Executive Summary
Legal Memorandum
Published by The Heritage Foundation
No. 8
August 7, 2003

Proposals for Implementing the Terrorism Information Awareness System

Paul Rosenzweig

The Terrorism Information Awareness (TIA) program under development by the Defense Advanced Research Projects Administration at the Department of Defense has generated substantial controversy. Much of that controversy is unwarranted, and concerns that the technology will be abused are speculative, at best. A number of analogous oversight and implementation structures already in existence can be borrowed and suitably modified to control the use of the new technology.

As six former top-ranking professionals in America's security services recently observed, we face two problems—both a need for better analysis and, more critically, "improved espionage, to provide the essential missing intelligence." In their view, while there was "certainly a lack of dot-connecting before September 11" the more critical failure was that "[t]here were too few useful dots." TIA technology can help to answer both of these needs. Thus, TIA can and should be developed if the technology proves usable.

The technology can be developed in a manner that renders it effective, while posing minimal risks to American liberties, if the system is crafted carefully, with built-in safeguards that act to check the possibilities of error or abuse. In summary they are:

- Congressional authorization should be required before data mining technology (also known as Knowledge Discovery (KD) technology) is deployed;
- KD technology should be used to examine individual subjects only in compliance with internal guidelines and only with a system that "builds in" existing legal limitations on access to third-party data;
- KD technology should be used to examine terrorist patterns only if each pattern query is authorized by a Senate-confirmed official using a system that: a) allows only for the initial examination of government databases, and b) disaggregates individual identifying information from the pattern analysis;
- Protection of individual anonymity by ensuring that individual identities are not disclosed without the approval of a federal judge;
- A statutory or regulatory requirement that the *only* consequence of identification by pattern analysis is additional investigation;
- Provision of a robust legal mechanism for the correction of false positive identifications;
- Heightened accountability and oversight, including internal policy controls and training, executive branch administrative oversight,

This paper, in its entirety, can be found at:
www.heritage.org/research/homelandsecurity/lm8.cfm

Produced by the Center for Legal and Judicial Studies

Published by The Heritage Foundation,
214 Massachusetts Ave., NE, Washington, D.C. 20002–4999
(202) 546-4400 / heritage.org

enhanced congressional oversight, and civil and criminal penalties for abuse; and

- Finally, absolute statutory prohibition on the use of KD technology for non-terrorism investigations.

Critics of TIA are wrong to exalt the protection of liberty as an absolute value. That vision rests on an incomplete understanding of why Americans formed a civil society. As John Locke, the seventeenth-century philosopher who greatly influenced the Founding Fathers, wrote: "In all states of created beings, capable of laws, where there is no law there is no freedom. For liberty is to be free from the restraint and violence from others; which cannot be where there is no law; and is not, as we are told, a liberty for every man to do what he lists." Or, as Thomas Powers recently wrote: "In a liberal republic, liberty presupposes security; the point of security is liberty." Thus, the obligation of the government is a dual one: to protect civil safety and security against violence *and* to preserve civil liberty.

That goal can be achieved. To be sure, it is a difficult task. It is far easier to eschew the effort. But failure to make the effort—failure to recognize that security need not be traded off for liberty in equal measure and that the "balance" between them is not a zero-sum game—is a far greater and more fundamental mistake. Policymakers must respect and defend the individual civil liberties guaranteed in the Constitution when they act, but they also cannot fail to act when we face a serious threat from a foreign enemy.

Indeed, resistance to new technology poses practical dangers. As the Congressional Joint Inquiry into the events of September 11, pointed out in noting systemic failures that played a role in the inability to prevent the terrorist attacks:

> 4. Finding: While technology remains one of this nation's greatest advantages, it has not been fully and most effectively applied in support of U.S. counterterrorism efforts. Persistent problems in this area included a lack of collaboration between Intelligence Community agencies [and] *a reluctance to develop and implement new technical capabilities aggressively* . . . .

It is important not to repeat that mistake.

—Paul Rosenzweig is Senior Legal Research Fellow in the Center for Legal and Judicial Studies at The Heritage Foundation and Adjunct Professor of Law at George Mason University.

Legal Memorandum
Published by The Heritage Foundation
No. 8
August 7, 2003

Proposals for Implementing the Terrorism Information Awareness System

Paul Rosenzweig

The Terrorism Information Awareness (TIA) program under development by the Defense Advanced Research Projects Administration (DARPA) at the Department of Defense has generated substantial controversy. That controversy has led the Secretary of Defense to convene an advisory committee, the Technology and Privacy Advisory Committee (TAPAC), to provide him with advice on how, if at all, the TIA program should proceed. No doubt Congress will (and, indeed, ought to) weigh in as well.

It is therefore appropriate to begin asking a practical, concrete question: Can TIA be developed, deployed, implemented, and operated in a manner that allows it to be used as an effective anti-terrorism tool while ensuring that there is minimal risk that use of the TIA tool-set will infringe upon American civil liberties?

Some believe it is not possible. Critics of the TIA program believe that it is a "Big Brother" project that ought to be abandoned. They begin with the truism that no technology is foolproof: TIA may generate errors and mistakes will be made. And, as with the development of any new technology, risks exist for the misuse and abuse of the new tools being developed. From this, critics conclude that the risks of potential error or abuse are so great that all development of TIA should be abandoned. To buttress their claim that TIA should be abandoned, these critics parade a host of unanswered questions. Among them: Who will operate the system? What will the oversight be? What will be the collateral conse-

- Terrorism Information Awareness (TIA) is a potentially valuable tool in the war on terrorism under development by the Pentagon's Defense Advanced Research Projects Administration.
- "Knowledge Discovery" technology, if it proves successful, would enable the examination of a variety of databases in the effort to anticipate and prevent terrorist attacks on the United States.
- Critics of the TIA program believe that it is a "Big Brother" project that ought to be abandoned. An ill-considered Senate amendment to the pending Department of Defense appropriations bill would eliminate all funding for TIA programs.
- Much of the controversy is unwarranted, and concerns that the technology will be abused are speculative, at best. A number of analogous oversight and implementation structures already in existence can be borrowed and suitably modified to control the use of the new technology and protect our civil liberties.

This paper, in its entirety, can be found at:
www.heritage.org/research/homelandsecurity/lm8.cfm

Produced by the Center for Legal and Judicial Studies

Published by The Heritage Foundation
214 Massachusetts Ave., NE
Washington, DC 20002–4999
(202) 546-4400 heritage.org

quences for individuals identified as terrorist suspects?[1]

These questions are posed as if they have no answers when all that is true is that for a system under development, they have no answer *yet.* The same is true of any new government program; thus, we know that these implementation issues are generally capable of being resolved.

In fact, there are a number of analogous oversight and implementation structures already in existence that can be borrowed and suitably modified to the new technology. Thus, TIA can and should be developed if the technology proves usable. It can be done in a manner that renders it effective, while posing minimal risks to American liberties, if the system is crafted carefully with built-in safeguards to check the possibilities of error or abuse. This paper is an effort to sketch out precisely what those safeguards ought to be. In summary, they are:

- Congressional authorization should be required before data mining technology (also known as Knowledge Discovery (KD) technology) is deployed;
- KD technology should be used to examine individual subjects only in compliance with internal guidelines and only with a system that "builds in" existing legal limitations on access to third-party data;
- KD technology should be used to examine terrorist patterns only if each pattern query is authorized by a Senate-confirmed official using a system that: a) allows only for the initial examination of government databases, and b) disaggregates individual identifying information from the pattern analysis;
- Protection of individual anonymity by ensuring that individual identities are not disclosed without the approval of a federal judge;
- A statutory or regulatory requirement that the *only* consequence of identification by pattern analysis is additional investigation;
- Provision of a robust legal mechanism for the correction of false positive identifications;
- Heightened accountability and oversight, including internal policy controls and training, executive branch administrative oversight, enhanced congressional oversight, and civil and criminal penalties for abuse; and
- Finally, absolute statutory prohibition on the use of KD technology for non-terrorism investigations.

In short, TIA can be safely implemented. Failing to make the effort poses grave risks and is an irresponsible abdication of responsibility.[2] As six former top-ranking professionals in America's security services recently observed, we face two problems—both a need for better analysis and, more critically, "improved espionage, to provide the essential missing intelligence." In their view, while there was "certainly a lack of dot-connecting before September 11," the more critical failure was that "[t]here were too few useful dots."[3] TIA technology can help to answer both of these needs.

Indeed, resistance to new technology poses practical dangers. As the Congressional Joint Inquiry

1. *See e.g.* William Safire, "You Are a Suspect," *The New York Times.* November 14, 2002.; Susan Baer, "Broader U.S. Spy Initiative Debated," *The Baltimore Sun,* January 5, 2003. Senator Russell Feingold (D–WI) has already offered the Data-Mining Moratorium Act of 2003 (S. 188) to halt DARPA's research. See *http://thomas.loc.gov/cgi-bin/query/D?c108:1:./temp/~c108F8RHrG.* The Senate version of the pending Department of Defense authorization bill ends funding for the project for the coming year. *See* Paul Rosenzweig, Michael Scardaville & Ha Nguyen, "Senate Should Restore TIA Funding," Heritage Foundation *Web Memo* No. 315, July 17, 2003; *see also* Statement of Jay Stanley (ACLU) to the Terrorism and Privacy Advisory Committee, June 19 2003 (*http://www.sainc.com/tapac/library/June19JayStanley.pdf*) (urging that TIA be "shut down").
2. The Senate, in a remarkabe step, has recently taken this easy road. An amendment to the pending Department of Defense appropriations bill would eliminate all funding for TIA programs. This restriction is over-broad and ill-conceived and should be removed in conference. *See* Rosenzweig, Scardaville & Nguyen, *TIA Funding*. If, as one suspects, the funding limitation is the product of antipathy toward the housing of TIA research in a military organization, then rather than terminate the program prematurely, the research should be transferred to a civilian agency, like the Department of Homeland Security, for continuation.
3. Robert Bryant, John Hamre, John Lawn, John MacGaffin, Howard Shapiro & Jeffrey Smith, "America Needs More Spies," *The Economist*, July 12, 2003, p. 30.

into the events of September 11 pointed out in noting systemic failures that played a role in the inability to prevent the terrorist attacks:

> 4. Finding: While technology remains one of this nation's greatest advantages, it has not been fully and most effectively applied in support of U.S. counterterrorism efforts. Persistent problems in this area included a lack of collaboration between Intelligence Community agencies [and] *a reluctance to develop and implement new technical capabilities aggressively* . . . .[4]

It is important not to repeat that mistake.

