If ever evidence was required of the increasing utility of social media, the above report in the *The New York Times* establishes it for sure. In today’s scenario, social media such as Google, Facebook, Twitter, chat rooms, blogs and many more are no longer mere idle distractions. Instead, they are emerging and establishing themselves as powerful social networks, powerful enough to bring about revolutions by marshaling widely scattered resources. And as can be seen from the...
If you have been to a social web gathering and a day later experience symptoms like painless diarrhoea and vomiting of clear fluid with racing pulse, you may share your status with your circle of friends and in a moment of Internet enabled self-diagnosis.

NY Times excerpt, social media are even emerging as powerful data sources providing near real-time data to predict disease outbreaks.

Tracking and prediction of communicable diseases, commonly known as infectious disease surveillance, plays a pivotal role in management of public health and policy making for health issues. This practice is crucial for the establishment of patterns of progression of infections and adoption of necessary precautions and treatments to minimize the harm caused by epidemics. Early prediction or detection of epidemic outbreaks is very important when it comes to planning response strategies to be adopted and making health related policies.

There are various organisations across the globe that assume this responsibility.

Dr Nicholas Christakis and James Fowler (Source: Google Images)
Social networks if properly studied and exploited can yield valuable information on social behaviour and disease progression.

Specialized agencies like the Centers for Disease Control and Prevention (CDC), USA and World Health Organization (WHO) are two such organizations involved in the monitoring and tracking of epidemics worldwide.

Traditionally, these organizations make use of statistical data collected through health institutions and official reporting structures on reported symptoms, outpatient reporting and test results supplied by laboratories worldwide to generate mathematical models for predicting the possibility of disease outbreak and to monitor the progress of contagious diseases. This practice largely relied on survey techniques and manual reporting for data collection.

Though they have shown some utility, these methods are often theoretical and time consuming. In some cases, the data for a particular infection may not be available for weeks, hindering early epidemiologic assessment and thus delaying the planning and execution of disease response strategies and health policies. In addition, cases of new and emerging diseases where little or no historical data exists makes the collection of data by traditional methods complex and often unproductive. With advanced modes of transport causal organisms of epidemics like SARS and H1N1 can travel speedily from continent to continent in a matter of hours. Hence, it becomes mandatory to adopt faster and more flexible methods for disease surveillance.

In light of all these, currently researchers are involved in exploring the possibility of quicker methods of epidemic prediction in the form of virtual media and social networking sites. Scientists propose making use of search engines like Google and social networking media like Twitter, Facebook, Blogs and Chat rooms to predict epidemics.

Internet and social networking media have permeated all spheres of our professional and personal life. The internet today is one of those resources that is being used pervasively by a variety of people including clinicians, public health practitioners, researchers and even the common public, to seek health information. It is a common trend today to attempt a self-diagnosis of diseases making use of web-based searches or initiate discussion groups at social networking sites. Researchers propose that these currently trending practices can be used as data sources to estimate epidemic dynamics.

**Using Social Media – How it Works**

Scientists say, web searches and microblogs show a good analogy for public health events. According to Dr. David Fisman, Epidemiologist, University of Toronto, Canada, it is based on the simple concept that, “the way that information moves is very similar to the way disease moves.” Now, let us look at how it can be achieved.

Say you have been to a social gathering like the one mentioned above and a day later experience symptoms like painless diarrhoea and vomiting of clear fluid with racing pulse. You, like Nico Zelfang, may share your status with your circle of friends or may indulge in a gathering like the one mentioned above and a day later experience symptoms like painless diarrhoea and vomiting of clear fluid with racing pulse. You, like Nico Zelfang, may share your status with your circle of friends or may indulge in a moment of Internet enabled self-diagnosis. In the blink of an eye dozens of health-related websites will appear on your screen. The search will supply you with information – some useful and some not – but the search term you provide acts as a data point for those who survey disease outbreaks by monitoring how people report symptoms via virtual/social media. This is essentially like tapping into people’s communications about health events.

The data present in twitter streams, search queries, and discussions on Facebook, chat rooms and blogs can thus be effectively mined. This mined data is then processed using modern tools to extract key elements according to the geographical locations. The information generated is then used to predict epidemics.

**Interesting Developments**

Numerous studies have been carried out to understand and evaluate the role of online social and news media in public life. Research has been undertaken to determine how effective they are when it comes to early detection of diseases and to check whether data put forth by these methods correlate with officially reported disease measures.

