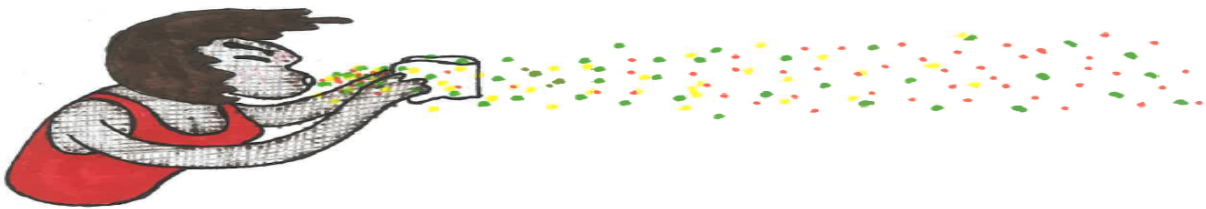# How do we mine social media data for insights into contagions?

National security is the first priority at Los Alamos National Laboratory. This leads some of our efforts into the field of epidemiology. Forecasting, preventing, and managing outbreaks of contagious diseases is an area of vast importance to our country and the people of the world.

Los Alamos researchers use supercomputers to make highly detailed models of geographical disease behavior. They use these models to study how diseases may spread, and to explore strategies to manage their spread and weaken their effects. Suppose avian flu arrived in an American city. Would closing that city's airport help contain the disease? LANL models suggest that closing the airport would not be of much help.

Traditionally, medical practitioners have shared information with governmental health agencies so the agencies might identify patterns in the data. This system has several problems. Not all providers report, or they may not recognize local patterns if they see only a few patients. There are always delays in data reporting. Often, patients may be reluctant to go to a doctor.

Some of those patients, however, may be less timid about going on line to ask Google or Wikipedia about their symptoms. They might tweet about how they feel or share symptoms on Facebook. Information from these channels is mined for specific words and phrases that may help researchers in early identification of public health issues, or help pin down the nature or magnitude of an event. We use this information to support the work of public health officials.