Understanding the Question: What TIA Technology Really Means

Any implementation structure must be based on an accurate conception of precisely what systems are being put into operation. Unfortunately, the popular understanding of the TIA program is wildly at odds with reality.[5] Few people, including many of TIA's critics, seem to understand what the TIA program entails or how it would work.

TIA Is a Broad Research Program

TIA is, in fact, a broad research program with several dozen different components—programs ranging, for example, from efforts to develop machine language capabilities to translate Arabic directly into English, to ones dedicated to the development of information technologies that will permit creation of a secure Virtual Private Network where classified information can be exchanged without threat of compromise.[6] A number of these programs—which no public observer has critiqued as unacceptable—have nonetheless been delayed or questioned because they fit within the broad TIA umbrella.[7] Thus, the first goal of any statutory or regulatory structure implementing TIA is to clearly define its various components and focus the systematic controls and limitations on the components that are most problematic and most in need of oversight.

4. *Report of the Joint Inquiry Into the Terrorist Attacks of September 11, 2001*, House Permanent Select Committee on Intelligence and Senate Select Committee on Intelligence, 107th Cong., 2nd Sess., S. Rept. No. 107–351 and H. Rept. No. 107–792, Dec. 2002, p. xvi (available at *http://www.fas.org/irp/congress/2002_rpt/911rept.pdf*) (emphasis supplied). The Joint Inquiry also critiqued the lack of adequate analytical tools, *id.* Finding 5, and the lack of a single means of coordinating disparate counterterrorism databases, *id.* Findings 9 & 10. Again, aspects of the TIA program are intended to address these inadequacies and congressional limitations on the TIA research program are inconsistent with the Joint Inquiry's findings.
5. *E.g.* D.W. Rejeski, Symposium Comments from "Enhancing National Security and Civil Liberties in the Information Age," Woodrow Wilson International Center for Scholars, May 22, 2003; Heather Mac Donald, "Total Misrepresentation," *The Weekly Standard,* January. 27, 2003. Even among computer professionals there is substantial misunderstanding about the nature of the TIA and those with seeming greater familiarity with the technology are less apocalyptic in their reactions. *Compare* Executive Committee, SIGKDD, "'Data Mining' is NOT Against Civil Liberties," June 30, 2003 (available at *http://www.acm.org/sigkdd/civil-liberties.pdf*) (statement of the Special Interest Group on Knowledge Discovery and Data Mining of the Association for Computing Machinery ("ACM") supporting data mining technology) *with* ACM Public Policy Committee, Letter to Sens. John Warner and Carl Levin, Jan. 23, 2003 (available at *http://www.acm.org/usacm/Letters/tia_final.html*) (expressing doubts about TIA).
6. For a more thorough summary of the various development programs within TIA, *see* Department of Defense, "Report to Congress regarding the Terrorism Information Awareness Program," May 20, 2003 (available at *http://www.darpa.mil/body/tia/tia_report_page.htm*); *see also* DARPA, Overview Presentation to TAPAC (available at *http://www.sainc.com/tapac/library/TerrorismInformationOverview.pdf*); Paul Rosenzweig & Michael Scardaville, "The Need to Protect Civil Liberties While Combating Terrorism: Legal Principles and the Total Information Awareness Program," Heritage Foundation *Legal Memorandum* No. 6, Feb. 5, 2003, Appendix; Dr. John Poindexter, DARPATech 2002, Anaheim, California, August 2, 2002 (available at *http://www.fas.org/irp/agency/dod/poindexter.html*); Dr. John Poindexter, Dr. Robert Popp, and Brian Sharkey, "Total Information Awareness," *IEEE Aerospace Conference*, Big Sky, Montana, March 2003.
7. It has been publicly reported that the Wyden amendment, prohibiting any use of TIA in a domestic context, has prevented the FBI from coordinating the use of VPN technology with other agencies. *See* Paul Rosenzweig, "Securing Freedom and the Nation: Collecting Intelligence Under the Law," Testimony before the House of Representatives, Permanent Select Committee on Intelligence, April 9, 2003 (failure of FBI to sign Memorandum of Understanding with DARPA because of Wyden amendment).

Another aspect of the TIA program that often gets obscured is that it is a *research* program, not a development or implementation program. DARPA specializes in "outside the box," speculative technological research and development. When its research pays off (as in the development of the Internet), the benefits can be spectacular. But with far greater frequency its research programs produce much more modest (and sometimes non-existent) results.[8] While the prospect that some of the more advanced technologies will be developed is sufficiently great that it should be taken seriously, it bears repeating that the research project very well might not succeed. The success of the endeavor has yet to be determined—but at this stage the significant step is to recognize the research nature of the program, and thus to avoid strangling nascent technology in its crib by imposing unreasonable and unrealistic "proving requirements" long before the technology has had a chance to be explored.[9]

To be sure, the ultimate efficacy of the technology developed is a vital antecedent question. If the technology proves not to work—if, for example, it produces 95 percent false positives in a test environment—than all questions of implementation may be moot. For no one favors deploying a new technology—especially one that threatens liberty—if it is ineffective. There may, however, be potentially divergent definitions of "effectiveness." Such a definition requires *both* an evaluation of the consequences of a false positive *and* an evaluation of the consequences of failing to implement the technology. If the consequences of a false positive are relatively modest (as this paper suggests), and if the mechanisms to correct false positives are robust (again, as recommended in this paper), then we might accept a higher false positive rate precisely because the consequences of failing to use TIA technology (if it proves effective) could be so catastrophic. In other words, we might accept 1,000 false positives if the only consequence is heightened surveillance and the benefit gained is a 50 percent chance of preventing the next smallpox pandemic attack. The vital research question, as yet unanswered, is the actual utility of the system and the precise probabilities of its error rates.

All of which is merely another way of saying that the implementation of any new technology must be cautious and carefully weigh costs and benefits: "Any new intrusion must be justified by a demonstration of its effectiveness in diminishing the threat. If the new system works poorly . . . it is suspect. Conversely, if there is a close 'fit' between the technology and the threat (that is, for example, if it is accurate and useful in predicting or thwarting terror), the technology should be more willingly embraced."[10] For purposes of the implementation questions discussed in this paper, it is *assumed* that TIA—and in particular pattern-recognition technology—has demonstrated its general utility in rigorous testing. But it must be recognized that this assumption is no more than an assumption—it has yet to be borne out by practice.[11]

Understanding Knowledge Discovery Technology

After making the important definitional distinctions already noted, the legal control structure

8. To be sure, some of the TIA research components are highly speculative, as with, for example, the idea of creating a "futures market" in terrorism analysis. *See* Audrey Hudson, "Hill Hits Gambling Program on Terror," *The Washington. Times*, July 29, 2003, p. A1. But the reflexive opposition to this sort of speculative research by some members of Congress is, as one observer recently put it, "downright un-American." *See* David Ignatius, "Back in the Safe Zone," *The Washington Post,* August 1, 2003, p. A19. It is an example of the "zero defect" culture of punishing failures, not reason. Though a market in terrorist activity might prove unavailing, the only certainty at this point is that no one knows. It is particularly unfortunate when America's elected leaders oppose basic scientific research without the least sense that in doing so they demonstrate a "lack [of] the essential American willingness to take risks, to propose outlandish ideas and, on occasion, to fail." *Id.* That flaw is the way to stifle bold and creative ideas—a "play it safe" mindset that, in the end, is a disservice to American interests.
9. Some may say that conceptualizing an appropriate control structure for a system that has yet to prove its utility is, perhaps, a bit like imagining how we will govern Mars when we settle it—but given the controversy the research program has generated, and the unfortunate prospect that unreasonable fears of technological development may short-circuit research that has great potential benefits, the effort seems worthwhile.
10. Paul Rosenzweig, "Principles for Safeguarding Civil Liberties in an Age of Terrorism," Heritage Foundation *Executive Memorandum* No. 854, January 31, 2003; *see also* Paul Rosenzweig, "A Watchful America," *The Responsive Community,* Fall 2002, p. 89.

developed should thereafter focus on the aspects of TIA technology that have generated the most controversy. Most emblematic of those technologies is the program known as Evidence Extraction and Link Detection—the technology that, if developed and deployed, will enable the use of hardware/software mechanisms to access disparate databases. Closely related to this technology are projects such as Proximity and the Group Detection Algorithm,[12] which might develop means of analyzing data to create a graphic depiction of social connections between various individuals who might be linked in committing, for example, terrorist acts. The characterization of these technologies as "data mining" is clearly a misnomer, as the name leads to a factually inaccurate inference: that data is removed from a source (in the same way that "gold mining" implies the removal of gold). For simplicity's sake and because those with the greatest expertise in the field use a different nomenclature, this paper refers to this grouping of technologies as Knowledge Discovery ("KD") technology.[13]

Within the TIA data collection research program there are three further preliminary distinctions that must be understood before the proper legal structure for implementing KD technology can be developed:

First, KD technology can be used to access data that resides either in government databases containing information already lawfully collected by the government, or in private databases containing information collected by non-governmental (either commercial or non-commercial) databases. Though there will be difficult issues at the margin—e.g., databases merging private and government information or information requests that require access to both government and private databases—the utility of the distinction is clear. Inquiries that access only government databases pose far fewer civil liberties concerns than those that seek access to private databases.[14]

Second, KD technology can be used to access both foreign and domestic databases. Those databases can contain information about both "foreign persons"[15] and American citizens. Again, the distinctions drawn (which may also suffer from marginal definition concerns relating to merged databases) have powerful utility. Using KD technology, for example, to secure information about al-Qaeda terrorists from the foreign database of a ter-

11. There is, of course, substantial debate as to the likelihood that all aspects of the TIA program will prove feasible. This paper is not the place for such a debate, which is, in any event, well beyond the competence of the author to evaluate. It is, however, noteworthy that those most familiar with the technical issues think that many of the problems are solvable. *E.g.* Remarks, Jeff Jonas, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003 (question of resolving identities—that is, ensuring that data all refer to a single unique individual—is a "solved problem"). And, of course, very rough forms of pattern-recognition are used daily by companies such as Amazon (to predict book purchasing preferences) and Visa (to uncover credit card fraud). The possibility of success for more ambitious technology is therefore sufficiently great that consideration of control mechanisms is useful and confident predictions of absolute failure somewhat premature.

