NOT FOR PUBLICATION UNTIL RELEASED BY THE SUBCOMMITTEE

Statement by

Dr. Tony Tether

Director
Defense Advanced Research Projects Agency

Submitted to the

Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census
Committee on Government Reform
United States House of Representatives

May 6, 2003

NOT FOR PUBLICATION UNTIL RELEASED BY THE SUBCOMMITTEE

Mr. Chairman, Subcommittee Members, and staff: I am Tony Tether, Director of the Defense Advanced Research Projects Agency (DARPA). I am pleased to appear before you today to talk about data mining and protecting the privacy of Americans. This is an important issue, and I hope that you will find my remarks helpful as your subcommittee looks into this complicated topic.

Some of you might be unfamiliar with DARPA. We are, essentially, tool makers, sponsoring high-payoff research for the Department of Defense (DoD). This research includes several new software tools that DARPA is developing to assist the DoD in its counterterrorism mission. We are developing new data search and pattern recognition technologies, which have little in common with existing data mining technology, and represent just one element of DARPA's counterterrorism research. Other critical areas of our research include secure collaborative problem solving, structured knowledge discovery, data visualization, and decision making with corporate memory.

It is important to remember that the technologies I will be discussing do not yet exist in their final form, and, no doubt, they will change. Some will succeed and some will fail, and we will learn as we go along. That is the nature of research.

Moreover, unlike some of the other agencies represented by my fellow panelists today, DARPA is not an agency that will actually use these tools, if they work. Other agencies in the DoD, Federal government, or Congress will decide *if* they want to use the tools we create and *how* they will use them.

DARPA's Approach to Data Search and Pattern Recognition

When most people talk about "data mining," they are referring to the use of clever statistical techniques to comb through large amounts of data to discover previously unknown, but useful patterns for building predictive models. This is typically done in the commercial world to better predict customer purchases, understand supply chains, or find fraud – or address any number of other issues where a better understanding of behavior patterns would be helpful. The basic approach is to find statistical correlations as a means of discovering unknown behavior patterns, and then build a predictive model.

At first, one might think that data mining would be very helpful for the most general attempts to find terrorists. It would appear ideal to have software that could automatically discover suspicious, but previously unnoticed patterns in large amounts of data, and which could be used to create models for "connecting-the-dots" and predicting attacks beforehand. However, there are fundamental limitations to expanding today's data mining approaches to the challenge of generally finding and interdicting complex and meticulously well-planned terrorist plots that involve various individuals.

Skeptics believe that such techniques are not feasible because it is simply too difficult to program software to answer the general question, "Is that activity suspicious?" when terrorist plans are so variable and evidence of them is so rare. The results, skeptics say, will contain unmanageable numbers of "false positives" – activities flagged as suspicious that turn out to be innocent.

Beyond the skeptics, critics claim that such an approach must inevitably lead to "fishing expeditions" through massive amounts of personal data and a wholesale invasion of Americans' privacy that yields, basically, nothing in terms of finding terrorists. In previous testimony, this approach has been referred to as "mass dataveillance."

In fact, these objections are among the reasons why DARPA is *not* pursuing these techniques, but is developing a different approach in our research.

DARPA is *not* trying to bring about "mass dataveillance," regardless of what you have read or heard. We believe that the existing data mining approach of discovering previously unknown patterns is ill-suited to ferreting out terrorist plans.

The purpose of data mining is, typically, to find previously unknown but useful patterns of behavior in large amounts of data on activities that are narrowly defined and identified, such as credit card usage or book purchases. These behavior patterns relate to individual transactions or classes of transactions (but not to individuals, themselves), again in narrowly defined and identified areas of activity.

The counter-terrorism problem is much more difficult than this. To detect and prevent complex terrorist plots, one must find *extremely rare* instances of patterns across an *extremely wide* variety of activities – and *hidden* relationships among individuals. Data mining is ill-suited to

this task because the domains of potentially interesting activity are so much more numerous and complex than purchasing behavior.

Accordingly, we believe that better tools and a different approach are needed for the most general efforts to detect and prevent complicated, well-planned terrorist plots, particularly if we are to prevent them well before they can occur and long before they can reach U.S. shores. Consequently, our research goal to create better counterterrorism tools will not be realized by surveilling huge piles of data representing a collection of broad or ill-defined activities in the hope of discovering previously unknown, unspecified patterns. Instead, we are pursuing an approach of searching for *evidence* of specified patterns.

Detecting Data that Fits Specified Patterns

Our approach starts with developing attack scenarios, which are used to find specific patterns that could indicate terrorist plans or planning. These scenarios would be based on expert knowledge from previous terrorist attacks, intelligence analysis, new information about terrorist techniques, and/or from wargames in which clever people imagine ways to attack the United States and its deployed forces. The basic approach does not rely on statistical analysis to discover unknown patterns for creating predictive models. Instead, we start with expert knowledge to create scenarios in support of intelligence analysis versus an data mining approach that scans databases for previously unknown correlations.

