CAN THE USE OF FACTUAL DATA ANALYSIS STRENGTHEN NATIONAL SECURITY? PART ONE

HEARING

BEFORE THE

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS

OF THE

COMMITTEE ON GOVERNMENT REFORM

HOUSE OF REPRESENTATIVES

ONE HUNDRED EIGHTH CONGRESS

FIRST SESSION

MAY 6, 2003

Serial No. 108–72

Printed for the use of the Committee on Government Reform

Available via the World Wide Web: http://www.gpo.gov/congress/house
http://www.house.gov/reform

U.S. GOVERNMENT PRINTING OFFICE

90–399 PDF WASHINGTON : 2004

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON GOVERNMENT REFORM

TOM DAVIS, Virginia, *Chairman*

DAN BURTON, Indiana
CHRISTOPHER SHAYS, Connecticut
ILEANA ROS-LEHTINEN, Florida
JOHN M. MCHUGH, New York
JOHN L. MICA, Florida
MARK E. SOUDER, Indiana
STEVEN C. LATOURETTE, Ohio
DOUG OSE, California
RON LEWIS, Kentucky
JO ANN DAVIS, Virginia
TODD RUSSELL PLATTS, Pennsylvania
CHRIS CANNON, Utah
ADAM H. PUTNAM, Florida
EDWARD L. SCHROCK, Virginia
JOHN J. DUNCAN, JR., Tennessee
JOHN SULLIVAN, Oklahoma
NATHAN DEAL, Georgia
CANDICE S. MILLER, Michigan
TIM MURPHY, Pennsylvania
MICHAEL R. TURNER, Ohio
JOHN R. CARTER, Texas
WILLIAM J. JANKLOW, South Dakota
MARSHA BLACKBURN, Tennessee

HENRY A. WAXMAN, California
TOM LANTOS, California
MAJOR R. OWENS, New York
EDOLPHUS TOWNS, New York
PAUL E. KANJORSKI, Pennsylvania
CAROLYN B. MALONEY, New York
ELIJAH E. CUMMINGS, Maryland
DENNIS J. KUCINICH, Ohio
DANNY K. DAVIS, Illinois
JOHN F. TIERNEY, Massachusetts
WM. LACY CLAY, Missouri
DIANE E. WATSON, California
STEPHEN F. LYNCH, Massachusetts
CHRIS VAN HOLLEN, Maryland
LINDA T. SANCHEZ, California
C.A. "DUTCH" RUPPERSBERGER, Maryland
ELEANOR HOLMES NORTON, District of Columbia
JIM COOPER, Tennessee
CHRIS BELL, Texas

BERNARD SANDERS, Vermont (Independent)

PETER SIRH, *Staff Director*
MELISSA WOJCIAK, *Deputy Staff Director*
ROB BORDEN, *Parliamentarian*
TERESA AUSTIN, *Chief Clerk*
PHILIP M. SCHILIRO, *Minority Staff Director*

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS

ADAM H. PUTNAM, Florida, *Chairman*

CANDICE S. MILLER, Michigan
DOUG OSE, California
TIM MURPHY, Pennsylvania
MICHAEL R. TURNER, Ohio

WM. LACY CLAY, Missouri
DIANE E. WATSON, California
STEPHEN F. LYNCH, Massachusetts

EX OFFICIO

TOM DAVIS, Virginia

HENRY A. WAXMAN, California

BOB DIX, *Staff Director*
SCOTT KLEIN, *Professional Staff Member*
URSULA WOJCIECHOWSKI, *Clerk*
DAVID MCMILLEN, *Minority Professional Staff Member*

CONTENTS

CAN THE USE OF FACTUAL DATA ANALYSIS STRENGTHEN NATIONAL SECURITY? PART ONE

TUESDAY, MAY 6, 2003

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY,
INTERGOVERNMENTAL RELATIONS AND THE CENSUS,
COMMITTEE ON GOVERNMENT REFORM,
Washington, DC.

The subcommittee met, pursuant to notice, at 3 p.m., in room 2154, Rayburn House Office Building, Hon. Adam Putnam (chairman of the subcommittee) presiding.

Present: Representatives Putnam, Miller and Clay.

Staff present: Bob Dix, staff director; John Hambel, senior counsel; Scott Klein, Chip Walker, Lori Martin, and Casey Welch, professional staff members; Ursula Wojchechowski, clerk; Suzanne Lightman, fellow; David McMillen, minority professional staff member; and Jean Gosa, minority assistant clerk.

Mr. PUTNAM. A quorum being present. The hearing of the Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census will come to order. Good afternoon and welcome to today's hearing, "Can the Use of Factual Data Analysis Strengthen National Security? Part One." First of all, I'd like to thank everyone for bearing with us as we've had to make time and room changes for today's hearing. We appreciate your cooperation.

In an effort to prevent future terrorist attacks and enhance law enforcement efforts, deputies and agencies throughout the Federal Government have begun developing strategies that will assist in the identification of potential risks through the use of technology and information sharing. The truth is that there is a tremendous amount of information that already resides in the public venue. However, due to past practices of stovepipe mentalities and turf issues, much relevant information that could be of use or interest to law enforcement officials has not been easily accessible.

In particular, since September 11, 2001, it has been imminently clear that we must do a better job of compiling and sharing information that will provide, enhance the opportunities for law enforcement and national security officials to identify potential risks in advance. Federal agencies have utilized methodologies that facilitate data base exploration for quite some time in an effort to root out waste, fraud and abuse. In fact the recent highly public case of government credit card abuse was flushed out, and the perpetra-

tors identified through the use of data mining or factual data analysis, as some call it.

Now, a number of Federal agencies with the responsibility for homeland security and law enforcement are employing the lessons learned through the use of factual data analysis or the conclusions drawn from this analytical process to increase their ability to detect patterns and relationships within the masses of data they have access to in an effort to increase risk assessment capabilities. This hearing will examine whether the use of this process will successfully enhance efforts to strengthen law enforcement and national security.

Does factual data analysis contribute to increase the risk detection? As we have previously established, factual data analysis is not a technology in and of itself. It is an analytical process that utilizes technology in an effort to identify patterns and relationships that were previously unknown. It has been used successfully in the private sector to craft specific marketing and sales programs. It has been used successfully in the public sector to identify and address instances of waste, fraud and abuse.

The hope is that these same technological advances that aid marketers in identifying customers for their products and law enforcement in catching tax evaders or identifying welfare fraud will also detect patterns that should raise suspicion among those working would improve our Nation's security. Today we have witnesses representing the FBI, the Transportation Security Administration, and Defense Advanced Research Projects Agency.

Each of these three agencies proposes to use factual data analysis or conclusions drawn from the process to enhance homeland security. Specifically, we will be examining the FBI's Trilogy and related technology analysis tools, TSA's computer-assisted prescreening process system [CAPPS] II, and DARPA's total information awareness [TIA]. We have asked each of these witnesses to explain their agency's program and talk about the role factual data analysis is envisioned to play. While each of these agencies proposals is different in its construct and each may generate varying responses and levels of interest, the subcommittee will seek to learn more about the source, accuracy, reliability, and security of the data that is accessed to determine risk assessment.

Let me be clear, we are not here to compare one project to the other, nor are we here to evaluate the strategic basis for these projects. We are here to examine the use of technology in the facilitation of this process and the techniques, processes and outcomes that are produced. We hope to listen and learn from these expert witnesses and hear factual information about these projects. We also recognize that there is clearly some concern and reluctance on the part of some of the witnesses to even be here today because of some of the press coverage about these projects. Today we will examine the facts about how data will be compiled, what data will be assembled, what steps are taken to ensure the accuracy and reliability of the data, how the data will be analyzed, and what will be done with the results as well as how the privacy and personal freedom of the public will be protected by the process itself. We expect full and complete disclosure.

In 2 weeks after gathering and evaluating the information that will be presented today, the subcommittee will reconvene and examine this issue from a standpoint of privacy and personal freedom concerns in part 2 of this hearing. The subcommittee believes this is a good place to start from. From an oversight perspective, we look forward to working with these agencies as they continue to plan and implement their proposals for enhancing homeland security.