12. For a brief description of these programs see the DARPA Overview Presentation to TAPAC (available at *http://www.sainc.com/tapac/library/TerrorismInformationOverview.pdf*). More detailed summaries are available in the DOD *Report to Congress.* A prototype analog of these link detection programs, Coplink, has already been commercially developed and is in use by police departments around the Nation. *See* Gareth Cook, "Software Helps Police Draw Crime Links," *The Boston Globe*, July 17, 2003 (online edition). The commercial development of TIA-like technology demonstrates another flaw in the critics' argument: They are attempting to sweep back the tide. This sort of link technology, already commercially in use, will only improve and become more readily available, even if DARPA ceases all research efforts immediately. In a very real sense, the search for privacy may be an impossible battle. *See* David Brin, *The Transparent Society* (1999).

13. Remarks, David Jensen, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003. Strictly speaking, data mining involves only the examination of past cases to supply a rule for a predictive model of the future. Applying that predictive model to new data sets is decision-making analysis. *Id.*

14. Rosenzweig & Scardaville, *Civil Liberties*, at p. 15–16.

15. A "foreign person" is defined to include both individuals and foreign corporations or other entities. *See* 50 U.S.C. § 2369(d)(1). As discussed below, different rules should apply for the use of KD technology against wholly foreign entities and individuals.

rorist enemy is substantially less problematic than using the same technology to access information about an American citizen in a private, commercial American credit card database.

Third, and most significant, KD technology may be used in two distinct ways—means that have been described as "subject-oriented" and "pattern-based" data inquiries. Here, the need for a precise distinction is vital.

Through a subject-oriented query of databases containing information, KD technology might, by focusing on a specific individual (identified by name or other unique identifying characteristic) attempt to gain a more complete understanding of a suspect, his activities, and his relationships with others. In other words, KD technology would be used to access databases with information about a particular individual (e.g., government driver's license records, telephone toll records, or airplane flight records) to develop an understanding of his conduct. From this, one could develop a greater understanding of those he associated with—who shares his apartment? who does he call long distance? where do the funds in his bank come from? —that would produce additional leads and could, potentially, allow the development of a complete picture of an individual suspect, his associates, and their potential connection to terrorist activity.

By contrast, a pattern-based query is not focused on a specific uniquely identifiable individual or individuals. Rather, using existing intelligence data, intelligence analysts will develop detailed models of potential terrorist activities. The models will be developed in an iterative process: A group of analysts intended to replicate the conduct of potential terrorists (called "Red Teams" because American enemies traditionally are colored red on charts and maps) will conduct operations in a virtual (i.e., fictitious and artificial) world of cyberspace, creating data transactions (by securing driver's licenses, boarding airplanes, purchasing goods, etc.). They will repeat these operations for as many different terrorist scenarios as their imaginations will support. Then a separate team of analysts (the "Blue Team"), using the same intelligence data and their own imaginations, will try to develop database search inquiries that are capable of identifying the terrorist operation patterns created in virtual data space with a high degree of accuracy and distinguishing them from patterns of innocent transactional activity.

Thus, the utility of the effort will turn on its ability to accurately sift terrorist patterns from innocent patterns of activity. This accuracy may be usefully defined as the degree to which a database search inquiry identifies false positives (i.e., denotes as a potential terrorist pattern a pattern that is wholly innocent of terrorist connection) and false negatives (i.e., fails to identify an actual terrorist pattern). When (and if) a pattern-based data query proves successful in virtual, artificial cyberspace it might then (with the controls identified below) be used to examine real-world data.

Here, too, the distinction drawn is one that is likely to have legal and policy significance. A subject-oriented inquiry using KD technology raises questions about enhancing government efficiency, but it is, fundamentally, little different from existing law enforcement or intelligence practice. This sort of inquiry is most akin to a classic law enforcement "lead" inquiry. Some predication for suspicion about a particular individual (or individuals) exists and that predication leads law enforcement or intelligence officials to develop a more thorough understanding of the individual's actions. In a murder, case, for example, a detective starts with a list of known associates of the victim. Working outward from that list the detective develops information about each individual (and other individuals whose identity is revealed through association to those under initial inquiry) in an effort to determine who the killer is. KD subject-oriented technology does little more than render this process of expanding inquiry more effective (so that the "detective" misses fewer promising leads) and more efficient (so that the "detective" can conduct the inquiry more quickly).

Pattern-based inquiries, by contrast, are fundamentally a new conception—they focus not on predication about an individual but on predication about a pattern of conduct based upon analysis of the pattern, not the individual. This new conception of how law enforcement and/or intelligence agencies may develop investigative leads raises far more questions: Use of KD technology in this way is not merely an efficiency enhancement; it has the potential for substantially new and different intrusions into American privacy. [16]

The Scope of the Problem: What Controls Are Necessary?

Any structure of control and implementation will need to appreciate and accommodate these significant distinctions. Within the context of KD technology these distinctions give us the following general outline of principles that should guide any structure.

For foreign uses:

- Uses of KD technology against foreign targets are not generally subject to the same concerns as domestic uses. As a general rule such foreign uses should follow existing regulations and limitations on the collection of foreign intelligence.[17]

For domestic uses:

- Subject-based queries of government databases are the least problematic, as they involve queries of a form traditional in law enforcement or intelligence gathering on pre-existing databases to which the government already has almost unlimited access.
- Subject-based queries of private, non-government databases, though posing somewhat greater risks of privacy intrusion are, if suitably constrained, likely to pose no significant risk to civil liberty. Such queries, again, mirror traditional methods of inquiry and a wide-ranging consensus already exists that they are well supported in law and policy.[18]
- Because of the new and unique nature of pattern-based queries, the use of such inquiries on government-only databases poses special concerns (though less than those posed by inquires of private databases) and should be subject to particular limitations.
- Pattern-based queries in private databases are the most problematic and pose the greatest potential for misuse, mistake, and abuse. Thus, the greatest limitations and most careful checks need to be implemented before such technology is deployed.

The next question is: What are the mechanisms and structures that can be put in place that will protect civil liberties concerns regarding the use of KD technology in these four domestic contexts, while allowing the use of KD in narrow areas where it will do significant good in combating terrorism?[19]

As the prospect of new and different information technologies has grown, it has become apparent that existing laws do not adequately address limits on their use or establish suitable control structures

16. It bears emphasis that this technology—both the ability to access disparate databases and particularly the "pattern-based" aspects of KD technology—is highly speculative. It will require, for example, the development of software that is very sophisticated at finding relatively weak data "signals" in a sea of data "noise." It may well prove to be technologically unfeasible. But the utility of the research into the possibility cannot be doubted, any more than could be the utility of research into nuclear energy and weapons, though they were as highly speculative in their time as the KD project is today.
17. Important restrictions continue to exist on the authority of foreign intelligence agencies to conduct surveillance or examine the conduct of American citizens. *E.g.* Exec. Order No. 12333, 3 C.F.R. 200 (1982), *reprinted at* 50 U.S.C. § 401 note; Attorney General Guidelines for FBI Foreign Intelligence Collection and Foreign Counterintelligence Investigations (April 1983) (available at *http://cryptome.org/fbi-fic-fci.htm*); *see generally* National Academy of Science, "Legal Standards for the Intelligence Community in Conducting Electronic Surveillance (Feb. 2000) (available at *http://www.fas.org/irp/nsa/standards.html*). Conversely, however, the courts have recognized that in the national security context the requirements of the Fourth Amendment apply somewhat differently than they do in the context of domestic law enforcement. *See United States v. United States District Court (Keith)*, 407 U.S. 297 (1972) (applying Fourth Amendment in context of domestic national security surveillance); *In re Sealed Case*, 310 F.3d 717 (Foreign Int. Surv. Ct. Rev. 2002) (applying Fourth Amendment in context of foreign intelligence surveillance For the most part this paper focuses on the more controversial use of KD technology in the domestic context. Foreign uses of KD technology should continue to be governed by existing legal limitations.
18. *See* Markle Foundation "Protecting America's Freedom in the Information Age," October 2002 at p. 32.
19. Some doubt that the terror threat is sufficiently great to warrant this effort. The available evidence suggests that they are wrong. *E.g.* Michelle Malkin, *Invasion: How America Still Welcomes Terrorists, Criminals, and Other Foreign Menaces to Our Shore* (Washington, D.C.: Regnery Publishing Inc., 2002). More important, the consequences of accepting that vision and being proven wrong are potentially catastrophic. Prudent policy planning requires the contemplation of worst-case scenario possibilities.

for their deployment. Neither the Fourth Amendment nor privacy laws appear to have any general applicability to KD technology used on American databases or American citizens.[20] An appropriate legislatively or regulatorily created structure would have components addressing the following areas:

- **Pre-use restrictions.** What limits should be imposed before a search of domestic non-governmental databases is conducted using KD technology?
- **Identifying the subject.** What limits should be imposed before the results of a pattern-based search query, and in particular, uniquely individual identification data derived from a pattern-based query, are disclosed to government authorities?
- **Data Use and Error Correction.** What consequences should arise from a positive identification? What programs should be implemented to ensure that false positives that mistakenly identify individuals as potential terror suspects are quickly and permanently corrected?
- **Accountability and Oversight.** What structures and systems need to be put in place to ensure that any KD-based system is used accountably, for appropriate purposes, and that mistakes are corrected and misuse or abuse suitably punished?
- **Mission Creep.** How might a KD-based system be best structured to ensure that it is used only to fight terrorism and that it does not become a general law enforcement tool for whatever "war" on crime is currently in fashion?[21]

Mechanisms and Structures for Controlling the Domestic Use of KD Technology

Any suitable structure for controlling and limiting the use of a new investigative tool must answer two questions: what standard should be applied? and who decides? In other words, in considering how best to effectively use a new technology such as KD while ensuring that it is used within bounds that respect American liberties, we must ask the fundamental questions of when it is appropriate to use the technology (the standards question) and who is in charge (the authorization question). The control mechanisms envisioned in this paper propose answers to these two questions based upon: a) when the authorization is sought (before or after the data query), and b) which sort of query is proposed (a subject-oriented or a pattern-based inquiry).[22]

20. On the inapplicability of Fourth Amendment restrictions *see* Rosenzweig & Scardaville, *Civil Liberties* at p. 12–13; Paul Rosenzweig, "Anti-Terrorism Investigations and the Fourth Amendment After September 11," Testimony Before the United States House of Representatives Committee on the Judiciary Subcommittee on the Constitution (May 20, 2003). On the failure of existing privacy laws to address new information and data acquisition technologies *see* James X. Dempsey, *Privacy's Gap: The Largely Non-Existent Legal Framework for Government Mining of Commercial Data*, Center for Democracy and Technology (May 19, 2003).