The scenarios would then be reduced to a series of questions about which data would provide evidence that such attacks were being planned. We call these scenarios "models," and they are, essentially, hypotheses about terrorist plans. Our goal is to detect data that supports the hypotheses.

Contrast this approach with trying to discover a suspicious pattern without having a model as a starting point – when the pattern is not known in advance. Consider a truck bomb attack, involving a rental truck filled with fertilizer and other materials. Trying to get software to discover such an attack in its planning stages by combing through piles of data – not knowing what it was looking for, but trying to flag "suspicious" activities suggestive of terrorist planning – is unlikely to work. Terrorist activity is far too rare, and spotting it across many different

activities by broadly surveilling all available data requires enormous knowledge about the world in order to identify an activity or individual as being "suspicious."

DARPA's approach, instead, focuses a search on detecting evidence for the scenario model or hypothesis, "Are there foreign visitors to the United States who are staying in urban areas, buying large amounts of fertilizer and renting trucks?" Again, the model or hypothesis is not created by meandering through vast amounts of data to discover unknown patterns.

Finding the evidence of a suspicious pattern is, of course, not as simple as I have made it sound. DARPA's counterterrorism research in the areas of data search and pattern recognition is based on two basic types of queries that, as a practical matter, would probably be used in combination.

The first type of query is subject-based and begins with an entity, such as people *known* to be suspects. Analysts would start with actual suspects' names and see if there is evidence of links with other suspects or suspicious activities. Current technology and policy pertaining to subject-based queries are fairly well developed and understood. One method of subject-based query with enormous potential is link analysis, which seeks to discover knowledge based on the relationships in data about people, places, things, and events. Link analysis makes it possible to understand the relationships between entities. Properly assembled, these links can provide a picture of higher-level terrorist networks and activities, which, in turn, forms a basis for early indications and warning of a terror attack. Data mining offers little as a tool for investigating such relationships – it creates models by finding statistical correlations within databases without using a starting point, and then applies these models indiscriminately over entire data sets. Link analysis differs because it detects connectedness within rare patterns using known starting points, reducing the search space at the outset.

The second type of query is strictly pattern-based. Analysts would look for evidence of a specified pattern of activity that might be a threat.

It is crucial to note that both types of queries start with either known, identified suspects or known, identified patterns. The focus is *investigative* as opposed to broad surveillance. In both cases, the data that one is looking for is likely to be distributed over a large number of very different databases. Querying distributed, heterogeneous databases is not easy, particularly if we are trying to detect patterns, and we do not know how to do it right now. Pattern query

technology is a critical element of our counter-terrorism research; it is rather immature, as are the policies governing its application.

The data that analysts get back in response to a query might not tell them everything. The response may depend on who is doing the analysis and their levels of authorization. This brings me to the second aspect of our approach, detecting in stages.

Detecting in Stages

We envision that analysts will search for evidence of specified patterns in stages. They will ask questions, get some results, and then refine their results by asking more questions. This is really just common sense, but it is worth highlighting that detecting in stages offers a number of advantages: it uses information more efficiently; it helps limit false positives; it can conform to legal investigative procedures; and it allows privacy protection to be built-in.

Detecting in stages helps deal with the crucial challenge of false positives – that is, mistakenly flagging activities and people as suspicious that are, in fact, innocuous. False positives waste investigative resources and, in the worst cases, can lead to false accusations. Unfortunately, much of the discussion of false positives and counter-terrorism has tended to emphasize technology as the key issue by implicitly assuming a caricature of an investigative process in which a computer program fishes through massive piles of data, officials press the "print" button, and out pop a bunch of arrest warrants. Of course, such an approach is unworkable.

We recognize that false positives must be considered as a product of the whole system. They result from how the data, the technology, the personnel, *and* the investigative procedures interact with each other – they are not solely the result of the application of less-than-perfect technology. DARPA's research seeks to provide analysts with powerful tools, not replace the analysts themselves. Moreover, how we react to positives and what we plan to do with the result is what matters enormously to this issue.

It is also important to remember that all investigations – whether they use databases or not – will yield false positives. Therefore, the relevant question is, "Can we improve our overall ability to detect and prevent terrorist attacks without having an unacceptable false positive rate at the system level?" That is the key challenge to be answered by our research.

No doubt many of the "positives" found during the first queries that analysts make will be false ones. The positives must be further examined to start weeding out the false ones and confirming the real ones, if there are any. This will require analysis in several stages to find independent, additional evidence that either refutes or continues to support the hypothesis represented by the model. Moreover, the level of proof depends, in part, on the nature of the planned response to a positive. We do not, for example, arrest everyone who sets off the metal detector when entering this building.

