Dr. Trevor Cohen

Email:cohenta@uw.edu

Quantifying Linguistic Anomalies in Neuropsychiatric Diseases: This project concerns the development and evaluation of computational methods to identify linguistic manifestations of cognitive changes in neuropsychiatric diseases (dementia and psychotic disorders). Cognitive changes in these disorders are reflected in language, which is used as a basis for diagnostic assessment. For example, patients with Alzheimer's Disease Dementia (ADD) may revert to higher level categories of concepts as their access to more specific terms diminishes over the course of disease. In contrast a cardinal sign of psychotic disorder is disorganized thinking, which may manifest as diminished semantic relatedness between units of text in a transcript of a speech sample. The goal of this project is to develop and evaluate methods that can detect and characterize such linguistic anomalies, leveraging recent advances in neural representations of language.

Dr. John Gennari

Email: gennari@uw.edu

John Gennari would be excited to mentor and work with a postdoctoral fellow in a wide variety of research directions. He has active research projects in the realm of **re-usability and reproducibility**, including funding from the NIH-sponsored Center for Reproducible Biomedical Modeling. His area of expertise includes **knowledge reuse, knowledge representation, ontologies, and semantic web technologies.**

Dr. Gang Luo

Email: luogang@uw.edu

Using Computational Approaches to Optimize Asthma Care Management: The study will develop more accurate, computational predictive models and a novel automatic explanation function to better identify patients likely to benefit most from care management. For many chronic diseases, a small portion of patients with high vulnerabilities, severe disease, or great barriers to care consume most healthcare resources and costs. To improve outcomes and resource use, many healthcare systems use predictive models to prospectively identify high-risk patients and enroll them in care management to implement tailored care plans. For maximal benefit from costly care management with limited service capacity, only patients at the highest risk should be enrolled. But, current patient identification approaches have two limitations: 1) Low prediction accuracy causes misclassification, wasted costs, and suboptimal care. If an existing model were used for care management allocation, enrollment would miss >50% of those who would benefit most but include others unlikely to benefit. A healthcare system often has insufficient data for model training and incomplete data on many patients. A typical model uses only a few risk factors for adverse outcomes, despite many being known. Also, many predictive variables on patient and system characteristics are not found yet. 2) No explanation of the reasons for a prediction causes poor adoption of the prediction and busy care managers to spend extra time and miss suitable interventions. Care managers need to understand why a patient is predicted to be at high risk before allocating to care management and forming a tailored care plan. Existing models rarely give such explanation, forcing care managers to do detailed patient chart reviews.

To address the limitations and optimize care management for more high-risk patients to receive appropriate care, the study will: a) improve accuracy of computationally identifying high-risk patients and assess potential impact on outcomes; b) automate explanation of computational prediction results and assess the impact on model accuracy and outcomes; c) assess automatic explanations' impact on care managers' acceptance of the predictions and perceived care plan quality. The use case will be asthma that affects 9% of Americans and incurs 439,000 hospitalizations, 1.8 million emergency room visits, and $56 billion in cost annually. Asthma experts and computer scientists will use data from three leading healthcare systems; a novel, model-based transfer learning technique needing no other system's raw data; a novel, pattern-based automatic explanation technique that also improves model generalizability and accuracy; a new data source, PreManage, to make patient data more complete; and novel features on patient and system characteristics.