[The prepared statement of Hon. Adam H. Putnam follows:]

COMMITTEE ON GOVERNMENT REFORM

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY, INTERGOVERNMENTAL RELATIONS AND THE CENSUS

CONGRESSMAN ADAM PUTNAM, CHAIRMAN

OVERSIGHT HEARING
STATEMENT BY ADAM PUTNAM, CHAIRMAN

Hearing topic: *"Can the Use of Factual Data Analysis Strengthen National Security? -Part One"*

Tuesday, May 6, 2003
3:00 p.m.
Room 2154 Rayburn House Office Building

OPENING STATEMENT

In an effort to prevent future terrorist attacks and enhance law enforcement efforts, departments and agencies throughout the federal government have begun developing strategies that will assist in the identification of potential risks through the use of technology and information sharing. The truth is that there is a tremendous amount of information that already resides in the public venue. However, due to past practices of stovepipe mentalities and turf issues, much relevant information that could be of use or interest to law enforcement officials has not been easily accessible. In particular since September 11, 2001, it has been imminently clear that we must do a better job of compiling and sharing information that will provide enhanced opportunities for law enforcement and national security officials to identify potential risks...in advance. Federal agencies have utilized methodologies that facilitate database exploration for quite some time in an effort to root out waste, fraud and abuse. In fact, the recent highly public case of government credit card abuse was flushed out, and the perpetrators identified, through the use of "data mining" or "factual data analysis" as some would prefer it was called.

Now, a number of Federal agencies with responsibility for homeland security and law enforcement, are employing the lessons learned through the use of factual data analysis, or the conclusions drawn from this analytical process, to increase their ability to detect patterns and relationships within the masses of data they have access to in an effort to increase risk assessment capabilities. This hearing will examine whether the use of this process will successfully enhance efforts to strengthen law enforcement and national security. Does factual data analysis contribute to increased risk detection?

As we have previously established, factual data analysis is not a technology in and of itself. It is an analytical process that utilizes technology in an effort to identify patterns and relationships that were previously unknown. It has been used successfully in the private sector to craft specific marketing and sales programs. It

has been used successfully in the public sector to identify and address instances of waste, fraud, and abuse. The hope is that these same technological advances that aid marketers in identifying customers for their products and law enforcement in catching tax evaders or identifying welfare fraud, will also detect patterns that should raise suspicion among those working to improve our nation's security.

Today we have witnesses representing the Federal Bureau of Investigation (FBI), the Transportation Security Administration (TSA) and Defense Advanced Research Projects Agency (DARPA).

Each of these three agencies proposes to use factual data analysis or conclusions drawn from the process to enhance homeland security. Specifically we will be examining the FBI's Trilogy and related technology analysis tools, TSA's Computer Assisted Prescreening Process System (CAPPS II) and DARPA's Total Information Awareness (TIA). We have asked each of these witnesses to explain their agency's program and talk about the role factual data analysis is envisioned to play.

While each of these agencies proposals is different in its construct, and each may generate varying responses and levels of interest, the Subcommittee will seek to learn more about the source, accuracy, reliability, and security of the data that is accessed to determine risk assessment. Let me be clear...we are not here to compare one project to the other. We are not here to evaluate the strategic basis for these projects...we are here to examine the use of technology in the facilitation of this analytical process and the techniques, processes, and outcomes that are produced. We are here to listen and learn from these expert witnesses and to hear factual information about these projects. We also recognize that there is some concern and reluctance on the part of some of these witnesses to even be here today because of some of the press coverage about these projects. Today, we will examine the facts about how data will be compiled; what data will be assembled; what steps are taken to insure the accuracy and reliability of the data; how the data will be analyzed; and what will be done with the results, as well as how the privacy and personal freedom of the public will be protected by the process itself.

In two weeks, after gathering and evaluating the information that will be presented today, the Subcommittee will reconvene and examine this issue from a standpoint of privacy and personal freedom concerns in Part II of this hearing.

The Subcommittee believes this a good place to start from. From an oversight perspective, we look forward to working with these agencies as they continue to plan and implement their proposals for enhancing homeland security.

Mr. PUTNAM. It's my pleasure now to yield to the gentleman from Missouri, the ranking member, Mr. Clay, for any opening remarks that he may have.

Mr. CLAY. Thank you, Mr. Chairman. And thank you for calling this hearing. I look forward to today's testimony. When government agencies collect information about American citizens, then those citizens have a right to see that information and correct it if there are errors. I would like to begin by quoting one of our witnesses from our last hearing on data mining, Professor Jeffrey Rosen. At that hearing, Professor Rosen opened his testimony with this statement: "It's possible to design data mining technologies in ways that strike better or worse balances between liberty and security. But there is no guarantee that the executive branch or the technology just left to their own devices will demand and provide technologies that strike the balance in a reasonable way. Congress, therefore, has a special responsibility to provide technological and legal oversight of data mining to ensure that the most invasive searches are focused on the most serious crimes."

Our job today is to gather as much information as possible about these three programs so that we can assure that the balance between liberty and security is a good one. Over the past 2 years, we have seen a heavy thumb on the balance scale in favor of security. However, it is not clear that we are necessarily more secure because of it. We have also seen the liberty of individuals abused in ways we have not seen in this country since the internment of the Japanese during World War II. We learned in hindsight that breach of liberty was a terrible misuse of government power. The government quietly admitted so when it turned to those people interned and asked them to serve in the military or to work as translators.

Much later our government officially apologized. President Clinton, in issuing that policy said, "we recognize the wrongs of the past and offer our profound regret to those who endured such grave injustice. We understand that our Nation's actions were rooted in racial prejudice and wartime hysteria. And we must learn from the past and dedicate ourselves as a Nation to renewing and strengthening equality, justice and freedom."

Today our government faces a threat to our national security that many have compared to World War II. President Bush compared the attack on the World Trade Center to the bombing of Pearl Harbor. In the days that followed that attack, the President's speech writers used President Roosevelt's speeches from December 1941 to shape President Bush's remarks. We must learn from the past and not allow our fears to destroy the very liberties for which we fight. The descriptions of the programs we are considering today with secret filings and warrantless searches of our electronic lives puncture that thin wall between liberty and security. At the same time, these programs have not proved that they have a benefit strong enough to justify that breach.

Finally, I'd like to thank the Defense Advance Research Project Agency for providing this testimony to the subcommittee in a timely fashion. It's a shame that the FBI and TSA did not show the same respect for this subcommittee. Again, let me thank the wit-

nesses for their testimony. And I ask that my statement be included in the record. Thank you, Mr. Chairman.

Mr. PUTNAM. You're very welcome. Thank you, Mr. Clay. We appreciate your interest.

[The prepared statement of Hon. Wm. Lacy Clay follows:]

1

STATEMENT OF THE HONORABLE WM. LACY CLAY
AT THE HEARING ON
GOVERNMENT USES OF DATA-MINING

MAY 6, 2003

Thank you, Mr. Chairman for calling this hearing. I look forward to today's testimony. When government agencies collect information about American citizens, then those citizens have a right to see that information and correct it if there are errors.

I would like to begin by quoting one of our witnesses from our last hearing on data mining, Professor Jeffery Rosen. At that hearing, Professor Rosen opened his testimony with this statement:

> "it's possible to design data mining technologies in ways that strike better or worse balances between liberty and security. But there is no guarantee that the executive branch or the technologists, left to their own devices, will demand and provide technologies that strike the balance in a reasonable way. Congress, therefore, has a special responsibility to provide

technological and legal oversight of data mining, to ensure that the most invasive searches are focused on the most serious crimes."

Our job today is to gather as much information as possible about these three programs so that we can assure that the balance between liberty and security is a good one. Over the past two years we have seen a heavy thumb on the balance scale in favor of security. However, it is not clear that we are necessarily more secure because of it. We have also seen the liberty of individuals abuse in ways we have not seen in this country since the internment of the Japanese during World War II.

