability, fault-tolerance, a richer data model, query expressiveness, and non-standardization. A CDN would allow a single query to be issued across multiple data sources. Each data provider has complete control of its data and can employ local access control policies for protecting the privacy of patients. Atop a CDN, we can build decision support and data mining tools to operate over multiple, non-standardized data sources. Our framework can enable fast, high quality, cost-effective patient care by assessing best practices, monitoring side-effects,

leveraging knowledge from clinical trials, superior decision support tools, and so forth. Our framework can be migrated to operate in a cloud computing environment, thereby reducing the cost of infrastructure setup and maintenance.

Our proposed framework can advance the efforts of NIH, and a joint-effort of AHRQ and FDA – to create an integrated framework for data sharing and coordination of large-scale, evolving health care and medical data. We believe our efforts can result in an incremental advancement towards a “learning health care system.” Our team of investigators has expertise in clinical research, decision support for clinical environments, IT for clinical data management, and peer-to-peer data management. We highlight potential opportunities for collaboration during the poster session.