INVITED SPEAKERS
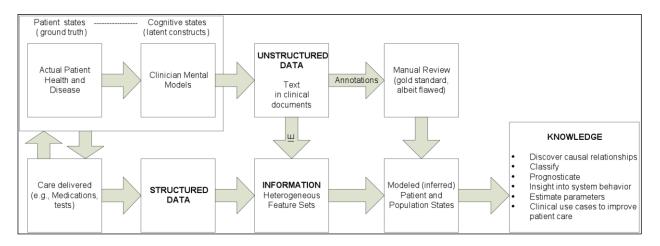
Big Data and Arthritis

Big Data Analytics – Veterans Affairs Informatics and Computing Infrastructure (VINCI)
By Yelena Yesha, Amanda Niskar, Michael Grasso, Scott DuVall, Naphtali Rishe
U.S. Arthritis Foundation, University of Maryland, University of Utah, NOA Data Science Institute Inc, Florida International University
yelenayesha@gmail.com

We work with clinical data from the Department of Veterans Affairs (VA) to perform Big Data Analytics for Arthritis and other conditions. The VA dataset is one of the largest clinical repositories available, and provides detailed patient information for approximately 25 million patients who have received care at 152 medical centers and more than 800 outpatient clinics across the United States over the past 15 years. The repository includes more than 4 billion progress notes, 2 billion procedure and imaging reports, 1.6 billion medication fills, and 1.5 billion diagnoses.   The repository also contains patient demographics, diagnostic codes, outpatient visits, hospital admissions, patient orders, vital signs, laboratory testing, inpatient and outpatient pharmacy data, clinical consults, immunizations, mental health screening, associated physicians, and payment information. This is a large and comprehensive data set, whose size and attributes create analytic and data mining issues that are beyond the capabilities of traditional software tools. The data comes from 163 Hospitals, 800 Clinics, 135 Nursing Homes, 43 Domiciliaries, 180,000 Healthcare Professionals, Hospital sizes vary from 100 to 1000 beds. Outpatient visits vary at the facility from 30,000 to 450,000 visits per year. There are 46.5 million outpatient visits per year, 564,000 inpatient admissions per year, 167 million prescription-months filled. The following diagram depict the knowledge discovery workflow.

This research has been sponsored in part by the Arthritis Foundation.