

Privacy concerns could limit benefits from real-time data analysis, researcher says

December 17 2009

Society will be unable to take full advantage of real-time data analysis technologies that might improve health, reduce traffic congestion and give scientists new insights into human behavior until it resolves questions about how much of a person's life can be observed and by whom, a Carnegie Mellon University computer scientist contends in a commentary published Friday in the journal *Science*.

In a "Perspectives" column, Tom M. Mitchell, head of the Machine Learning Department in Carnegie Mellon's School of Computer Science, notes that data-mining techniques, once used for scientific analysis or for detecting potential credit card fraud, increasingly are being applied to personal activities, conversations and movements, such as information that can be deduced about an individual by monitoring that person's smart phone.

"The potential benefits of mining such data range from reducing [traffic congestion](#) and pollution, to limiting the spread of disease, to better using public resources such as parks, buses, and ambulance services," Mitchell wrote. "But risks to privacy from aggregating these data are on a scale that humans have never before faced."

Technical means can help limit threats to privacy and misuse of data, Mitchell said. One approach is to mine data from many different organizations without ever aggregating the data into a central repository. For instance, individual hospitals might analyze their medical records to see which treatments work best for a particular [flu strain](#), then use

cryptography to encode the results and protect patient privacy; only then would the findings be combined with those from thousands of other hospitals.

"Perhaps even more important than technical approaches will be a public discussion about how to rewrite the rules of data collection, ownership, and privacy to deal with this sea change in how much of our lives can be observed, and by whom," Mitchell wrote. "Until these issues are resolved, they are likely to be the limiting factor in realizing the potential of these new data to advance our scientific understanding of society and human behavior, and to improve our daily lives."

Mitchell pointed out that the use of real-time data from individuals already has begun. In many cities, anonymous location data from [smart phones](#) is being used to provide up-to-the-minute reports of traffic congestion. Researchers have shown that by analyzing health-related Google queries from particular geographic areas, they can estimate the level of flu-like illnesses in regions of the U.S. before government agencies such as the Centers for Disease Control and Prevention can provide estimates. Scientists are beginning to use real-time sensing of routine behavior to study interpersonal interactions as people go about their daily lives.

Combining data sets could open up many new possibilities, as well as new [privacy](#) issues, Mitchell said. "For example, if your phone company and local medical center integrated GPS phone data with up-to-the-minute medical records, they could provide a new kind of medical service using phone GPS data to detect that you have recently been near a person who is just now being diagnosed with a contagious disease — then automatically phoning to warn you."

A former president of the Association for the Advancement of [Artificial Intelligence](#) (AAAI), Mitchell is a member of an AAAI panel that is

exploring the potential societal impacts of advances in artificial intelligence. A pioneer in artificial intelligence and machine learning, Mitchell was named a University Professor, the highest distinction that faculty can achieve at Carnegie Mellon, in May 2009. He has been head of the Machine Learning Department since the first-of-its-kind department was established in 2006. His research focuses on statistical learning algorithms for understanding natural language text and on understanding how the human brain represents information.

Provided by Carnegie Mellon University

Citation: Privacy concerns could limit benefits from real-time data analysis, researcher says (2009, December 17) retrieved 20 March 2024 from <https://phys.org/news/2009-12-privacy-limit-benefits-real-time-analysis.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--