We learned in hindsight that that breach of liberty was a terrible misuse of government power. The government quietly admitted so when it turned to those people interned and ask them to serve in the military, or to work as translators. Much later, our government officially apologized. President Clinton, in issuing that apology said, "We recognize the wrongs of the past and offer our profound regret to those who endured such

grave injustice. We understand that our nation's actions were rooted in racial prejudice and wartime hysteria, and we must learn from the past and dedicate ourselves as a nation to renewing and strengthening equality, justice and freedom."

Today our government faces a threat to our national security that many have compared to World War II. President Bush compared the attack on the World Trade Center to the bombing of Pearl Harbor. In the days that followed that attack, the president's speechwriters used President Roosevelt's speeches from December 1941 to shape President Bush's remarks.

We must learn from the past and not allow our fears to destroy the very liberties for which we fight. The descriptions of the programs we are considering today with secret files, and warrantless searches of our electronic lives, puncture that thin wall between liberty and security. At the same time, these programs have not proved that they have a benefit strong enough to justify that breach.

Finally, I would like to thank the Defense Advanced Research Project Agency for providing its testimony to the

Subcommittee in a timely fashion. It is a shame that the FBI and TSA did not show the same respect for the Subcommittee.

Again, let me thank the witnesses for their testimony, and I ask that my statement be included in the record.

Mr. PUTNAM. And obviously this is going to be an interesting hearing. At this time, I'd like to recognize the vice chair of the subcommittee, the gentlelady from Michigan, Mrs. Miller.

Mrs. MILLER. Thank you, Mr. Chairman. Just a brief opening statement if I may. I look forward to hearing the testimony of all the witnesses. I'm very appreciative of all of you coming here today. I think this will be a fascinating hearing. I think the issue of factual data analysis as a tool to strengthen national security is certainly one of the most significant issues facing our Nation. In fact, today in our society. And with the implementation of the E-government Act of 2002 and the growing importance of information technology in our world establishing investigative techniques such as factual data analysis are vital, absolutely vital if our Nation is to successfully prosecute the war on terror.

As we know, the terrorists seem to have an uncanny ability to adapt to our methods of prevention. In many instances, they are using our freedoms against us. In my view, I think we need to focus all of our resources and attention to ensure we are always at least one step ahead of them if possible. Federal officials currently have at their disposal the resources and knowledge to implement systems that assist us in this process. These officials currently occupy the vanguard of our defenses and need not to necessarily be hampered by bureaucracy in their efforts.

However, the American people must have confidence that the Federal Government is using this new source of information in a very ethical and proper and effective way. It's absolutely essential that a proper balance be made between the operation of the government as it prosecutes the war on terror and the disclosure of operations to citizens that it was set up to protect. And for this reason, it's the responsibility of this subcommittee to ensure that the factual data analysis as a tool be not abused.

For example, I'm certainly very encouraged that the Total Information Awareness Project [TIA], is being conducted by the Defense Advanced Research Projects Agency [DARPA], and it has been relatively transparent. Why only in its beginning phases this program provides a hope that we can analyze the patterns of a terrorist to anticipate their next move? And some have expressed concern about programs such as these. But the mere fact that Mr. McCraw, Admiral Loy, and Dr. Tether have agreed to testify certainly shows that the Federal Government is concerned about the perception that Congress and the public has about data analysis as well.

I'm very much looking forward to working with the chairman and members of the subcommittee and full committee to ensure that this program receives the proper congressional oversight and I certainly will be interested to hear the testimony provided today. Thank you, Mr. Chairman.

Mr. PUTNAM. You're very welcome, Mrs. Miller. We thank you.

With that, we will move to the witnesses. You're all experienced with congressional testimony. You understand the light system. We'll ask you to adhere to the timing out of respect for everyone's schedules. Today each witness will testify on his own panel. After each witness has given the 5-minute statement, the subcommittee will ask questions particular to that witness's agency. After all three panels have testified and answered this initial round of ques-

tioning, the three witnesses will return to the witness table to answer future rounds of questions.

As you are aware, we swear in our witnesses. So if Mr. McCraw would please rise for the swearing in.

[Witness sworn.]

Mr. PUTNAM. Note for the record the witness has responded in the affirmative. I will ask if there are associates from your agency who intend to provide supporting evidence or testimony for the subcommittee, that you rise and be sworn in also at the appropriate time. Our first witness today is Steven C. McCraw, a 20-year FBI veteran. He's assisted this year to the newly created Office of Intelligence. His office will be responsible for implementing FBI intelligence strategies, making sure that intelligence is properly collected, managed and shared within the FBI, with State and local law enforcement through the 66 Joint Terrorism Task Forces, and with the intelligence community, including the new Terrorist Threat Integration Center.

Previous to his current appointment, he was special agent in charge of the FBI San Antonio field office and served as the director of the Foreign Terrorist Tracking Task Force before that. We're pleased to have you and you're recognized for your statement.

STATEMENT OF STEVE McCRAW, ASSISTANT DIRECTOR, OFFICE OF INTELLIGENCE, FEDERAL BUREAU OF INVESTIGATION, ACCOMPANIED BY WILLIAM HOOTEN, DEPUTY EXECUTIVE ASSISTANT DIRECTOR

Mr. MCCRAW. Thank you, Mr. Chairman. First, I owe yourself and the members of the subcommittee an apology because you didn't have my statement well in advance. I have been notified that it has been cleared. I ask your permission that we do submit it for the record.

Mr. PUTNAM. How quickly can you get it to the subcommittee using all the miracles of technology?

Mr. MCCRAW. Well, I have to turn around here, Mr. Chairman, with your permission. Making copies and driving it here right now at this time.

Mr. PUTNAM. I'm quite certain we have a fax machine. If they want to get it to us that way, we'll be able to afford the audience and others the opportunity to review it as well in a timely manner. We look forward to that. Thank you. You're recognized.

Mr. MCCRAW. Thank you. First, I'd like to take the opportunity to thank each and every one of you for your support in enabling the FBI to modernize its information technology systems. I think, in fact, in all the statements that you made, you know, previously including the letter asking the FBI's participation in this important hearing, I noted the importance and value added benefits of utilizing technology. Clearly, the FBI's focus is trying to utilize these advances to manage and to find links, relationships, and patterns of individuals within its own data systems. To that end, it's the information in terms that the FBI legally and lawfully collects in the course of its investigations that becomes a part of its system of records.

As background in terms of discussion of data mining, I believe it's important to understand the term "data mining" as it's used

commonly today. It's defined as technology that facilitates the ability to sort through masses of amounts of information through data base explorations, extract specific information in accordance with defined criteria, and then identify patterns of interest to users.

Also, as I mentioned before, it's an outstanding tool to be able to go through those data sets and identify links, relationships, and associations between individuals of interest. And in effect, what it does is automate what analysts and agents have had to do for years. So what it does, it allows analysts and agents to work cheaper, faster, smarter. And that is how the FBI, in terms of its Trilogy, is going about it.

Now, one thing that has been critically important to the FBI and we have a strong commitment to and that is the rule of law, the Constitution, the statutes that you in Congress have passed, the Attorney General guidelines, the Privacy Act, and all of the laws and statutes that clearly delineate and guidelines for the FBI in terms of how they can properly and lawfully collect information. Because in effect, that's the information that we would be utilizing this technology on, its own internal information.

One of the advantages, the FBI, from lessons learned over the years, is that we have, and the reasons we have what the Office of General Counsel has, an Administrative Law Unit, we have an Investigative Law Unit, the reason we have in the field the Chief Division Counsels so that these rules and regulations that are closely adhered to and followed up on is to ensure that an agent doesn't go out on an fishing expedition in terms of looking at an individual, arbitrarily looks at a person or surveils a person, but there's predication, a reason for doing it. It's the same thing that we're talking about in terms of data mining. There's a reason, first, to collect the information. Once we collect it lawfully, then naturally we want to exploit the latest technology so that we can work better, in protecting Americans from terrorism, from crime, and from foreign intelligence activities.

Ensuring the appropriate controls to protect the privacy of the Federal Government data, we must also look at in terms of public source data. And that is data that is derived and sold by public companies that people have access to, you and I, certainly the private industry utilizes this, and clearly the FBI does utilize public source data as a tool, and clearly as a tool for leads.