21. The dangers of mission creep are significant. *See* Paul Rosenzweig, "Can The Use of Factual Data Analysis Strengthen National Security?" Testimony Before the United States House of Representatives Committee on Government Reform Subcommittee on Technology Information Policy, Intergovernmental Relations, and the Census (May 20, 2003); Rosenzweig & Scardaville, *Civil Liberties*, at p. 10–11. It is, therefore, imperative that certain applications that may be developed in the KD research program be limited to use in fighting terrorists who pose a true threat to national security. As detailed below, structures to achieve this limitation can and should be imposed.

22. A point of clarification is in order at the outset on how the system envisioned by this paper will interact with existing legal requirements. Current law distinguishes, broadly, two classes of information—law enforcement information and "intelligence" information, both of which are subject to separate regulatory regimes. *See generally* Exec. Order No. 12333, 3 C.F.R. 200 (1982), *reprinted at* 50 U.S.C. § 401 note. KD technology will, if it operates appropriately, have access to both categories of information and operate to collate and integrate them. The structures and mechanisms outlined in this paper suggest, in effect, that this process creates a third category of information—call it "domestic KD information" for want of a more euphonious name—and that this third category will be subject to a distinct (though complementary) legal and regulatory structure. The information accessed remains law enforcement or intelligence information insofar as it resides in place in the original data base where it was first collected and stored, but it take on a new character as KD information when combined through the operation of KD technology and used in a domestic context. As noted earlier, to the extent KD technology is used in a wholly foreign context, against foreign enemies and databases, then the existing foreign intelligence legal structure will suffice.

Pre-Use Authorization

As already noted, the general effectiveness of KD technology—and particularly pattern-based recognition technology—will need to be demonstrated before KD technology is authorized and implemented. But a general demonstration of effectiveness is different from an assessment of the utility of any particular search query. Both aspects of effectiveness will need to be addressed before KD technology is deployed.

With respect to the broader question of the general utility of KD technology, in light of the underlying concerns over the extent of government power, the best answer is that formal congressional consideration and authorization of the use of KD technology, following a full public debate, should be required before the system is deployed.[23] Before any KD technology—with both great potential utility and significant potential for abuse—is implemented, it ought to be affirmatively approved by the American people's representatives.

In making this decision, Congress should ensure that it bases its judgment on a sound technical understanding of the program. It might, for example, have the technology reviewed by an independent board consisting of non-governmental experts. Staffing for such an oversight board could come from experts identified by the National Academy of Sciences or some other respected scientific body, with additional representatives from the law enforcement and intelligence communities.

The remainder of this section addresses principally the second question: What pre-use authorization should be required before KD queries are deployed in specific cases? More specifically, what standards should be applied to assess the appropriateness of a particular proposed use, and who will decide whether or not the standard has been met?

Subject-Oriented Inquiries. In general, a subject-oriented inquiry is closely akin to the "normal" operation of routine law enforcement or intelligence activities—based upon some predication government agents have a factual basis (albeit one insufficient to be conclusive) for believing that a particular, identifiable individual or individuals are engaging or will be engaging in criminal/terrorist activity.

In the law enforcement context, the standard for the initiation of an investigation and for conducting further examination of a particular individual is minimal. No judicial authorization is needed for a government agent to initiate, for example, surveillance of a suspected drug dealer. All that is generally required is some executive determination of the general reliability of the source of the predication and, within the context of a particular agency, approval for initiation of an investigation from some executive authority. For example, the initiation of various anti-terror investigations is governed by guidelines promulgated by the Attorney General (in the case of domestic investigations) and an Executive Order (in the case of foreign intelligence investigations).[24] By contrast, routine drug or robbery investigations are often authorized at the local level by a unit supervisor.

To the extent that any such subject-oriented investigation requires access to data regarding the subject of the investigation, the same executive authorization suffices to permit a search of existing government databases for information regarding the individual. Such searches might include, for example, a check of the National Criminal Information Center for data on prior criminal convictions and a search of a state driver's license database for information on residence. By contrast, when the information is being collected from third-party data holders (such as telephone records or credit-card information) the government inquiry must proceed by way of subpoena—a method that affords the data holder the opportunity to object to production of the data if it is unduly burdensome or if the government seeks irrelevant information.[25]

23. *See* Rosenzweig & Scardaville, *Civil Liberties*, p. 9.

24. *See* Attorney General's Guidelines on General Crimes, Racketeering Enterprise and Terrorism Enterprise Investigations (May 2002) and Exec. Order No. 12333, 3 C.F.R. 200 (1982), *reprinted at* 50 U.S.C. § 401 note; *see also* Attorney General Guidelines for FBI Foreign Intelligence Collection and Foreign Counterintelligence Investigations (April 1983) (available at *http://cryptome.org/fbi-fic-fci.ht*).

With respect to subject-oriented queries of this sort, KD technology is best understood as enhancing the efficiency of the information gathering process. When an individual subject is identified, the use of KD systems simply accelerates the speed at which the information may be collected. Instead of taking hundreds of FBI agents several months to track the activity of 19 terrorists, the same information might (if KD works) be gathered in a few minutes, thereby affording the government the opportunity to short-circuit a terrorist strike proactively, rather than gather information reactively after the attack has already occurred. The best example of the potential success of this process is private technology that has already been used to identify non-obvious relationships among individuals. Using that technology and publicly available information, one could have readily "linked" the 19 terrorists.[26] Notably, the same analysis using non-public, government-held data identified the same 19 terrorists and 11 additional suspects.[27]

The only conceivable objection to the use of KD technology in this form is an objection against enhanced government efficiency. This is not an objection to be slighted as negligible—the entire conception of separation of powers and checks and balances that inheres in our constitutional system is premised on the notion that barriers to effective government action enhance individual liberty.[28] It cannot, however, be gainsaid that such principles stem from a time when the potential consequences of government inaction were less significant—in the current environment, the consequences might, literally, be the destruction of a city. Thus, in the unique context of terrorist threats that hold the potential for mass destruction, it appears advisable to relax our disapproval of an efficient government if suitable controls can be implemented. [29]

In principle, then, the subject-based inquiry using KD technology appears to be a welcome prospect. But the technology should not be deployed without limit, even in this traditional context. Rather, KD should be implemented only in a manner that mirrors existing legal restrictions on the government's ability to access data about private individuals—nothing more and nothing less. In other words, the implementing regulations and/or legislation should require that:

- Authorization for a subject-oriented use of KD technology be based only upon adequate predication to believe that criminal or terrorist activity is or is likely to occur;
- Any authorization be issued only on the basis of internal guidelines;
- Any authorization be embodied in writing, the records of which will be available for oversight and review;

25. *See* Rosenzweig & Scardaville, *Civil Liberties*, at p. 13–14. As more fully set out in *Civil Liberties*, the use of a subpoena authority in the KD context suggests the development of other substantive limitations including restrictions on access to protected government databases and an absolute prohibition on the creation of new government databases as a means of evading limitations on access to private databases.

26. *See* Remarks, Jeff Jonas, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003.

27. *See* Remarks, James Zimbardi, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003.

28. Some have called this the argument that increased efficiency is, in effect, the practical end of obscurity or anonymity. *E.g. Justice Dept. v. Reporters Committee*, 489 U.S. 749, 780 (1989) ("There is a vast difference between public records that might be found after a diligent search . . . and a computerized summary located in a single clearing house."). Notably this objection is in tension with another strand of the Framers' conception: One of the purposes of the constitutional structure was to ensure "energy in the Executive." *See Federalist No. 70* ("Energy in the executive is a leading character in the definition of good government.").

29. Some make another objection. They suggest that allowing KD technology access to information that is already "legally obtained and useable by the federal government under existing law" is a "major loophole to data mine 'everything under the sun.'" Audrey Hudson, " Pentagon to Dig Into Marketing Data on Citizens," *The Washington Times,* July 15, 2003 p. A1 (quoting Sen. Ron Wyden). On its face this argument lacks persuasive value: It is one thing to argue against greater government efficiency on power-enhancement grounds, but it is disappointing when one of the principal critics of TIA confuses that argument with one suggesting that KD technology will provide the government with access to databases to which it does not currently have access.

- Authorization may be granted only by a supervisory level officer within the investigating agency;
- KD technology that is developed for the use of accessing data held by non-government third parties must incorporate hard-wired programming that prevents access to data in the possession of non-government data holders without notice to and the consent of the data holder; and
- In the event the non-government data holder interposes an objection to access to the data, KD technology must provide a mechanism (whether electronically or, if necessary, through more conventional means) for the objection to be referred to an Article III judge (that is, one appointed by the President, confirmed by the Senate, and holding life tenure) for resolution using the same standards and principles that would govern such a data request today if it were made by means of a paper subpoena rather than, in effect, an electronic KD subpoena. [30]

Pattern-Based Queries. Unlike subject-oriented queries, pattern-based KD queries have no ready analog in contemporary law enforcement practice. Like the investigation of "tips" about a particular subject, pattern analysis is predictive in nature. Both may look to collect information about possible future criminal or terrorist conduct. But unlike subject-oriented analysis, which takes an identity and works outward to examine whether a criminal pattern exists, pattern-based queries presuppose the ability to successfully identify criminal patterns and propose, as it were, to work inward from the pattern to individual identities.

As has already been noted, the general utility and prospects for success of this sort of analysis are by no means certain. Many critics have suggested that it will routinely generate far too many false positives and miss too many terrorists.[31] Researchers in the field believe, however, that because of the high correlation of data variables that are indicative of terrorist activity, a sufficient number of variables can be used in any model to create relational inferences and substantially reduce the incidence of false positives.[32]

But even assuming the utility of the practice generically, KD pattern-based technology will also need to be vetted and approved on a more particularized basis. In other words, just because testing has established that some pattern-based inquiries may generate useful information about potential terrorist threats, it does not necessarily follow that any particular pattern-based query is well-constructed and ought to be deployed.

