Effective Counterterrorism and the Limited Role of Predictive Data Mining

by Jeff Jonas and Jim Harper

Executive Summary

The terrorist attacks on September 11, 2001, spurred extraordinary efforts intended to protect America from the newly highlighted scourge of international terrorism. Among the efforts was the a way to discover planning and preparation for terrorism. Data mining is the process of searching data for previously unknown patterns and using those patterns to predict future outcomes. Information about key members of the 9/11 plot was available to the U.S. government prior to the attacks, and the 9/11 terrorists were closely connected to one another in a multitude of ways. The National Commission on Terrorist Attacks upon the United States concluded that, by pursuing the leads available to it at the time, the government might have derailed the plan.

Though data mining has many valuable uses, it is not well suited to the terrorist discovery problem. It would be unfortunate if data mining for terrorism discovery had currency within consideration and possible use of "data mining" as national security, law enforcement, and technology circles because pursuing this use of data mining would waste taxpayer dollars, needlessly infringe on privacy and civil liberties, and misdirect the valuable time and energy of the men and women in the national security community.

What the 9/11 story most clearly calls for is a sharper focus on the part of our national security agencies -- their focus had undoubtedly sharpened by the end of the day on September 11, 2001--along with the ability to efficiently locate, access, and aggregate information about specific suspects.

Jeff Jonas is distinguished engineer and chief scientist with IBM's Entity Analytic Solutions Group. Jim Harper is director of information policy studies at the Cato Institute and author of Identity Crisis: How Identification Is Overused and Misunderstood.