Again, so we can work more economically, we can save time, and we can be more efficient in what we're doing in terms of investigations. Now, we've learned from lessons learned, and as you well know, is that public source data is not always accurate. In fact, many times there are errors. So we have to be mindful it's a tool that requires followup investigation.

I'm sure, there have been instances were they've come across erroneous information. I'm applying for a mortgage at this time—and they've identified inaccurate information. In so doing, it makes you mindful of how much other information was linked to a particular credit card inappropriately. Well, you know, the systems aren't perfect. They're run by people. That's why it's only a tool for FBI. And that when we use public source data that we extract the relevant components of it before we bring it into the FBI system of records. We don't, and we won't, go out and purchase wholesale data sets

that are publicly available and incorporate that with the names of myself, my family, you and hundreds of thousands of other Americans in our system of records for convenience sake.

Clearly we have an obligation to be mindful of those things. I look forward to any questions that you might have later. Thank you very much for your time.

Mr. PUTNAM. Thank you, Mr. McCraw. I'm informed that we are either electronically or by fax receiving your testimony so we appreciate that.

Mr. MCCRAW. We don't have the greatest success rate in technology, Mr. Chairman. I apologize for that. We're trying to get better with it.

Mr. PUTNAM. We're going to try to help you.

Mr. MCCRAW. Thank you, sir.

[The prepared statement of Mr. McCraw follows:]

STEVEN C. MCCRAW

ASSISTANT DIRECTOR

OFFICE OF INTELLIGENCE

FEDERAL BUREAU OF INVESTIGATION

HOUSE GOVERNMENT COMMITTEE

SUBCOMMITTEE ON TECHNOLOGY

INFORMATION POLICY, INTERGOVERNMENTAL

RELATIONS AND THE CENSUS

"DATA MINING: PROTECTING THE HOMELAND,

SAFEGUARDING AMERICAN VALUES"

MAY 6, 2003

Good morning Mr. Chairman and Members of the Subcommittee:

My name is Steve McCraw, and I am the Assistant Director of the FBI's Office of Intelligence. I am pleased to have an opportunity to appear before you today to discuss the FBI's use of "data mining" and the safeguards it has in place to protect the privacy and personal information of American citizens.

First, I would like to take this opportunity to thank you and the Members of the Subcommittee for the support you have provided to the FBI in modernizing our information technology infrastructure. The FBI for too long has not been able to capitalize on the tremendous advances in information technology to quickly locate essential elements of information and identify previously unknown links, relationships, and associations hidden within the vast amount of data collected by the FBI in the performance of its investigative responsibilities.

Second, I would like to thank the Subcommittee for this opportunity to address the issue of "data mining" and personal privacy. As witnessed both here in Congress and in the press, it is clearly an issue in the hearts and minds of the American people. The FBI is proud of its commitment to privacy concerns and I am happy to discuss the efforts made in this regard.

"Data Mining"

As background to this issue, I believe it is important to understand the term "data mining," as it is commonly used today. "Data mining" is defined as technology that facilitates the ability to sort through masses of information through databases exploration, extract specific information in accordance with defined criteria, and then identify patterns of interest to its user. In general it refers to the

ability to work with larger amounts of data, at faster speeds, in ways that were previously not possible computationally due to size or speed limitations. The private sector often uses data mining to make sense of the wide breadth of data that companies and industries have available. For example, data mining is often used by industry to analyze potential goods and services that are in demand. Companies that use data mining shorten response time to market changes, which allows for better alignment of their products with their customer's needs. Both government and industry have also successfully used data mining to identify and protect against fraud. A key principle of the FBI's information technology business plan calls for the use of effective industry practices and off the shelf private sector technology to allow the FBI to work more efficiently, more effectively, and more economically. Further, prior to procuring these technologies we carefully review the purchase to ensure they meet privacy laws, policies and regulations.

In recent debates, however, people have begun to use the term "data mining" as a shorthand reference to the specter of abusive searches through vast amounts of publicly available data on innocent private citizens. As your letter requesting FBI attendance at this hearing astutely recognized, these are not the same things. The term "data mining" should not be viewed as always connoting any such abuse.

Uses of "Data Mining"

The United States Constitution and the United States Congress, through legislation, have carefully delineated acceptable conduct in law enforcement investigations and intelligence activities. The FBI has an unwavering commitment to adhere to those requirements, as well as those mandated by Federal regulations and the Attorney General guidelines.

Whether the work is performed manually or in an automated fashion, the commitment does not change.

I'd like to provide some examples of the ways in which the FBI uses, or plans to use, "data mining":

- The FBI's Integrated Automated Fingerprint Identification System (commonly known as IAFIS) has long been available to the law enforcement community to permit verification of arrestee identity. When necessary, IAFIS will compare the fingerprints of a non-cooperating arrestee against millions of known fingerprints in order to provide law enforcement with the individual's identity. It also conducts pattern searches to link on unknown latent fingerprints taken from a crime scene with known individuals.
- The new SCOPE/Integrated Data Warehouse project will allow an agent or analyst to search within the FBI's existing datasets (that is, information collected in lawful investigations and stored in computer format) for links, associations, and relationships among the individuals.
- The Information Sharing Initiative begun in St. Louis will permit sharing of state, local, and Federal data to enable officers to quickly search through multiagency investigative data to identify links between subjects of terrorism and criminal investigations.

Risk Assessment Technology

The Subcommittee has specifically inquired about the use of risk assessment technology. To the FBI, this term connotes the use of technology to identify individuals who present a particular risk, for

example, a risk to the national security. As a general rule, the FBI until recently has not been able to participate in the use of such technology because of inadequate information technology capabilities.

The FBI intends to use the advances in risk technology to identify FBI employees whose activities include a pattern of possible misuse of the FBI's computer systems to access information. The use of this technology and increased audit functions will substantially increase FBI internal security. This ability would have assisted in earlier detection of convicted spy, former FBI Special Agent Robert Hanssen.

The FBI will also leverage risk assessment technology to increase its oversight of human sources. As part of a reengineering project to expand the FBI's Human Intelligence base while providing greater oversight, the FBI will utilize information technology to identify potential problems in the operation of human sources.

Ensuring the appropriate controls to protect the privacy of Federal government data, public source data that contains personal information on American citizens, as well as a combination of the two remains a top priority. Certainly, the FBI uses information collected by public source companies to obtain information on individuals during the course of its terrorist, criminal, and foreign intelligence investigative activities. Often, these systems provide leads that enable the FBI to save valuable investigative and analytical time and resources. Again, acceptable conduct in the collection and use of this information has been clearly defined by statute and Attorney General guidelines. Although public source data is a useful tool, the FBI is well aware that the data is sometimes outdated and/or inaccurate and follow-up investigation to confirm the information is essential.

However, the FBI neither has in the past nor intends in the future to purchase personal information on American citizens who are not part of an ongoing investigation from public source companies and placed it within the FBI's system of records. The FBI has no interest in gathering data on law abiding citizens. Such information is not required to protect the nation.

In closing, I want to thank you for the opportunity to testify before you today and I look forward to any questions you may have for me.

Mr. PUTNAM. I'd like to recognize the gentleman from Missouri to begin his round of questions. Mr. Clay, you're recognized.

Mr. CLAY. Thank you, Mr. Chairman.

Mr. McCraw, a few weeks ago, the FBI issued a final rule that exempted information held by the FBI like the information in Trilogy from the Privacy Act requirements that the information held on individuals be accurate and timely. In other words, the FBI is going to make no effort to assure that the information they hold on an individual is correct. Now, some of these records are available to local law enforcement officials. When someone is stopped for a traffic violation, the officer runs that person's identity through a number of systems, one of the systems checked is the FBI's National Crime Information Center. Under this new rule, the FBI no longer is obligated to assure that the information held by the National Crime Information Center is correct. That means that people will be detained and arrested based on inaccurate information when all they have done is roll through a stop sign. Will you explain to this committee why the FBI believes it is no longer necessary to verify the accuracy of the information it holds on individuals?