Plainly, the definition of a well-constructed query is subject to substantial debate. It reflects a policy judgment as to the degree to which false positives are tolerable in a new investigative system and an assessment of the consequences arising from tightening the inquiry to avoid false positives with a concomitant increase in the number of false negatives or in the failure to detect potential terrorist activity because the technology has not been deployed at all. Defining that balance point is difficult and it would be hubristic for this paper to offer a resolution. It would, however, be appropriate for our elected representatives to make that value judgment by providing guidance on the characteristics of an effective technology that would identify KD

30. Just as with subject-oriented queries, the examination of non-government databases for pattern-based KD inquiries must adhere to existing legal rules and provide for both notice and the opportunity to object. That requirement is a process value that does not distinguish based upon the nature of the KD search inquiry. With pattern-based queries, however, the level of "predication" that should be required and the identity of the authorizer should be different from those for subject-oriented queries.

31. *E.g.* Statement of Jay Stanley (ACLU) to the Terrorism and Privacy Advisory Committee, June 19, 2003 (*http://www.sainc.com/tapac/library/June19JayStanley.pdf*).

32. *See* Remarks, David Jensen, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003; *see also* K. A. Taipale, "Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data," 5 Colum. Sci. & Tech. L. Rev. (forthcoming December 2003) (critiquing false positive argument) (to be available at *http://www.stlr.org/cite.cgi?volume=5&article=2*); David Jensen, Matthew Rattigan, Hannah Blau, "Information Awareness: A Prospective Technical Assessment," SIGKDD '03 August 2003 (forthcoming).

pattern search methodology that ought to be deployed.[33]

The other question—and perhaps the more important one—is identifying who must approve the use of a particular pattern inquiry before it is deployed. Who, in other words, must make the decision that a pattern query is likely to be effective given the guidance provided in law or regulation? Here, many potential decision makers can be identified and the choice among them is, essentially, one of preference and practicality.

Operational constraints limit the ability for particularized oversight from outside the executive branch. Yet unfettered discretion is unacceptable. Particular pattern-based queries must be deemed effective prior to use—but requiring approval by either the legislative or judicial branch prior to their use would be cumbersome (and judges may be particularly ill-suited to the scientific judgments that will be required). Given that post-use structures (described below) can provide significant protection of individual liberty, the best analogous model would appear to be the requirement for "high-level" approval before a particular pattern-based query is deployed. As with, for example, the use of a subpoena to seek records from a press organization,[34] no action should be taken without the authorization of a Senate-confirmed officer of the Department of Justice, Homeland Security, FBI, or CIA (such as an Assistant Attorney General or the FBI Director).[35]

If this official deemed a particular pattern-based query sufficiently well-designed that its use was warranted, that determination should suffice to allow its deployment in the real-world data stream. The question then would arise as to how the response to the query should be structured in the real world and how those results should be handled.

Identifying the Subject: Breaking the Anonymity Barrier After a Data Query

One can envision any number of ways in which the response to a data query would be structured. It could, for example, be a comprehensive response, with uniquely identifying individual data about each of the individuals whose activities fit the pattern tested in the real-world data space. Equally plausibly, the response could be simply a number, representing the number of positive matches. The first would be overly intrusive, the latter meaningless. Indeed, the response structure might vary with different applications and uses for the technology. The general question, however, will remain the same: What response structure is best at enhancing liberty values while providing useful information?

Privacy and Anonymity. To answer that question one must have in mind the nature of the liberty value at risk. For some, the liberty in question is an absolute prohibition on government scrutiny—a pure privacy model, if you will. The law, however, has long recognized that the Constitution does not protect this conception of privacy against government intrusion when the object of the intrusion is the disclosure of criminal conduct.

Rather, since 1967, the Supreme Court has recognized that the Fourth Amendment protects only those things in which someone has a "reasonable expectation of privacy" and, concurrently, that anything one exposes to the public (i.e., places in public view or gives to others outside of his own personal domain) is not something in which he has a "reasonable" expectation of privacy—that is, a legally enforceable right to prohibit others from accessing or using what one has exposed.[36] Thus, an individual's banking activity, credit card purchases, flight itineraries, and charitable donations are information

33. Caution should be used in statutorily carving into stone a particular measure of effectiveness. One can envision a law that defines a system as effective if it is at least *X%* effective at identifying terrorist patterns in a virtual simulation and also is at least *Y%* effective in excluding innocent patterns from identification in the same virtual simulation. But, as noted earlier, that valuation must also take into account the magnitude of the consequences of a false negative. *X* and *Y* may differ depending upon the terrorist threat being addressed. Inasmuch as this paper recommends congressional authorization before deployment, this complex value judgment of multiple variables will, in the end, be conducted by our elected representatives, who will each reach their valuation in, no doubt, slightly different ways.

34. *See* United States Attorney's Manual § 9-13.400 (requiring authorization of Attorney General before issuance of subpoena to news organization).

35. This requirement would also be consistent with the authorization requirement outlined below that demands a certification that the pattern-based query is being used for anti-terrorist purposes, as a means of preventing mission creep.

that the government may access because the individual has voluntarily provided it to a third party. According to the Supreme Court, no one has any constitutionally based enforceable expectation of privacy in them. The individual who is the original source of this information cannot complain when another entity gives it to the government. Some thoughtful scholars have criticized this line of cases, but it has been fairly well settled for decades.

Congress, of course, may augment the protections that the Constitution provides, and it has with respect to certain information. There are privacy laws restricting the dissemination of data held by banks, credit companies, and the like.[37] But in almost all of these laws (the Census being a notable exception),[38] the privacy protections are good only as against other private parties; they yield to criminal, national security, and foreign intelligence investigations.

That balance should change, at least somewhat, when the question becomes the implementation of KD pattern-based technology. Most Americans readily understand that their individual privacy is subject to limits when there is predication to believe that they have committed criminal or terrorist acts. The existence of such predication is taken as a justification for breaching the wall of privacy. Where KD technology differs, however, is in the prospect that the search engine seeking patterns in data space will necessarily be obliged to scan and, if there is no match discard, data relating to the conduct of many innocent individuals as to whom there is no predication at all. The difference is between asking for the bank records of Al Capone and asking for all the records of all the customers at Capone's bank. If KD pattern-search technology is to be used, it can only be used on the understanding that data about innocents will be examined—how then to square that reality with the liberty and privacy expectations of Americans?

The answer lies in the concept of anonymity.[39] American understanding of liberty interests necessarily acknowledges that the personal data of those who have not committed any criminal offense can be collected for legitimate governmental purposes. Typically, outside the criminal context, such collection is done in the aggregate and under a general promise that uniquely identifying individual information will not be disclosed. Think, for example, of the Census data collected in the aggregate and never disclosed, or of the IRS tax data collected on an individual basis, reported publicly in the aggregate, and only disclosed outside of the IRS with the approval of a federal judge based upon a showing of need.[40]

What these examples demonstrate is not so much that our conception of liberty is based upon absolute privacy expectations,[41] but rather that government impingement on our liberty will occur only with good cause. In the context of a criminal or terror investigation, we expect that the spotlight of scrutiny will not turn upon us individually without some very good reason. Sampling of data that is discarded poses substantially less threat to individual liberty than sampling data and retaining the sample or using unverified samples to subject an individual to heightened scrutiny.

This idea of preserving anonymity unless and until a good reason for breaching the anonymity barrier arises can and must be hardwired into any

36. *Katz v. United States*, 389 U.S. 347 (1967). So, for example, federal agents need no warrant, no subpoena, and no court authorization to have a cooperating witness tape a conversation with a third party (because the third party has exposed his words to the public), *United States v. White*, 401 U.S. 745 (1971); attach a beeper to someone's car to track it (because the car's movements are exposed to the public), *United States v. Karo*, 468 U.S. 705 (1984); or search someone's garbage, *California v. Greenwood*, 486 U.S. 35 (1988).

37. *E.g.* 12 U.S.C. §§ 3402, 3403 (bank disclosure); *id.* § 3407 (subpoenas to bank);

38. *E.g.* 13 U.S.C. §§ 8, 9 (prohibition on disclosure of Census data); *id.* § 214 (penalties for disclosure).

39. *See* Phillip Kurland, "The private I," *The University of Chicago Magazine*, Autumn 1976, p. 8 (characterizing three facets of privacy, broadly characterized as anonymity, secrecy, and autonomy) *quoted in Whalen v. Roe*, 429 U.S. 589, 599 n.24 (1977).

40. *E.g.* 26 U.S.C. § 7213 (prohibiting disclosure of tax information except as authorized for criminal or civil investigations).

41. *But cf. Lawrence v. Texas*, -- U.S. --. 123 S.Ct. 2472 (2003) (recognizing that certain intrusions into individual privacy are beyond governmental power).

KD technology. The protections should take three forms:

- First, the system must be built to ensure that no data scanned for match to a pattern inquiry is retained if the data does not fit the pattern. The algorithms must scan the requisite databases for information matching the pattern query but absolutely no data that fails to match the pattern query should be retained in any independent government database.
- Second, the system must also ensure that data scanned that does not match the pattern query is never presented to an analyst for examination. The initial scanning must be automated and structured to prevent unauthorized access of this sort.
- Third, and more important, the implementing legislation or regulations should mandate that any KD technology developed should structure the pattern query response in a manner that:
  - uses the search query in successive iterations, looking first only in government or other public databases (or perhaps subsets of government or public databases) and permiting the query to be used on private, non-government databases only after pattern-based inquires on government and public databases provide the basis for expansion of the universe of data examined to private; and
  - initially disaggregates the pattern-based inquiry results from the identity of the individual(s) who match the pattern. In other words the response provided must include only information about the activities identified that form the alleged pattern of terrorist activity—absolutely no uniquely identifying individual information should, at the outset, be provided.

These later requirements provide a dual structural mechanism that will have the effect of preserving civil liberty and privacy. The successive iteration requirement (sometimes called the use of primary and secondary databases)[42] will ensure that those databases containing information that individuals consider the most private will not be examined absent a suitable showing of cause. The disaggregation requirement will have the effect of absolutely preserving anonymity at the initial stage of any pattern-based inquiry. Thus, those developing KD should be required to construct a system that initially searches non-public databases and disaggregates individual identifiers from pattern-based information. Only after the pattern is independently deemed to warrant further investigation should the individual identity be disclosed and more sensitive non-government databases examined. At the first iteration, those using KD technology will initially be made aware only of a potentially suspicious pattern, but not of the identity of the individual whose pattern has been identified.[43]

Thus, one aspect of the TIA program, the Genisys Privacy Protection program, is to be welcomed by everyone on both sides of the discussion. The Genisys program is developing filters and other protections to keep a person's identity separate from the data being evaluated for potential terrorist threats. In authorizing KD technology, Congress could mandate that a trusted third party governmental agency (as suggested below in the related context of accountability, perhaps the Inspector General of the agency where KD-technology is housed) rather than the organization's database administrator control these protections. This methodology would ensure that the privacy protections are not being circumvented.[44]

We must also realize that the use of enhanced technology does not uniformly impose costs on lib-

42. *E.g.* Remarks, David Jensen, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003.