An analogy we sometimes use to illustrate this is submarine detection. In submarine warfare, we do not simply attack something based on first indications that a single sensor has detected an object. We refine the object's identification in stages – from "possible" enemy submarine, to "probable" enemy submarine, to "certainly" an enemy submarine. To be sure of our actions, we confirm the identification over time, using different, independent sensors and sources of information. Our approach to data searching and pattern recognition would proceed in a similar fashion.

Proceeding in stages also means that the entire process can conform to required, legal procedures or steps. In fact, many of these steps exist *precisely* to protect people's rights and weed out false positives. We envision hard-wiring many of the required procedures, permissions, or business rules into the software to ensure that they are actually being followed at each stage of the process.

Let us go back to the truck bomb example. One might incorporate a process called "selective revelation" into data queries. In selective revelation, the amount of information revealed to the analyst depends on who the analyst is, the status of the investigation, and the specific authorization the analyst has received. The analyst's credentials would be automatically included with the query, and the level of information returned would vary accordingly.

Perhaps the result of the truck bomb query I talked about earlier is that 17 people fit the truck bomber pattern, but no personal information about those 17 is revealed. To retrieve additional personal information, a higher level of authorization might be required, based on an independent evaluation (by a court, for example) of the evidence that the analyst is actually "on to" something suspicious.

This suggests that there is a special class of business rules and procedures that could be put into the technology to strengthen privacy protection, so let me turn to that now.

Built-in Privacy Protection

From the very start of our research, we began looking for ways to build privacy protection into DARPA's approach to detecting terrorists.

We had two motivations. First, we knew that the American public and their elected officials must have confidence that their liberties will not be violated before they would accept this kind of technology.

Second, much of what Federal agencies need to share is *intelligence* data. Historically, agencies have been reluctant to share intelligence data for fear of exposing their sources and methods. Accordingly, protecting privacy and intelligence sources and methods are integral to our approach.

We are putting policies into place that will highlight protecting privacy. As I previously alluded, DARPA does not own or collect any intelligence or law enforcement databases. Our policies will address the development and transition of new tools to the agencies authorized by law to use those databases, reinforcing to everyone the importance of privacy. Moreover, we are fully aware of and intend for the tools to be only used in a manner that complies with the requirements of the Privacy Act, as well as the privacy provisions of the E-Government Act regarding a Privacy Impact Assessment where such an assessment is required. And we recognize that under Office of Management and Budget policy, major agency information systems employing the technology will have to be justified by a business case that addresses how privacy and security are built into the technology.

To further assist agencies that have collected the data for analysis, we are developing other tools that will help them protect the integrity of the information – even during searches. I previously mentioned "selective revelation" as one way to protect privacy, and we are looking at other related techniques as well, such as separating identity information from transaction information. These separate pieces of information could only be reassembled after the analyst has received the proper authorizations.

Until then, an analyst might only know the basic facts but not the identity of who was involved. We are also looking at ways to anonymize data before it is analyzed. We are evaluating methods for filtering out irrelevant information from the analysis, such as the use of "software agents" that utilize experience-based rules. These software agents would automatically remove data that appears to be irrelevant before the analyst even sees it.

Going beyond privacy protection, we are also looking into building-in indelible audit technology that makes it exceedingly difficult to abuse the data search and pattern recognition technology without the abuse being detected. This audit technology would answer the question, "Who used the system to retrieve what data?"

Some ideas that we are pursuing include cryptographically protecting audit information and perhaps even broadcasting it to outside parties, where it cannot be tampered with. We are also looking into software agents that would watch what analysts are doing to ensure that their searches and procedures are appropriate and that they are following established guidelines.

Another interesting idea is data that reports its location back to the system. One might even include a unique identifier for each copy ("digital watermark"), so that if unauthorized copies were distributed their source could be traced. Still another concept is giving control of database querying a trusted third party, who could not be subject to organizational pressure to provide unauthorized access.

We take privacy issues very seriously. DARPA is, in fact, one of the few Federal agencies sponsoring significant research in the area of privacy protection technologies.

You will often hear talk in this debate about how there are trade-offs – for instance, that we may need to trade less privacy for more security. People may disagree about the proper balance, but DARPA's efforts in developing privacy protection technology are designed, in fact, to improve prospects for providing both improved privacy protection and improved security by the legally relevant agencies

In closing, I would like to emphasize two points:

First, remember that what I have been describing here today is research, and exactly how the technology will work – indeed, *if* it works – will only be shown over time.

Second, because of the high profile of DARPA's research in this area, in February 2003 the Department of Defense announced the establishment of two boards to provide oversight of our Information Awareness programs, including our data search and pattern recognition technologies. These two boards, an internal oversight board and an outside advisory committee, will work with DARPA as we proceed with our research to ensure full compliance with U.S. constitutional law, U.S. statutory law, and American values related to privacy.

This concludes my remarks. I would be happy to answer any questions.