Mr. MCCRAW. Mr. Clay, one reason without discussing the statute for years and maintaining these high standards in terms of what data is entered into the NCIC is for that very specific reason that you gave, is that we have police officers out there, they stop for a traffic ticket, run the name and they have been advised all of a sudden this person is wanted. In accordance with officer safety and established guidelines, they have this person arrested. Clearly, there has to be, you know, strict guidelines when you use NCIC, and in fact, it has to not be just for the FBI, but all users of NCIC that have access and enter records into it. Those requirements are, and it's been held in place and there will be no shifting of those requirements in terms of agents being obligated in terms of the accuracy of the information that goes in there.

That's critical in terms of the FBI in terms of how it operates NCIC. It's also important for us in our own system of record that we have accurate information. But sometimes what we may find in our own system of records is a report, a lead that someone is suspected by somebody of doing something and that we're obligated as FBI, in the FBI, to followup on. And we find during the course of it that's not accurate.

In fact we found that for other reasons that the allegation was made that wasn't accurate. I'm not in a position to discuss some of the technology issues or the statute that you referred to, but obviously, we want everything in our data base to be correct. But I can assure you there is instances where we collect information that when we do further investigation, we find out, in fact, that statement was not correct to begin with.

Mr. CLAY. So then what happens? Do you go back and correct your records? I mean, do you delete that information that is incorrect from the records or do you put it in a different category?

Mr. MCCRAW. No, sir. We just include in the record that we have gone out and done this and determined that, in fact, this allegation was not true.

Mr. CLAY. So that's the system the FBI has in place.

Mr. MCCRAW. Because we're obligated to keep all information that we collect and also show the business process of what we actually did or did not do.

Mr. CLAY. What about some information that is so inaccurate that you list some motorist stopped as being on the FBI 10 most wanted list, and then you detain this person and you find out he or she is not the right person. Then how do you correct that?

Mr. MCCRAW. We're still obligated to report the facts and the facts are that in your scenario that the FBI made a mistake. We still need to maintain the fact that we had a report, we acted inappropriately, we made a mistake and it's still there and it's documented.

Mr. CLAY. It would seem that out of efficiency to law enforcement you would go and clean up that error and take it out of that record so that the next law enforcement officer doesn't pull that information up.

Mr. MCCRAW. I couldn't agree more, Congressman. As it relates to NCIC, and what State and local law enforcement have access to, absolutely. If there's a mistake that has been made, or if someone has been located they have to be immediately taken out of NCIC because that issue could have been resolved. We have to be mindful and the rules and guidelines require that it is immediately corrected if there's an inaccuracy found at NCIC.

Mr. CLAY. And that's what happens now.

Mr. MCCRAW. Absolutely.

Mr. CLAY. Thank you, Mr. Chairman.

Mr. PUTNAM. At this time, I recognize the Vice Chair, Mrs. Miller.

Mrs. MILLER. Thank you, Mr. Chairman. Mr. McCraw, pleasure to have you here. First of all, let me just say that my experience with your agents in the Detroit area has been remarkable. You have some really fabulous folks there that have done a very excellent job. And I felt like I was achieving nirvana with them because I was speaking to them every single day after September 11 in my previous capacity as the Michigan Secretary of State, where we do the motor vehicle kinds of things.

As you might recall, after September 11 there were nine individuals that appeared on every newspaper in the Nation where these fellows had—this is about a week after September 11—these individuals had obtained commercial driver's licenses with hazardous material endorsements. And they were all ostensibly from Michigan. As we found out later, I think there were only two that actually got a CDL with a HAZMAT endorsement from us. The rest weren't Americans, but they were here of Arabic descent that had just gotten a driver's license through the sources that they should have not from another State, actually.

But at any rate, you know you look at some of these things. And that actually led us to make a proposal. I wasn't a Member of Congress then, but something that we had talked about actually did become a part of the Patriot Act in my State, if you want a concealed weapons permit, we do a criminal background check. But if you want a commercial driver's license with a hazardous material endorsement and drive around with 10,000 gallons of liquid propane, no problem, just fill out a form.

So we thought then to try to think like terrorists ourselves. As I said in my opening statements many times, these individuals are using our freedoms against us.

And I sort of preface that with asking you, if you were aware if the Federal motor carrier division has promulgated rules or implemented them in which there is a requirement for an FBI criminal background check for anyone who is receiving not a CDL, Commercial driver's license, but a hazardous material endorsement. We have a large population of people who are of Arabic descent, and I compliment the FBI and Justice. It is very common knowledge there was a large group of individuals who were called in to be questioned in the Detroit area, and I think it was handled with a high degree of sensitivity by the FBI. I certainly again commend the Detroit agents for how they handled that.

But do you have any knowledge if anything is happening in that area where there is a criminal background check now required for those that are getting those kinds of things? What has your experience been as you are creating these data bases as you interact with State agencies such as a DMV or others?

[The prepared statement of Hon. Candice Miller follows:]

Congresswoman Candice S. Miller
Opening Statement
Committee on Government Reform
Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census
May 6, 2003

Thank you, Mr. Chairman.

The issue of factual data analysis as a tool to strengthen national security is one of the most significant issues facing our society today. With the implementation of the E-Government Act of 2002 and the growing importance of information technology in the world, establishing investigative techniques such as factual data analysis are vital if our nation is to implement our war on terror successfully.

As we all know, terrorists have an uncanny ability to adapt to our methods of prevention. In my view, we need to focus all of our resources to ensure that we are always at least one step ahead of the nefarious creatures.

Federal officials currently have at their disposal the resources and knowledge to implement systems that will assist in this process. These officials currently occupy the vanguard of our defenses, and need not be unnecessarily hampered by the bureaucracy in their efforts.

However, the American people must have confidence that the federal government is using this new source of information in an ethical, proper and effective way. It is essential that a proper balance be made between the operations of the government as it prosecutes the war on terror and the disclosure of those operations to citizens it was set up to protect. For this reason, it is the responsibility of this subcommittee to ensure that factual data analysis as a tool is not abused.

For example, I am encouraged that the Total Information Awareness project (TIA) being conducted by the Defense Advanced Research Projects Agency (DARPA) has been relatively

transparent. While only in its beginning phases, this program provides a hope that we can analyze the patterns of the terrorists to anticipate their next move. Some have expressed concern about programs such as these, but the mere fact that Mr. McCraw, Admiral Loy, and Dr. Tether have agreed to testify today, shows that the federal government is concerned about the perception Congress and the public has of data analysis.

I look forward to working with the Chairman and members of the subcommittee and full committee to ensure that this program receives proper congressional oversight.

Mr. MCCRAW. First, I want to thank you for your kind words. I appreciate that. I'll relay that to the special agent in charge in Detroit, and hopefully to his agents as well. I'm not familiar with what has been done right now in terms of regulations. And it may be better asked to the TSA, they may know better. If I'm not mistaken I think that's in their bailiwick, by statute. I do know one of the things that the FBI has been criticized for over the years, that we've been trying to correct, and clearly the director stepped ahead in doing so in your question in terms of how we're using data bases. We take—and also addresses one of Mr. Clay's concerns as well—individuals under investigation for terrorism that we actually have a predicated subject, and they're investigated by the full field investigation, the FBI has taken, the Director ordered and has taken those names and put them into NCIC.

For two reasons, there is predication, there is not an arrest scenario with it, but there is predication so the State and local officers have access to that information who are really the front lines of public safety and need that information.

So those are the types of things that we're doing, and there's a number of other information sharing initiatives that really are technology-based, like a national alert system. We want to be able to reach out to those chiefs of police using the latest technology through their cell phones and PDAs and text messaging, and let them know, we're in the process of doing those similar type of things and even looking, the chairman noted earlier, at the importance, in post September 11, of sharing information in ways that we can do it better and collocate investigative data, multi agencies, use encryption point-to-point over the Internet backbone to access information and provide details. I hope that was responsive to your question.

Mrs. MILLER. It was. Because I think it's so important, as you mention, as we saw since September 11, we've had individuals that have been picked up for routine traffic violations that were suspected terrorists. And the patrol in the black and white car didn't have the information or whatever as they're trying to share some of these data bases. My understanding is that the FBI currently has 31 different data bases, separate data bases which you're trying to combine under this Trilogy project. First of all, why do you have 31 different data bases? How is it working as you try to notify some of those?