Result of a pioneering effort in this field is available on the website HealthMap ([http://healthmap.org/](http://healthmap.org/)) developed by Dr John Brownstein and collaborators (Children’s Hospital Boston, USA) in 2006. HealthMap works on the principle of mining news websites, microblogs, government alerts, eyewitness accounts, and other data sources for outbreaks of infections reported globally and aggregates those cases on a global map. HealthMap displays outbreaks in real time on the global map.

Brownstein’s team recently launched an iPhone application called “Outbreaks Near Me”. This application delivers HealthMap directly to cell-phone users. Their newest endeavor is the website called Flu Near You ([https://flu near you.org/](https://flu near you.org/)), created in collaboration with the American Public Health Association and the Skoll Global Threats Fund of San Francisco, USA. This unique website allows users to serve as potential disease sentinels by reporting their health status on a weekly basis.

Various groups across the globe have conducted experiments in this subject of crowd-sourced epidemiology. In September 2008, encouraged by the observation that increase in flu queries on the web and disease outbreaks often coincide, Google launched Google Flu Trends ([http://www.google.org/flutrends/](http://www.google.org/flutrends/)) and later Google Dengue trends ([http://www.google.org/denguetrends/](http://www.google.org/denguetrends/)). Both these resources work on the principle that certain search terms are good indicators of flu and dengue activity respectively particularly with respect to geographical location of the users. These tools use automated algorithms to discover influenza/dengue related search.

The aggregated Google search data is then used to arrive at a disease pattern to estimate current flu/dengue activity around the world in near real-time. This website allows people to compare volumes of flu/dengue related search
activity against reported incidence rates for the illness displayed graphically on a map and hence acts as potential source for early warnings in infection-prone locations.

Recently, Dr John Brownian and his team conducted a study to assess the correlation of volume of cholera-related HealthMap news media reports, Twitter postings, and government cholera cases reported during the first 100 days of the 2010 Haitian cholera outbreak. It was observed that trends in volume of data from informal sources like Twitter postings significantly correlated in time with official case data and was available up to two weeks earlier. Hence, the researchers came to the conclusion that these unconventional sources can be used complementarily with official data in an outbreak setting to get timely estimates of disease dynamics.

In today’s scenario, social media such as Google, Facebook, Twitter, chat rooms, blogs and many more are no longer mere idle distractions. Instead they are emerging and establishing themselves as powerful social networks, enough to bring about revolutions by marshalling widely scattered resources.

One of the most recent works in this area involves the work done by scientists at Harvard University in Cambridge, USA and The University of California, USA, who aim to determine a method of predicting the spread of contagious diseases using social networking sites such as Facebook. Dr Nicholas Christakis at Harvard Medical and James Fowler at UC San Diego have come up with analysis on how social networks like Facebook can be used to predict the progress of disease.

Their work is based on the concept of human social behaviour that the position of an individual in the social network is determined by their popularity. Hence individuals towards the centre of a social network are more popular with more connections in the virtual and real world. These individuals owing to more connections are more prone to acquiring an infection rather than individuals towards the periphery. Thus monitoring the individuals towards the centre of a social network can aid in early detection of epidemics. The central individuals can be easily identified utilizing the friendship paradox – the idea that a friend of a randomly chosen individual is likely to have more friends than the random person himself, following whom one can move towards the centre of a network without undertaking the tedious task of fully mapping social networks.

An Indian Effort
An indigenous development in using social network to predict contagious diseases is the online resource developed by scientists at the Institute of Genomics & Integrative Biology (CSIR-IGIB). The resource called DATE (acronym for Data Analytics for Tracking Epidemics) is an online resource that tracks social media posts by more than a million population spread across 43 Indian cities to create epidemics alerts in India.

The resource has been developed in an open source mode with community participation. The datasets and software codes are freely shared and available for further development. The resource uses custom algorithms developed by the community to track, analyse and predict the occurrence of epidemics in different parts of India using posts on Twitter, news websites, blogs and other public social media resources.

Social networks if properly studied and exploited can yield valuable information on social behaviour and disease progression. In the wake of ever increasing use of virtual and social media, the existing disease surveillance practises are poised for a massive transformation thus providing new perspectives with respect to responses to disasters and pandemics.

Ms Geetha Sugumaran is Project Assistant, Open Source Drug Discovery Project, Council of Scientific & Industrial Research, Room No 122, DGTC, 2 Rafi Marg, CSIR HQ, Anusandhan Bhawan, New Delhi-110001