43. Thus, as already noted, access to data is not necessarily equated with a loss of privacy. To be sure, it may in some instances amount to the same thing, but it need not. There is, for example, a sense in which the automated screening of personal data by computer enhances privacy: It reduces the arbitrariness or bias of human screening and insures that an individual's privacy will be disrupted by human intervention only in suspicious cases.

44. *See* ISAT, "Security with Privacy," December 2002 (noting potential for technological protection of privacy) (summary available at *http://www.darpa.mil/iao/secpriv.pdf*); *see also* Markle Foundation "Protecting America's Freedom in the Information Age," October 2002 (same).

erty—there are potential benefits as well. One could, conceivably, adopt a purely preventative mode in responding to terrorist threats, enhancing security at airports, government buildings, and the like and relying on increased physical intrusions and identity cards as a means of forestalling the next attack. But if we are not to condemn ourselves to the "citadelization" of America, we must also consider a different tack—the use of predictive technologies to attempt to anticipate and thwart terrorist attacks before they occur. These technologies come at some potential costs to liberty, but with the very real prospects of gains in other forms of liberty. Absolute protection for electronic privacy necessarily leads to even less physical privacy.

Judicial Review. The final step—and the critical one for striking a suitable balance between the use of technology to enhance the ability to predict terrorist attacks while adequately protecting liberty interests—lies in the interposition of a judicial officer before the barrier of anonymity is broken. In other words, once a pattern of potential terror activity is identified, the user of a KD pattern search ought to be obliged to present that information to a court—in effect, the equivalent of the court currently used to implement the Foreign Intelligence and Surveillance Act (FISA), if not that court itself.[45] Only after the judge determines that a basis exists for concluding that the pattern identified is, in fact, a pattern of potential terrorist activity and not merely a coincidental pattern of innocent activity ought the identify of the actor whose pattern is in question be provided to law enforcement or intelligence officials. This mechanism—sometimes called selective or progressive revelation—should be built into the KD technology from the beginning, not added as a supplement after the structure of the technology has been substantially developed.[46]

It is absolutely vital to both the success of the KD technology and the protection of civil liberty that this judgment is made by a neutral third party. And because of the nature of the inquiry—determining the basis for official scrutiny and intrusion by law enforcement or intelligence officials—is one that has historically been conducted by judicial officers, it is fully appropriate to rely on that model for control of the system. It has proven quite successful in the FISA context, and there is no reason that it could not be directly adapted to the use of KD pattern technology.

What standard, then, should the "KD judge" apply in determining whether or not to disclose the identity of the pattern creator? What showing ought to be required of the government? The answer to those questions turns directly on the answer to another question—what are the consequences of a positive identification? For the more severe the consequences, the higher the standard necessary, and conversely, the less significant the consequences, the lower the standard acceptable.

The law has dealt with this in a variety of traditional ways: Before one may be arrested, for example, the government must have probable cause to believe that a specific individual has committed an offense.[47] Similarly, before the police may intrude in an individual's home, they must secure a search warrant, supported by probable cause.[48] By con-

45. FISA establishes a legal regime for "foreign intelligence" surveillance separate from ordinary law enforcement surveillance. *See* Pub. L. No. 95-511, 92 Stat. 1783 (codified as amended at 50 U.S.C. § 1801 *et seq.*). Under FISA, a separate court, the Foreign Intelligence Surveillance Court, issues warrants and authorizations for governmental activity relating to counter-intelligence or counter-terrorism investigations. *See* 50 U.S.C. § 1801(f) (wiretaps); *id.* §§ 1821-1829 (covert physical entries); *id*. §§ 1841-1846 (pen registers); *id.* §§ 1861-62 (business records).

46. *See, e.g.* Taipale, *Data Mining and Domestic Security* (arguing that selective revelation will secure privacy). The utility of selective or progressive revelation can be enhanced through other technological means of insuring privacy, some of which are currently under development. One can conceive, for example, of "one-way hash" functions (akin to those used to create encryption algorithms) that would allow anonymous pattern matching to occur between two distinct data holders who would be advised only of the match and would be obliged to go through a legal process similar to that described in text to secure the matching identification information. *See* Remarks, Jeff Jonas, "Data Mining in the Private Sector," Center for Strategic and International Studies, July 23, 2003.

47. *United States v. Watson*, 423 U.S. 411 (1976).

48. U.S. Const. Amend. IV; *Illinois v. Gates*, 462 U.S. 213 (1983).

trast, if the nature of the intrusion is less—if, for example, the police are making a brief investigative stop for the purposes of questioning—then all that law enforcement requires is some "reasonable suspicion" or "articulable suspicion" that will form the basis for investigation.[49]

Using this paradigm as a model, the most reasonable answer relies on the use of KD pattern identification solely as a predicate for further investigation. In other words, pattern-based identification is an investigative, not an evidentiary tool.[50] In this case, the "KD judge" should be required to determine whether the pattern presented raises a reasonable suspicion of potential terrorist activity. If the KD pattern identification is used for something more—for example, to place a named individual on a watch list for flights or to deny a named individual employment in a secure facility, then a higher standard of probable cause ought to be required.

The Consequences of Identification

KD technology will operate, if it ever successfully operates at all, on the basis of predictive analysis. As such, it will never provide more than a basis for suspicion from a pattern of activity. And, qualitatively, there can be little doubt that the predictive power of KD identification is of substantially less probative value than is direct information about the conduct of a particular subject. Indeed, KD advocates have never contended otherwise—to the contrary they have often characterized pattern-recognition analysis as akin to the very difficult task of picking out a barely audible signal in a sea of noise.[51] Thus, all recognize that positive pattern identification provides one of the weaker forms of inference about a suspect's potential activity.

Given this ineluctable nature of the pattern inference, its limitations should be recognized and acknowledged through restrictions on the uses to which it may be put. In particular, the KD pattern identification is most analogous to a policeman's investigative suspicion—the product of observations, experience and other factors, many of which are capable of limited articulation and quantification. As the Supreme Court recognized in *Terry,*[52] such suspicions, while frequently in error, are, if based upon articulable observations, often the cornerstone of preventative law enforcement activity. They ought to be fostered as good police work, but given only limited scope. In the context of KD pattern recognition, that means that the *only* consequence of a positive identification for a particular individual should be that the individual be subject to further investigation by normal investigative means within the existing confines of the law. Depending on the information disclosed, that investigation might involve surveillance, interviews with the subject or his associates, or more intrusive investigative techniques (most of which will independently require the approval of a judicial officer). *No other collateral consequences should be permitted absent some corroboration of the initial predication by other investigative means.*

As a corollary to this principle, it also follows that the initial step of breaking the anonymity barrier should only be authorized if a judicial officer finds that there is a reasonable basis for suspicion that a specific individual has engaged in terrorist activities. Requiring more would impose a burden of proof on KD-based predication that is more substantial than in any other investigative context. Requiring less would effectively eliminate any constraint on the technology's use.[53]

---

49. *Terry v. Ohio*, 392 U.S. 1 (1968).

50. Taipale, *Data Mining and Domestic Security.*

51. Robert Popp, Deputy Director of DARPA IAO, *quoted in* Erika Jonietz, "Total Information Overload," *Technology Review,* July/August 2003 (available at *http://www.technologyreview.com/articles/impact0703.asp*)

52. 392 U.S. at 21–22.

53. The TIA program has generated a substantial (though unfounded) fear. As a consequence, though this paper recommends a "reasonable suspicion" standard, if, as a political matter, it were deemed appropriate to enhance the restrictions on the use of KD technology by raising the standard required for disclosure of an individual's identity to "probable cause," that change would diminish the utility of the system but, one hopes, not to an appreciable degree.

Dealing with False Positives

No system constructed by man is perfect. The only certainty is that there will be false positives—both in investigations and possibly (though less likely) in the mistaken imposition of collateral consequences on a misidentified subject. To some, this prospect is a reason to forgo the entire possibility of KD technology. To others, it is no ground for any concern. Neither view is tenable. A more balanced approach would recognize certain particular aspects of the question:

General Principles. First, and foremost, the question of false positives is not unique to KD. Indeed, all law enforcement or intelligence activity will, on occasion, result in the identification of a subject who proves, upon closer examination, to have done nothing wrong. Thus, the dilemma posed by the problem of false positives through KD technology is, in one sense, nothing new.

As a consequence, implementing laws or regulations should specify that, to the degree that it recapitulates already encountered problems with investigative activity, the law applicable to KD should embrace the same remedies that have been used in the past. For misidentification as a subject that is the product of a good faith inquiry, the law currently allows little or no redress—for the good and sufficient reason that we do not want to deter good faith examination of criminal conduct.[54] All the more so, it would seem, for investigations of terrorist activity. However, as a general matter, the grossly or willfully negligent identification of a subject can, and should, subject one to tort remedies, just as it would outside the context of a KD inquiry.[55]

Nor should the use of KD technology change the rules for those who are subject to "traditional" law enforcement or intelligence inquiry because they are identified as a suspect based upon information (a law enforcement tip or intelligence information) and are named in a subject-based data inquiry using KD technology. To the extent that inquiry relies upon information from already existing government databases, these individuals, even if later determined to have been mistakenly named as a subject, will have no independent basis for seeking to correct the government databases themselves—the information contained in them was lawfully collected for other purposes and is not subject to correction.[56] Put another way, they should have no greater right to redress based upon the use of KD technology to conduct the examination of their conduct because the KD technology itself was not the source of the misidentification.

And, of course, if the subject-oriented query and further investigation produce evidence of involvement in a terrorist act or of another crime, the individual who is named in the subject-based query will have no cause for complaint. Identifying a potential terrorist is, after all, precisely what the KD technology is for. And if evidence of other crimes (e.g., a deportable immigration violation or drug crime), is developed collaterally, there ought to be no cause for concern, *so long* as the standards already articulated regarding the initiation and conduct of a subject-oriented KD investigation are upheld. In these circumstances, collateral discoveries of crime should be utilized, as their disclosure is a mere fortuity.