Mr. MCCRAW. It's not working. Fortunately behind me, I have the Deputy Assistant Director, who we are fortunate to have in the FBI, whose job is to make those things work. And for years we've had antiquated systems, stovepipe systems, and certainly, Mr. Hooten is in a better state to describe the state of affairs he inherited, but clearly it was a detriment in terms of what we needed to do in our mission.

So I have no defense for it. Clearly it was a problem. Certainly this Director recognizes the need for technology. And fortunately for us, and thankful to you that you've empowered people from the outside that come with the latest technological skills like Mr. Hooten, Mr. Lauer to come in and address this important issue.

Mrs. MILLER. Thank you. I don't have any further questions, but I certainly look forward to working with you and make sure you

have the resources that you need to work these systems. It's critical. No use sitting here pointing fingers on what we should have done 3, 4 or 5 years ago. We need to look to the future. We've got a new enemy. These terrorists are different. They live in the shadows and prey on the innocent. We do need to utilize technology to assist you in doing your job. Thank you very much.

Mr. PUTNAM. Thank you very much. We do have your testimony so thank you.

Mr. MCCRAW. Again, apologize for the delay.

Mr. PUTNAM. My understanding for the record, it was held up at OMB, not at FBI; is that correct?

Mr. MCCRAW. I don't know that I'm supposed to comment on why it's held up. I know one thing, if I had done it sooner, it would have been likely cleared in time. So it's really my fault.

Mr. PUTNAM. Well, we're glad that we have it now.

In your testimony you indicated that you have traditionally used factual data analysis, the collection of data bases prior to Trilogy. Could you please compare what you have done in the past with the technology that Trilogy will provide for you.

Mr. MCCRAW. Certainly. Currently in the FBI we have the system called ACS. It provides an antiquated software over a full text data that allows you to go through a number of green screens. It allows someone to try to do a full text query, just like you would a search on the Internet, a search engine. However, it is so cumbersome and is so difficult and you are overloaded with a tsunami of information that comes back hardly useful.

Moreover, there is no visual link type of tools or link analysis tools that are common use; certainly the Department of Defense uses it; a number of different agencies have been using it successfully over the years.

So right now the type of technology that is being brought on board and actually being used, even though development is being utilized right now by our counterterrorism analysts, is a tremendous benefit to our analysts. They will actually be able to go to specific sets of information for query. One of the major advances using a software package that actually works is that there is no question that we would have gotten to the Phoenix memo just asking the question, in terms of are there any threats in aviation.

Also it's really important to utilize push technology, which the private industry has been using for years.

Again, I'm just describing what we're doing now versus what I was able to do before. And let's say an analyst has a certain issue or topic, let's say it's ricin, anything in the FBI records gets loaded up into if it had ricin, whether it came in the community, whether it was an FD 302, which was an investigative report that the FBI did, or whether it was an insert or electronic communication, that information is pushed to the analyst.

Now, the scope is the prototype, and Mr. Hooten and Mr. Lauer and others in their professional project managers are working on perfecting the technology. And it improves every day. Already we're seeing some tremendous advancements when we standardize the data, provide it in a useful format, and apply these state-of-the-art technology tools on top of it.

Mr. PUTNAM. In quickly reviewing your written testimony, the term Trilogy is never mentioned. Would you define Trilogy for the subcommittee?

Mr. MCCRAW. I think Mr. Hooten is probably better to define it, but Trilogy, I'll attempt it, my understanding is that it is the entire modernization of information technology in the FBI.

If you don't mind, could we have Mr. Hooten.

Mr. PUTNAM. Please stand.

[Witness sworn.]

Mr. PUTNAM. Note for the record the witness responded in the affirmative.

Mr. HOOTEN. Trilogy is two very specific contracts. One is to redraw our infrastructure of networks, which is virtually non existent. That part has already been done. The second part is going through and upgrading our hardware inlcuding our old PCs. The third part is several applications of software, the main one is the virtual case file, which is the replacement, as Steve said, for the old ACS system. So it will be our new system of records. It's the management of our cases. But it's not a data analysis tool which is what this particular subject is. That system that Steve has been referring to is SCOPE, which is a development system that we're currently working on that's made up of the series of COTS products that we are just basically buying off the shelf and doing some quick modifications so that these analysts can have something in their hands that they can use right away. That is closer to this sort of data mining idea, going through sets of data, multiple data bases and looking for particular things.

Mr. PUTNAM. The SCOPE would be more of what you would traditionally define as data mining than Trilogy?

Mr. HOOTEN. Yes.

Mr. PUTNAM. And SCOPE stands for what?

Mr. HOOTEN. I was afraid you would ask me that. I can't tell you off the top of my head. We've been calling it SCOPE so long I forget what it stands for. I can find out for you though.

Mr. PUTNAM. OK. That would be helpful. What new data bases would be searched through SCOPE or through the new Trilogy program that are not currently accessed or utilized today or prior to the deployment of those two programs?

Mr. HOOTEN. Nothing new. It's the same thing we're doing now. The first one is our ACS, which is our system of records, that's our main data base. The nine other things that are very helpful to the analyst are called SAMnet, which are all the cables coming in.

Mr. PUTNAM. According to your written testimony, "the FBI uses information collected by public source companies to obtain information on individuals during the course of its terrorist criminal and foreign intelligence investigative activities." What type of public source companies have data bases that are accessed prior to an event that would trigger an investigation?

Mr. MCCRAW. That we would use as an investigative tool? To name some of the public names, LEXIS/NEXIS, Choice Point. There's several of them out there that have information, driver's license information, government information that they've purchased and that through a query over the Internet and for a fee, you're able to find out additional information about a name. Again, it's a

nice tool. It saves valuable lead time. But it has to be done not on a fishing expedition, it's done based upon a reason. There has to be a reason why you decided to run somebody through that data base.

Mr. PUTNAM. Could you please elaborate some on what role these improvements, and you've outlined two or three different programs, how will they contribute to better collaborative efforts between the FBI and the CIA with the Terrorist Threat Investigation Center?

Mr. MCCRAW. Well first and foremost, it allows the FBI to properly manage its information so we can extract the essential elements of information and to get that in through reports officers, and subject matter experts to get that to the TTIC, also to the Counterterrorism Center as well and to other customers out in the community that need that information.

From the technology standpoint, Mr. Hooten can explain a lot better. It has been standardized so when there's sets of information that the FBI is legally able to provide the intelligence community with, it can do in a standardized format, that it can then use without additional, you know, contractors having to rewrite the format. Better——

Mr. PUTNAM. Before you do, I think, for the record, if you would, please give your full name and your position.

Mr. HOOTEN. William L. Hooten. My position is Deputy Executive Assistant Director over at administration.

Mr. PUTNAM. Thank you. I appreciate that. Thanks for your help.

Mr. HOOTEN. Sure.

Mr. PUTNAM. Mr. Clay raised some interesting issues about the accuracy. What is the level of sophistication of technology today that an Arabic name, for example, would be case sensitive or would certain persons who have the same name and perhaps even the same middle initial and perhaps even the same middle name, what level of sophistication is there to prevent people from being caught up in a mistaken identity?

Mr. MCCRAW. Well, obviously, transliteration has been a problem that all of us face in the government in terms of names. I mean, Waheed Alshiri, I know of at least 14 different types in juxtapositions of the name itself and in public source data alone in which it it appeared. And many times there is insufficient data that you can actually make a determination that it was, in fact, that person. Because there is no date of birth, biographical data or other relational type of data that you can be assured it's that person. That's why it's careful, especially if your operating in the public and proprietary data bases, that there is always followup along those lines, and that it's properly characterized, that information.

Again, a tool within our own system now that we're bringing on greater and advanced tools, there is varying degrees of software that has greater success in terms of discerning those differences, in providing a greater ability of analysts to be able to try to get the transliterations, the juxtapositions or incorrect spellings during the course of an investigation that it was captured.

So clearly, technology has improved. It empowers the analyst and agent to do things that we couldn't do in the past, but it still requires followup work on every piece. And certainly it's an analyt-

ical judgment, an investigative judgment when you brought this information together.

