United States—National Security

Constitution Project Recommends Reformed Data Mining Procedures for Federal Agencies

Privacy and intelligence experts Dec. 7 sparred over the effectiveness of data mining to predict possible terrorist activity during a panel discussion held by the Constitution Project in conjunction with its release of a report making recommendations for reforming data mining procedures to ensure the preservation of civil liberties.

The report, Principles for Government Data Mining: Preserving Civil Liberties in the Information Age, calls on agencies to build in constitutional safeguards when they implement data mining programs or launch new programs, Sharon Bradford Franklin, senior counsel for the Constitution Project, told BNA Dec. 9. The report recommends that Congress, the president, and federal agencies incorporate critical protections for individual rights into all government data mining programs. The project offers two sets of principles as a starting point for agency-specific data mining regulations or for governmentwide rules.

The report is part of the Constitution Project’s overall efforts to ensure that as data mining technology increases, that the collection, acquisition, and use of data does not infringe on individual privacy rights and respects the constitutional rights of freedom of expression, due process, and equal protection, Franklin said.

In addition, the Constitution Project calls for Congress and the president to revive the Privacy and Civil Liberties Oversight Board.

The use of data mining—defined in the report as the use of computing technology to examine large amounts of data to reveal relationships, classifications, and patterns to identify potential wrongdoing—is a valuable tool to the government to uncover fraud or other criminal activity, the report says.

Efficacy of Predictive Analysis Questioned. Franklin, who served as panel moderator, said that data mining has long been used as a tool to detect fraud and other crimes such as tax evasion. Most recently it is being used as a counterterrorism weapon. However, the use of data mining in national security cases can be challenging. The problem is that Americans know little about data mining for anti-terrorist purposes.

Panel member Jim Harper, director of information policy studies, the Cato Institute, while endorsing the report, raised questions about the use of “predictive” pattern-based analysis in data mining as opposed to “link-based” analysis.

Predictive analysis is the process of searching for data for previously unknown patterns and using those patterns to predict future outcomes, he explained. The problem is that terrorism is unpredictable and that pattern-based data mining programs work only when there is a high level of confidence that the tracked behavior is predictive.

On the other hand, link analysis takes known information and uncovers relationships and associations between objects, events, or persons that are not apparent from isolated pieces of data.

Franklin told BNA that in the report, the Constitution Project raises the question of whether data mining is really an effective counter-terrorism tool. Part of the problem is that there are not many data points for terrorist activity. Also, terrorists tend to avoid patterns, she explained.

Paul R. Pillar, a former intelligence officer for the CIA and the Counterterrorist Center, emphasized that no technology can achieve the extent of performance that Americans want from government counter-terrorism efforts. Therefore, the government has to use every set of tools available to it.

Pillar said as far as counter-terrorism goes, experts are not talking about “predicting” terrorist activity. Rather, the goal is to increase the odds in favor of counter-terrorism efforts. The goal is to try to identify people who are not terrorists so that law enforcement can focus on the possible terrorists. He also opined that terrorism is not totally patternless. Pillar is visiting professor and director of studies of the Securities Studies Program in the Edmund A. Walsh School of Foreign Service, Georgetown University.

In the report, the Constitution Project says that the efficacy of predictive analysis for terrorism prevention has yet to be demonstrated to the public.
Mary Ellen Callahan, chief privacy officer, Department of Homeland Security, said that data mining is not used in the agency’s Secure Flight or airport scanning programs. She discussed one of three data-mining program at DHS, the Automated Targeting System, or ATS, which she explained, looks for patterns, but also incorporates human intervention.

“ATS compares traveler, cargo, and conveyance information against intelligence and other enforcement data by incorporating risk-based targeting scenarios and assessments. DHS uses ATS to improve the collection, use, analysis, and dissemination of information that is gathered for the primary purpose of targeting, identifying, and preventing potential terrorists and terrorist weapons from entering the United States,” according to the Privacy Impact Assessment for the program.

DHS’s TRIP is a redress system for individuals who believe they have been improperly targeted by ATS, such as having been denied entry to the United States, refused boarding on an airline, or identified for additional border screening. Through TRIP, individuals can request correction of erroneous data stored in ATS or other DHS databases, Callahan said.

Callahan said that DHS’s privacy compliance policies and procedures governing the use of personally identifiable information follow the Fair Information Practice Principles, known as “FIPPs.”

The FIPPs underlie the 1974 Privacy Act and many federal agencies’ approaches to privacy protection, according to the Constitution Project report. The report cites the Federal Trade Commission as an example of an agency that incorporates FIPPs into its policies and procedures.

DHS, in a Dec. 29, 2008, memo memorializing DHS’s adoption of FIPPs, described the FIPPs as a set of eight principles: transparency; individual participation; purpose specification; data minimization; use limitation; data quality and integrity; security; and accountability and auditing.

The Constitution Project’s recommendations are consistent, but more specific than the FIPPs, Franklin said.

Oversight Where Transparency Not Possible. Franklin said that the recommendations in the report would also apply to data mining programs that are protected from public scrutiny for national security reasons.

Where full transparency is not possible, the recommendations would substitute oversight by Congress or the PCLOB. Just because a program is not disclosable to the public, does not mean that there is no oversight and accountability, Franklin said.

The PCLOB, formed on the recommendation of the 9/11 Commission, assists the executive branch in ensuring that privacy and civil liberties concerns are appropriately considered in the implementation of laws, regulations, and executive branch policies regarding counterterrorism. The PCLOB, which was originally established as an advisory board, was strengthened when Congress in 2007 recreated the board as an independent agency. However, a full board was never nominated by then-President Bush, and the board lapsed at the end of 2007.

President Obama has yet to nominate any members to the PCLOB, Franklin said. The Constitution Project has been urging the president to take action.