Pattern-Based Queries and Watch Lists. This is not true, however, of the pattern-based query using

54. Generally, one may not secure damages for a violation of the Fourth Amendment by a law enforcement officer unless the officer has violated a clearly established constitutional norm. *Anderson v. Creighton*, 483 U.S. 635 (1987). Absent such a clear norm, the officer is immune from suit. Similarly, injunctive remedies are extremely difficult to secure. *Los Angeles v. Lyons*, 461 U.S. 95 (1983).

55. Such misconduct, because it violates clearly established constitutional norms, is actionable under 42 U.S.C. § 1983 if the officer is a state official and under the doctrine of *Bivens v. Six Unknown Named Agents*, 402 U.S. 388 (1971), if the officer is a federal official.

56. Under the Privacy Act, 5 U.S.C. §§ 552a(d)(2), (g) an individual has the right to request amendment and correction of a record pertaining to him and may sue civilly if the government refuses to amend the record. However, law enforcement records, classified information, and CIA records are exempt from this provision. *See id.* §§ 552a(j), (k). As set forth below, the Privacy Act will need to be amended to conform to the right of correction that will attach to law enforcement or intelligence records examined using pattern-based KD technology that prove after investigation to have been unfounded.

KD technology. Even if (as this paper advocates) there are no adverse collateral consequences to identification based upon a pattern (e.g., no placement on a no-fly list), there is, nonetheless, a different direct consequence of the use of the technology: The individual in question is subject to scrutiny, in the form of investigation and surveillance, to which he would not otherwise have been subject. Generally, the primary cause of misidentification in the traditional law enforcement or intelligence context is the bad information provided from an outside source. Using KD technology, one may become a misidentified subject of an investigation even based upon underlying data that is wholly reliable. It is the premise of the inquiry—the presumed pattern—that is necessarily in error.[57]

Society is willing to accept the risk to some individuals of enhanced surveillance if it is a consequence of good-faith reliance on predication developed through a traditional law enforcement means. But fitting a "pattern" is a circumstance subject to potential repetition and based upon less powerful inferences of predication. It is more likely to subject innocents to investigation than traditional law enforcement methods. For that reason, some method of correcting false positive errors derived from pattern-based queries must also be a part of any structure of controls implementing KD technology. That method should, in outline, have the following components (though other similar structures can also be conceived):[58]

First, when a pattern-based query is authorized and when disclosure of the identity of the individual has been approved by a judge, government officials will be obliged to conduct a further investigation of the individual identified. If those investigations lead the government to "clear" the individual of any connection to terrorist activity, the investigators should be under an affirmative obligation to ensure that no collateral consequences inhere to the identification and that if any collateral consequences have inhered, they are removed immediately. Thus, if identification has led by mistake to placement on any sort of watch list, the unique identification must be removed. Moreover, because pattern analysis of this sort is capable of repetition and thus capable of again mistakenly identifying a particular individual as a suspect, investigators should also be under an affirmative obligation to ensure that those cleared of any suspicion are not, again, identified.

The more difficult and challenging question arises when the results of the investigation are ambiguous—that is, when the investigation does not "clear" an individual, but the evidence collected is of insufficient strength to allow for prosecutive action. In other words, what happens if the answer after investigation is "maybe," such that it would be irresponsible of the government to ignore the evidence (that is, the individual should be placed on some form of "watch list" because of valid suspicions that are insufficient to allow for prosecution), yet equally inappropriate for the individual to be permanently affected without being advised of the effect. One can hope that such ambiguous situations are few, but they may prove fairly commonplace.

It bears emphasis, however, how narrow the range of cases discussed here is. First, it involves only individuals identified by a pattern-based query. Second, it involves only those individuals as to whom a judicial officer has already found reasonable suspicion of terrorist activity derived from the pattern-based query, thereby authorizing their individual identification. Third, it involves only those individuals as to whom, after subsequent investigation, the conclusions are ambiguous. And, fourth, it involves varying and sometimes minimal levels of residual suspicion—some watch-listed individuals may be placed on a "no fly" list, but others may only have heightened screening of their bags because the residual questions about them are comparatively less significant. If this system operates as envisioned, this narrow class of individuals will be one that most Americans will agree are justly subject to scrutiny and are not merely being scrutinized for random or invidious reasons.

57. The underlying data in the public or governmental domains may also be in error—an issue discussed *infra*. But that is a circumstance that might also apply outside the KD pattern search context. It is the use of a new search methodology as the causative basis for enhanced surveillance that renders KD different and unique.

58. As discussed more fully below, in addition to the mechanisms addressed here, the result of all pattern-based queries will need to be fully documented for oversight and review purposes.

Nonetheless, in such situations, the ultimate burden should be on the government to justify any permanent or lengthy deprivation of civil liberties (again, remembering that all intrusions are not equal in nature). And the government should also be under an affirmative obligation to afford the investigated individual notice of the investigation and the ambiguous nature of its resolution. To be sure, such notification need not be immediate—it might perhaps follow the end of the investigation by 90 or 120 days. But equally clearly, if as a result of the investigation, the government believes it is appropriate to impose upon an individual adverse, non-punitive collateral civil consequences, it ought not to be allowed to do so without providing the individual with notice of that decision.

Nor should it be able to enforce those consequences indefinitely. There ought to be a presumptive time frame, of (again) perhaps 90 or 120 days after notification to the individual is provided within which the individual's name could be maintained on a watch list, or other collateral consequences imposed, before that decision is reviewed and confirmed (or rejected) by an independent, neutral arbiter—that is, a judge. The time frame might be longer for less significant intrusions (such as enhanced baggage screening) or shorter for more intrusive ones (such as a "no fly" limitation).

As an initial matter, there should not be direct review by a court. The implementing legislation or regulations should, instead provide for administrative review of this essentially civil decision to impose collateral consequences. At this administrative hearing the individual should have the full panoply of due process protections, including the right to be heard and the right to be represented. More important, there should be a private right of action to appeal any adverse administrative decision to a federal district court. And there, unlike the normal case for the review of an administrative agency action,[59] the review by the federal court should be *de novo*.[60] One could, of course, imagine equivalent mechanisms for review that would be equally protective—the one proposed is merely one model.

In adjudicating any such case (through whatever mechanism adopted) the subject on whom adverse consequences are imposed cannot be placed with the burden of establishing his innocence—such a showing is virtually impossible as it would require proof of an almost unprovable negative. Thus, once an investigative subject comes forward with a prima facie case establishing a basis for believing that his continuing presence on any watch list is without foundation, the burden should shift to the government. In order to maintain an individual on any such list or continue the imposition of other collateral consequences, the government should be obligated to prove by clear and convincing evidence (as in the case of pretrial detention)[61] that: a) for significant intrusions such as a "no fly" determination, the subject poses a substantial risk to the community, or b) for more modest intrusions such as additional baggage screening, the subject poses a potential risk. Here, too, a full panoply of due process rights (as with any civil case) ought to be afforded to the subject.

Maintaining Records. Finally, there is the consideration of what to do with the records of the subject-based or pattern-based inquiry when the inquiry and further investigation produce no evidence of any crime or link to terror. In one sense, it seems appropriate to require that all such records be expunged from the system. But expunging the records will make accountability of the system, assessment of its effectiveness, and oversight of its use impossible—without data about mistakes the

59. *See Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc.*, 468 U.S. 1227 (1984) (requiring deference to agency decision making by courts reviewing decision).

60. As noted, these review and correction provisions will require amendment of the Privacy Act to permit them to operate.

61. This standard for pre-trial detention was approved in *United States v. Salerno*, 481 U.S. 739 (1987). The review proposed is consistent with this standard. Indeed, inasmuch as contemporary jurisprudence would almost certainly allow a suspect's detention based upon such a showing, the proposed review process is overprotective of civil liberty, as it requires the government to meet the same high standard of proving dangerousness in order to impose certain less onerous and intrusive collateral consequences. Again, an alternate model might approve a lesser standard without appreciably trenching on civil liberties—and that judgment, again, must be left to our legislative bodies.

system makes its operation cannot possibly be properly assessed.

Thus, the better solution is to designate a trusted third party government agency as the repository for information about false positives. The characteristics of such a trusted third party would be an entity already within the government (thereby ensuring security of the data) but with an independent role (thereby insuring neutrality). The obvious candidate for this is the Inspector General of the agency where the KD technology is housed, to whom should be transferred all records of "false positives" where either investigation or a judicial determination has resulted in a conclusion that no further records of the individual in question should be retained.

Accountability and Oversight

Having now set forth a stringent series of restrictions and limitations on the use of KD technology, it remains for any implementing legislation to set forth the mechanisms by which adherence to those restrictions and limitations will be ensured. There is, as in any system of governmental activity, a clear need for accountability within implementing agencies and for oversight by outside bodies on the use of KD. Several components will need to be a part of any effective accountability and oversight system.

First, there needs to be a series of internal controls on the use of KD technology. Those controls will begin at a tactical level with the creation of an internal organization, fully a part of the KD decision-making and utilization network, with responsibility for the development and implementation of operational guidelines. This internal policy node will provide support to the intelligence, law enforcement, and operational components of the system by specifying the circumstances in which various activities are authorized and by adjudicating policy issues and disputes as they arise in problem-specific instances.

This internal policy organization (call it the "KD Council," though doubtless better names can be proposed) should consist of high-level officials from all of the organizations that are implementing KD technology (FBI, CIA, etc.). For ease of service, it is likely that the Council would best be housed either in the new Terrorist Threat Information Center, which already has the function of coordinating counter-terrorism and law enforcement activities, or if this paper's suggestion of housing the operational control of the technology within the Department of Homeland Security is adopted,[62] within that Department's new intelligence bureau.

Beyond this tactical internal control there is also a need for other internal mechanisms that will foster accountability and prevent misuse. Guidelines, such as those to be developed by the Council are of little use if there is not a culture of accountability within the organizations that use the new technology. An understanding of the legal limitations applicable and a bona fide commitment to adherence to those limitations is worth far more than all the structural mechanisms and rules combined.

Creating such a culture is no easy task. And in some instances we clearly have failed—one thinks, for example, of some of the abuses of the FBI during the 1960s. But at the same time, it is equally clear that the task is achievable. The National Security Agency is a case in point. It has long operated under stringent legal rules that restrict its activity—its systems may not, for example, be used to conduct electronic surveillance of "United States persons." By all accounts, even those of its critics, this limitation has been successfully internalized to the culture of the organization and breaches of it are virtually unknown.[63] Much the same approach must be taken within the agencies that deploy pattern-based search technology.