Mr. PUTNAM. Mr. McCraw, we're going to have to move to the second panel recognizing, of course, that everyone will be on the panel together as soon as we have gone through these individually. Somebody had to be first and you drew the short straw. So thank you very much for leading us off on this hearing. At this time we'll excuse the first panel and seat the second.

Mr. MCCRAW. Thank you.

Mr. PUTNAM. Admiral, are you ready?

Admiral LOY. I'm ready to be sworn, sir.

Mr. PUTNAM. Let me introduce you first. We appreciate you being here. And look forward to your testimony. Admiral James Loy is the administrator of the Transportation Security Administration. Previous to his service in this position, Admiral Loy was Commandant of the U.S. Coast Guard and served as the Coast Guard chief of staff from 1996 to 1998. From 1994 to 1996, he was Commander of the Coast Guard's Atlantic area. His other flag assignments were as chief of personnel and training and commander of the 8th Coast Guard district. A career sea going officer, Admiral Loy has served tours aboard six Coast Guard cutters, including command of a patrol boat in combat during the Vietnam War and command of major cutters in both the Atlantic and Pacific Oceans.

Admiral Loy graduated from the U.S. Coast Guard Academy in 1964 and holds masters degrees from Wesleyan University and the University of Rhode Island. Certainly a very distinguished career serving our Nation. We look forward to your service at the new Department of Homeland Security. Please rise and I'll swear you in.

[Witness sworn.]

Mr. PUTNAM. Note for the record, the Admiral responded in the affirmative. Is there anyone with you that needed to do that also?

Admiral LOY. I don't think so.

Mr. PUTNAM. You don't need anybody to answer questions?

Admiral LOY. We'll see, sir.

Mr. PUTNAM. Very well. You're recognized.

STATEMENT OF ADMIRAL JAMES L. LOY, DIRECTOR, TRANSPORTATION SECURITY ADMINISTRATION

Admiral LOY. Thank you, Mr. Chairman and good afternoon. Congressman Clay. Good afternoon, sir. Thanks for the opportunity to discuss CAPPS II as a project with your subcommittee. Mr. Clay, I clearly got your message on time, sir, and we'll make sure that we follow that closely in the future. If I may, sir, I'll offer my written testimony for the record and simply try to emphasize a couple of important points with my oral testimony about this project, and then answer your questions if I may, sir.

First, it's important to recognize that the existing CAPPS system is seriously flawed and in need of replacement. We've studied this system at great length and replacement is the right word. It's too broken in both concept and execution to be upgraded or repaired. Discussing its shortcomings in a public hearing, I believe, is a bit inappropriate, sir, but I would like to offer the committee a follow-on closed briefing if there is any interest in that after our time together today.

My point here is that CAPPS II would be a huge security improvement. I believe of all the elements that we've put in place in designing a system of systems for aviation security first and for the rest of transportation of now and into the future, CAPPS II has the most potential to improve both security and customer service. Our goal is to simply keep foreign terrorists off airplanes. And CAPPS II is a key piece of our interlocking system of systems. And it's also very important to note, I believe, Mr. Chairman, that we are working hard with other detection and screening project owners in the new Department of Homeland Security to ensure good stewardship of the taxpayers' investment in all of these projects. We don't need redundancies; we don't need overlaps; but we do need gaps closed, and we're working very hard to see those goals come to closure.

Second, the goals of CAPPS II are simply twofold and very basic: They are to radically improve the identification authentication of travelers and with that, improve the identification and detection of known and unknown foreign terrorists before they board an airplane, including those that associate with terrorists.

Third, the Aviation Transportation Security Act directed TSA to focus on CAPPS, the current system, and any of its successor systems for improvement and to ensure that any such system would evaluate all would-be passengers before they board. That's precisely what we're doing. Our review to date again clearly indicates the requirement for a replacement system.

Fourth, CAPPS II will be a limited risk assessment tool based on dynamic intelligence information about the activities of known terrorists and their associates. It will be run by TSA inside a compartmented government firewall and sensitive to changing intelligence assessments or the alert conditions, for example, as set by DHS. Think of it, if you will, as a theostat, for the risk assessment scores returned by the CAPPS II tool can be compared to higher or lower limits set by intelligence inputs for the day.

Fifth, unlike the classic definition of data mining as outlined in your invitation to testify, where one searches through reams of data to detect or identify hitherto unknown patterns, CAPPS II will be a traveler-activated search where traveler-offered data elements provided to secure a reservation, name, address, phone number, and date of birth, will initiate the identification authentication score, as well as the risk assessment score. The search will be to determine if well-known patterns are found to drive that risk assessment score higher than the threshold acceptable for the security environment of the moment.

The products of the search, the result and risk score and identification score will translate to a simple direction to the checkpoint at the airport. Green, have a nice flight; yellow, provide secondary screening before boarding; or red, refer to law enforcement as one whose risk score deserves law enforcement scrutiny, and in any case, who will not board that flight.

Sixth, we have recognized the importance of privacy concerns and have conducted considerable outreach to the privacy and civil liberties arenas. At multiple day offsites, we have collaborated with a good number of those with privacy concerns because we have them as well. We have met with groups of privacy officers from the business community, with groups of stakeholders in the privacy

arena, and with individuals who have a deep conviction about the fourth amendment. We've also met with congressional representatives, with Senators and their staffs, and have listened intently so as to develop a privacy strategy for CAPPS II that will be a strength of our system, not a weakness.

Most recently, we have engaged the review of Ms. Nuala O'Connor Kelly, the newly established privacy officer at the Department of Homeland Security, to validate our commitment to doing this right. And as we speak, she is with our delegation to the EU to sort out international privacy concerns. And I am simply of the mind that we can design, if we put our minds to it, a solid program where security and privacy are complementary goals. That's why we designed our Federal Register notice to solicit the widest spectrum of comment possible. When our listening is done, we will reissue that notice based on what we've learned.

Seventh, CAPPS II will not build data bases on U.S. persons permitted to fly; will never see the commercial data being used to authenticate identification; will not search medical records or criminal records nor see credit ratings, overdue bills or any such data to assess risk; will not generate new intelligence; and will not keep even the risk scores after travel is complete.

Eighth, CAPPS II will be a serious resource allocation tool. It will allow us to better schedule Federal air marshals, to better schedule screeners, even the new Federal flight deck officers—"guns in the cockpit," if you will—so as to optimize the resources we have to throw against the security problems we face.

Last, Mr. Chairman, CAPPS II will be counterintuitive in the sense that we will not be looking for the proverbial needle in a haystack. Rather, we will be taking that haystack off the needle, identifying those thousands and thousands of perfectly innocent travelers, opting them in, if you will, thereby leaving only those we can't evaluate as OK to be the subject of added scrutiny.

Today about 15 percent of the travelers are dubbed selectees, and undergo secondary screening. We believe we can bring that percentage way down, and thereby make not only a significant difference in security, but also a significant difference in customer service as well.

Thank you for your attention, sir. I look forward to your questions.

Mr. PUTNAM. Thank you, Admiral.

[The prepared statement of Admiral Loy follows:]

U.S. DEPARTMENT OF HOMELAND SECURITY

STATEMENT OF ADMIRAL JAMES M. LOY
ADMINISTRATOR, TRANSPORTATION SECURITY ADMINISTRATION

Before the

SUBCOMMITTEE ON TECHNOLOGY, INFORMATION POLICY,
INTERGOVERNMENTAL RELATIONS, AND THE CENSUS
COMMITTEE ON GOVERNMENT REFORM

U.S. HOUSE OF REPRESENTATIVES

May 6, 2003

Good afternoon, Mr. Chairman and Members of the Subcommittee. Thank you for the opportunity to appear before the Subcommittee on behalf of the Transportation Security Administration (TSA) to discuss how the use of information technology can strengthen transportation security. TSA was established under the Aviation and Transportation Security Act (ATSA) just weeks after the tragic events of September 11, 2001. Since then, TSA has worked diligently to deploy Federal screeners and explosives detection systems at more than 429 airports across the country, dramatically expand the Federal Air Marshal (FAM) program, and enhance perimeter security at airports.

