Abstract: Delivering contextually-relevant knowledge resources into EHR systems at the point of care was proposed as Infobutton HL7 standard. However, the InfoButton standard does not consider social media data or the patients’ healthcare behaviors or practices. We present a Social InfoButtons system that collects patient-generated social media data and other open health data to provide insights on healthcare trends and patients’ practices and issues, using the Semantic Integration model, that supports Social Healthcare Knowledge for clinicians, patients and policy makers.

Introduction: A study [1] showed that 34% of the Internet users utilized social media to read other patients’ commentaries and experiences about health or medical issues. The InfoButtons system of Cimino et al. [3, 4] provides knowledge resources needed by clinicians but does not incorporate patient-generated social health data. With the Health 3.0 trend, it is increasingly becoming important to understand the patients’ actual health practices, behaviors, trends and concerns. Our Social InfoButtons system generates contextually summarized information about social health practices by geographic or temporal dimensions, providing end-users (e.g., patients, clinicians, or government officials) with healthcare information, such as treatments, practices, conditions, experiences, sentiments, and behaviors reported by other patients through social media.

Methods: Providing integrated social health knowledge is challenging, since the online health data sources vary in formats and platforms. We present a semantic model for representing the integrated social health knowledge, and use Linked Open Data [2] to integrating heterogeneous social media health data sources. We developed the Social InfoButtons architecture to provide a social health analytics platform that can summarize, visualize and compare the contextually relevant health information for patients, clinicians, and healthcare government officials. The architecture consists of the data collector, RDF triple store, and analytics module. (1) The data collector extracts publicly available health data from various data sources, such as the social network site PatientsLikeMe, the government-maintained CDC website, the PubMed website, and the patient resource portal WebMD; (2) The extracted entities from different sources were integrated with RDF triples; (3) The coverage of the analytics module of Social InfoButtons ranges from simple SPARQL queries to interactive geographic and temporal visualizations to gather insights about social health trends and anomalies.

Results: The integrated knowledge base currently has 612,017 triples representing 1228 conditions from different sources and their related information such as treatments, symptoms, and side effects of treatments. The figures below show the geographic knowledge retrieved by Social InfoButtons for Diabetes Type 2 (left), and a timeline trend of negative sentiments towards tuberculosis from Twitter users (right).

Conclusions: A Social InfoButtons prototype was developed to provide social health knowledge that is relevant to an end users’ context. A semantic integration model and social analytics module are presented for social health knowledge, extracted from data sources ranging from social network sites, research communities, and government sites to patient Web resources.

References