

The Event and Pattern Detection Laboratory (EPD Lab), directed by Professor Daniel Neill, is a research laboratory at Carnegie Mellon University's Heinz College, with funding support from the National Science Foundation, MacArthur Foundation, Richard King Mellon Foundation, UPMC Technology Development Center, and Disruptive Health Technologies Institute.

Our research focuses on the development of new statistical and computational techniques for scalable and accurate detection of emerging events and other patterns in complex, massive, and high-dimensional datasets. We are particularly interested in applications of this work for the public good, ranging from public health and patient care, to law enforcement and urban analytics, to human rights and conflict. We work directly with a variety of organizations in the public and private sectors, including public health practitioners, hospitals, police departments, and city leaders, to develop data-driven decision support systems that can improve public health, safety, and security.

The EPD Lab is currently engaging in multiple projects in the health care domain, with the primary goals of improving quality of care and population health outcomes while reducing healthcare costs. Some of these projects include:

- Improving the accuracy of disease diagnosis, prognosis and risk prediction. For example, we have developed new methods for **chronic disease risk prediction and visualization** that give clinicians a comprehensive view of their patient population, risk levels, and risk factors, along with the estimated effects of potential interventions for disease prevention. We have applied these approaches to diabetes, chronic kidney disease, and related complications.
- **Optimizing hospital processes** such as resource allocation and patient flow. For example, by early and accurate prediction of each patient's diagnosis-related group, we can better predict demand and allocate scarce hospital resources such as beds and operating rooms.
- Improving **medication reconciliation** to reduce adverse drug events, by automatic detection of errors and omissions in patients' medication lists.
- We have developed new methods to efficiently identify regions of interest in digital pathology slides, and demonstrated that these approaches are effective for **detecting prostate cancer**.

We are particularly excited about two of our newest projects:

- **Discovery of new medical knowledge**, such as clinical guidelines and best practices. We are currently working with UPMC and Highmark to analyze massive quantities of patient care data (electronic health records and health insurance claims data) to discover anomalous patterns of care that significantly impact patient outcomes. These impacts could be negative,

in which case we have discovered a pattern of suboptimal care to be corrected, or positive, in which case we have discovered a new potential best practice. In addition to medicine, these approaches could be applied to scientific discovery, enabling the data-driven generation of new scientific hypotheses which could be then be further evaluated.

- Use of hospital microbiome data to model and prevent the spread of **hospital-acquired illnesses**. We are working with a novel dataset from a Chicago hospital consisting of complete taxonomies of the micro-organisms present at various points in space and time (collected through swabbing and sequencing), such as patient beds, caregiver gloves, etc. Using these data, we hope to identify the primary means by which different serious hospital-acquired illnesses spread, such as direct person-to-person, insufficient sterilization of hospital beds, insufficient handwashing by caregivers, etc. These identified patterns and models of disease spread will suggest specific key preventive measures which may differ from strain to strain.

In addition to health care, other EPD Lab application focus areas include:

1) **Disease surveillance**. We work with state and local public health departments, using electronically available health data such as hospital visits, medication sales, and online user behavior to automatically identify and characterize emerging disease outbreaks. We are particularly interested in using emergency department chief complaint data for early discovery of “novel” outbreaks with unusual symptoms, or unusual geographic and demographic patterns, which may have high pandemic potential.

2) **Crime prediction and prevention**. We work with city law enforcement agencies to predict and prevent crime patterns, using data sources such as crime offense reports, 911 emergency calls, and online social networks. Our CrimeScan software has been used effectively by the Chicago Police Department for targeted patrolling to prevent hot-spots of violence. We are currently working with Pittsburgh police and city and county leaders to identify opportunities for crime prevention through a combination of precisely targeted policing and “non-punitive” interventions, such as targeted city clean-up efforts and better provision of human services.

3) **Urban analytics**. We work with city leaders in Chicago, Pittsburgh, and Baltimore to predict and proactively respond to patterns of citizen needs, as measured by 311 non-emergency calls for service. For example, we have deployed and are in the process of evaluating approaches for rodent prevention (through advance prediction and proactive baiting) in Chicago.

We have also applied our work to numerous other areas, including prediction of civil unrest, early detection of emerging patterns of human rights events, network intrusion detection, customs monitoring of container shipments, and physical infrastructure monitoring.

For more information on our current projects, please visit the Event and Pattern Laboratory web page, <http://epdlab.heinz.cmu.edu>, or contact Prof. Daniel Neill (neill@cs.cmu.edu).