Next, the architecture of the hardware and software of the KD system must be developed to ensure accountability. At least three components need to be "programmed in" to the system:

1. Any use of the KD system should produce an audit trail, so that a hard record of the use of the system is maintained for subsequent review and analysis;
2. The audit trail must, to the maximum extent possible, be tamper-proof, so that those who

62. *See infra* at p. 23.

63. Remarks, John Hamre, "Enhancing National Security and Civil Liberties in the Information Age," Woodrow Wilson International Center for Scholars, May 22, 2003.

misuse the system cannot conceal their activity by altering the audit trail; and

3. To the extent that the audit trail cannot be rendered tamper-proof, it must be rendered tamper-evident, so that even if a specific misuse cannot be conclusively identified, the very fact of misuse is apparent.

In short, the components of the system must be built so that abuse is difficult to achieve and easy to uncover.[64]

System architecture alone cannot do the job of controlling misuse. Implementing regulations should also provide for independent, internal review of the operation of the system. Ideally, the Inspectors General of the agencies using the KD system would be tasked with conducting periodic internal audits of the use of the system and reporting on the results of those audits.

This leads, inevitably, to the most important source of oversight—Congress. Since much of the operation of the KD-based system will involve classified information, the mechanism for oversight must account for that fact. But the fundamental point remains: Congress must commit at the outset to a strict regime of oversight of the TIA/KD program. This would include periodic reports on the technology's use once developed and implemented, frequent examination by the General Accounting Office, and, as necessary, public hearings on the use of KD technology. Congressional oversight is precisely the sort of check on executive power that is necessary to insure that KD-based programs are implemented consistent with the appropriate limitations and restrictions. Without effective oversight, these restrictions are mere parchment barriers. While potentially problematic, one can be hopeful that congressional oversight in this key area of national concern will be bipartisan, constructive, and thoughtful. Congress has an interest in preventing any dangerous encroachment on civil liberties by an executive who might misuse pattern-based technology.[65]

Added to this, there should be annual public reports on non-classified aspects of the program. The public has a general right to know information about the operation of KD technology. While the public disclosure of specific search parameters, for example, would compromise the utility of the system by disclosing its methodology to terrorists, the public and the press are clearly entitled to information about the technology's operation. As with any other intelligence activity, sources and methods should be protected—but the agency implementing KD search technology should report annually on aspects of the program whose disclosure will not endanger national security. For example, the agency can report on the gross numbers of such searches conducted, the frequency with which the searches produce positive results and the frequency with which positive KD searches have been borne out (or refuted) by subsequent investigation. Gross and aggregate information for assessing the utility of the technology must be available publicly.

Finally, any new law or regulatory system should have severe administrative and criminal punishments for misuse and abuse of the system. Termination of employment and imprisonment are sanctions of last resort and it is to be hoped that they will never be used. But as a matter of deterrence and just punishment, the law must provide for significant sanctions for the knowing misuse of the new KD technology. In this way, Congress will enlist the third branch of government, the courts, to serve as a further check on potential abuse. Thus, in addition to penalties, Congress should also authorize a private right of civil action for injunctive relief, attorneys' fees and, perhaps, monetary

64. *See* Taipale, *Data Mining and Domestic Security* (arguing for strong audit requirements).

65. The Heritage Foundation has written extensively on the need for reorganization of the congressional committee structure to meet the altered circumstances posed by the war on terror and the formation of the Department of Homeland Security. *See, e.g.*, Michael Scardaville, "The New Congress Must Reform Its Committee Structure to Meet Homeland Security Needs," Heritage Foundation *Backgrounder* No. 1612, November 12, 2002; Scardaville, "Congress Must Reform Its Committee Structure to Meet Homeland Security Needs," Heritage Foundation *Executive Memorandum* No. 832, July 12, 2002. Oversight of any program developed under TIA would most appropriately be given either to the committee which, after reorganization, had principal responsibility for oversight of that Department or, if involving classified materials, to the two existing intelligence committees.

damages by individuals aggrieved by a violation of the restrictions Congress imposes.[66]

Preventing Mission Creep

Perhaps the greatest source of concern to civil libertarians is that the pattern-recognition technology developed to fight the war on terrorism will be enlisted in the service of other perceived public "goods." As one participant in a recent forum observed, driving with an expired driver's license is a crime and, with enhanced pattern-recognition technology and access to government databases, it is one that might readily be uncovered and prosecuted.[67] Is KD technology to be used for so trivial a purpose? Or, indeed, is it to be used for more substantial purposes—such as to catch a sniper or fight the "war" on drugs?

Though the potential for using the KD tool in these circumstances exists, it *must not* be used in this manner. In every instance where the government intrudes on civil liberty we must ask three distinct questions:

- Is the intrusion an effective means of addressing the threat of concern?
- How significant is the threat? and
- How substantial is the intrusion on individual liberty and privacy?[68]

Evaluating each proposed use necessarily involves a relative ranking and value judgment as to the nature of a threat and the degree of an intrusion. Here, the calculus is necessarily inexact. But our best judgment is that KD is only acceptable as a technological response to the new, significant threat of terrorism at home and abroad.

After September 11, no one can doubt that domestic law enforcement and foreign intelligence agencies face a new challenge that qualitatively poses a greater threat to the American public than any other criminal activity. U.S. foreign counterintelligence efforts are responding to a new and different form of terrorism and espionage. It is appropriate, therefore, that the use of KD to pose pattern-based queries to non-government databases be limited to the exigent circumstances that caused it to be necessary. But because the intrusion is so great, technology being developed to query and correlate data and uncover potential terrorist activity should be used (whether for law enforcement or intelligence purposes) only to investigate terrorist, foreign intelligence, or national security activities, and should never be used for other criminal activity that does not rise to this level of national significance.[69]

Given the bona fide fears of increased government power, any systems that might be derived from KD should be used only for investigations in which there is substantial reason to believe that terrorist-related activity is being perpetrated by organizations whose core purpose is domestic terrorism and the primary purpose of the use of the technology is the discovery of that activity. That limitation should be enacted into law. And to ensure that the enactment is honored, Congress should authorize pattern-based data inquires only if they are certified at a high and responsible level of government to be necessary to accomplish the anti-terrorism objectives of the United States. Possible restrictions might require the

66. One model for enhanced sanctions would be section 224 of the USA PATRIOT Act, Pub. L. No. 107-56, which provided for the administrative discipline of federal officers and employees who violate prohibitions against unauthorized disclosures of information and allowed for civil actions against the United States for damages. With TIA, Congress should go beyond that statute and include provisions for individual civil and, in the case of deliberate misconduct, criminal liability.

67. "Good Cause" and the Timely Collection of Information about Terrorists, Potomac Institute for Policy Studies, June 24, 2003.

68. *See* Paul Rosenzweig, "Principles for Safeguarding Civil Liberties in an Age of Terrorism," Heritage Foundation *Executive Memorandum* No. 854, January 31, 2003.

69. Some have mocked the recent name change of the overall research program from "Total Information Awareness" to "Terrorism Information Awareness," deriding it as an Orwellian name change intended to conceal the truth. They are wrong. Though the announcement of the name change provides, again, an example of DARPA's "tin ear" for public relations, in this instance it illuminates admirably an important reality. By renaming the system to focus on terrorism, DARPA has reminded us of the underlying purpose of the program—to prevent another attack of the scope of that conducted on September 11, or worse. And by emphasizing that focus, the name change enhances (albeit only marginally) the prospects that the use of TIA technology will be limited to the dangers that brought the program into being.

authorization of a Senate-confirmed officer of the Department of Justice, Homeland Security, FBI, or CIA (such as a Assistant Attorney General or the FBI Director).

Beyond the authorization and certification requirement, limitations on "mission creep" can be advanced by housing operational control of the KD system in the appropriate part of government. Placing operational control with the Department of Justice, for example, would be a mistake, because the culture of the Department is such as to foster a desire to use a KD system for other law enforcement purposes. Because the KD technology is being developed primarily to prevent a repeat of terrorism within America, it makes best sense to give operational responsibility for use of KD technology to a new bureau within the Department of Homeland Security or to the Terrorist Threat Integration Center.[70] That structural decision will also help in assuring that use of the technology is limited—as will congressional oversight and continued public scrutiny.

To be sure this task, too, is not an easy one. But if it becomes apparent that the task cannot be accomplished—that is, if it becomes clear that TIA/KD technology cannot be limited to use in the war on terror—then the risks of further development of the technology are too great. One may run the risk of abuse and a diminution of liberty when the threat is especially significant—as the terrorist threat is—but those same risks are not worth running if all that we will accomplish is to empower government in ways that allow it to exercise authority to punish more mundane criminal conduct.

Conclusion

Critics of TIA are wrong to exalt the protection of liberty as an absolute value. That vision rests on an incomplete understanding of why Americans formed a civil society. As John Locke, the seventeenth-century philosopher who greatly influenced the Founding Fathers, wrote: "In all states of created beings, capable of laws, where there is no law there is no freedom. For liberty is to be free from the restraint and violence from others; which cannot be where there is no law; and is not, as we are told, a liberty for every man to do what he lists."[71] Thus, the obligation of the government is a dual one: to protect civil safety and security against violence *and* to preserve civil liberty.

That goal can be achieved. To be sure, it is a difficult task. It is far easier to eschew the effort. But failure to make the effort—failure to recognize that security need not be traded off for liberty in equal measure and that the "balance" between them is not a zero-sum game—is a far greater and more fundamental mistake. Policymakers must respect and defend the individual civil liberties guaranteed in the Constitution when they act, but they also cannot fail to act when we face a serious threat from a foreign enemy.

—Paul Rosenzweig is Senior Legal Research Fellow in the Center for Legal and Judicial Studies at The Heritage Foundation and Adjunct Professor of Law at George Mason University.

70. As noted previously, the KD Council should be co-located with the operational bureau implementing KD technology if it is not housed within the Terrorist Threat Information Center.

71. John Locke, *Two Treatises of* Government, ed. Peter Laslett (Cambridge: Cambridge University Press, 1988), 305; *see also* Thomas Powers, "Can We Be Secure and Free?" *The Public Interest* (Spring 2003) ("In a liberal republic, liberty presupposes security; the point of security is liberty.").