Now, having met the initial deadlines of ATSA, we are expanding our security efforts in other modes of transportation and finding ways to continually improve aviation security. One of the most promising opportunities for improving both efficiency and effectiveness in aviation security is to make greater use of information technology and risk analysis tools. Information technology can play a key role in protecting citizens from terrorist threats while protecting their privacy. This hearing is an important forum for developing a shared understanding that security and privacy are complementary, not conflicting, goals.

Currently, airlines operate the Computer Assisted Passenger Prescreening program (CAPPS) to identify passengers for enhanced screening before those passengers are permitted to board a commercial aircraft. In ATSA, Congress directed TSA to ensure that CAPPS or any successor system is used to evaluate all passengers before they board an aircraft and to include procedures to ensure that individuals selected by the system and their carry-on and checked baggage are adequately screened. As a result, TSA is developing an enhanced Computer Assisted Passenger Prescreening program (CAPPS II), which will far more effectively identify passengers that may be a risk to the aviation system.

The purpose of CAPPS II is to identify foreign terrorists and those with links to foreign terrorists that pose a threat to civil aviation security. CAPPS II also will allow TSA to

make more efficient use of screener resources by using dynamic intelligence information to select passengers for enhanced screening.

CAPPS II will be a limited, automated screening tool that will be operated under the direction of TSA and that will form a critical element in our strategy of providing overlapping layers of security to protect aviation from curbside to cockpit. Passenger pre-screening under CAPPS II is an essential component of our system-of-systems approach to aviation security.

CAPPS II will establish a more standardized risk assessment system, dramatically reducing what some travelers view as arbitrary selections of passengers for enhanced screening at airport security checkpoints. Indeed, by using this tool to focus screeners' efforts on passengers who appear to pose a heightened risk, much of the additional screening that is now performed may be eliminated. At present, under CAPPS, some 15 percent of passengers traveling within, through, or out of the U.S. undergo enhanced screening. Under CAPPS II, we expect that percentage to drop significantly, thus expediting travel for many passengers without compromising security.

CAPPS II will reduce much of the current confusion involving persons with similar names. Public trust and confidence in the security of air travel will increase with a more robust and fully audited pre-screening system. Implementation of CAPPS II also will relieve air carriers of the financial burden of operating the current CAPPS system, a cost airlines estimate at over $150 million annually.

In all we do, TSA strives to provide world-class security and world-class customer service. We cannot be successful in serving our customers without taking great strides to protect their privacy. TSA is mindful that privacy protections must be built into the CAPPS II system from its very foundation. We have been working with Congress and stakeholders in the privacy and civil liberties communities, and have made good progress toward that end. As Secretary Ridge has stated, we will not implement CAPPS II until the Department of Homeland Security's chief privacy officer has reviewed and approved the privacy protections in the program. I am pleased to report that the Department's new privacy officer, Ms. Nuala O'Connor Kelly, has already begun her review of CAPPS II. Under the E-Government Act, TSA is working to finalize its CAPPS II business case, which will detail how privacy and security are built into the system. TSA also will conduct a Privacy Impact Assessment.

Essentially, CAPPS II will be a passive system that produces a general indication of the level of terrorist risk each airline passenger might pose to civil aviation security. It will be activated by a traveler's airline reservation request. Airlines will ask passengers for specific reservation information that will include a passenger's full name, plus other identifiers including date of birth, home address, and home phone number. Passengers will not be asked to provide social security numbers, and TSA will not look at credit worthiness.

The CAPPS II process will then authenticate each passenger's identity through publicly and commercially available databases. Once a passenger's identity is authenticated and the passenger's information is run against terrorist or other appropriate Federal government systems, an aggregate numerical threat score will be generated that TSA will use to determine which passengers should proceed through the ordinary screening process and which passengers should be asked to submit to a somewhat more thorough screening. In extremely rare cases, the system may identify an individual who is a known foreign terrorist or the associate of a known foreign terrorist. In such a case, law enforcement authorities would be notified and given the opportunity to take appropriate action.

The entire risk assessment process will be conducted in less than five seconds. It is important to stress that TSA will rarely see the public-source information that is checked to authenticate passenger identity. The exception may be in the extremely rare case of passengers who are positively identified as known terrorists or associates of known terrorists. In such a case, the Federal government would need this information for enforcement purposes.

The CAPPS II process will allow the vast majority of passengers to simply go through ordinary screening. Fewer passengers will be subject to enhanced screening under CAPPS II, and this will lead to shorter lines. CAPPS II may be compared to an electronic lock protecting a secured area. You must identify yourself and satisfy the system before you are allowed entry. The system's algorithms will be designed to confirm passenger identity information, identify known and unknown foreign terrorists, and recognize connections to known terrorist-related activities or individuals.

CAPPS II is a passenger-screening tool only. It will not ingest or store large quantities of data. Very importantly, CAPPS II is not data mining in that it will not explore databases to extract information to identify patterns of behavior among travelers.

CAPPS II will operate under a stringent privacy protection protocol being developed through discussions with privacy groups, both in the U.S. and internationally, with Congress, and with the public. Strict firewalls and access rules will protect a traveler's information from inappropriate use, sharing, or disclosure. CAPPS II will not retain data on U.S. passengers that are permitted to fly. Once travel is completed, CAPPS II records on these passengers will be purged.

I want to recognize the valuable contributions the privacy community has made in the development of this system. In March, TSA held a three-day privacy summit that was attended by many leading privacy experts. The discussion was frank. TSA listened carefully to all views, and we are giving these views full consideration as decisions about CAPPS II are made. In the weeks ahead, we are holding public meetings around the country to get the view of as many people as possible so that our privacy protections respond to the concerns and have the support of most Americans. Briefings for officials, privacy advocates, and opinion leaders have been and will continue to be conducted on a regular, on-going basis. We are reviewing comments submitted in response to our

Federal Register notice and will issue a new notice based on comments received. TSA, in conjunction with the State Department, is also working with the European Commission to ensure that international privacy concerns are fully addressed.

Based on TSA's outreach to the privacy community and the public, additional privacy measures are being incorporated into CAPPS II. CAPPS II will minimize the amount of information on travelers that ever comes into the system, using only the information that is necessary to conduct an identity authentication and risk assessment. The base information needed to operate CAPPS II will be provided by passengers themselves. The CAPPS II authentication function will be conducted for the most part outside government databases using commercially available data, and very importantly, data in those systems will not be viewed by TSA. Employees of commercial data companies assisting with the authentication process will never directly view or acquire records of traveler personal information from TSA. A system of firewalls will prevent these companies from ever directly using or retaining personal information.

TSA wants travelers to fully understand at the outset how information they provide will be used. To that end, we are developing procedures consistent with the Privacy Act to provide timely notification to individual airline ticket purchasers of the purpose for which we are obtaining information about them and the need for such information.

TSA will maintain a policy of openness and public accountability. When a passenger feels that he or she is being singled out for heightened scrutiny, complaint procedures will enable that passenger to bring a grievance to the attention of TSA. TSA is committed to affording any aggrieved passenger prompt access to appropriate redress or assistance.

It will not always be possible to inform a passenger of the reason for any additional screening that may be performed. However, TSA's Passenger Advocate will be empowered to look into any issues or concerns raised by a passenger.

Security is a primary concern for TSA in the construction of the CAPPS II program. TSA takes the responsibility for handling personal data very seriously and will use the best technology to ensure that data is handled properly and that inadvertent disclosures do not occur. To that end, CAPPS II will be a policy and security based system with real-time auditing. Access to the system will be limited to those with an appropriate need, and the system will monitor and identify precisely who accesses the system, when it was accessed, and for how long. Data input will be validated to ensure that it is correct, authorized, and appropriate in light of applicable restrictions. The CAPPS II design will ensure that data is securely transferred between systems and end-users.

We are also working aggressively across the Department to ensure that CAPPS II technology is appropriately leveraged with other DHS investments in screening technology. This Department-wide review and planning will help achieve a key DHS goal of preventing unnecessary duplication and wasted taxpayer dollars.

The CAPPS II system is still under development. We are now testing the technology to ensure that it functions properly, but we are not piloting the system itself. Passenger data is not being transferred or processed. TSA expects to have CAPPS II fully implemented by the summer of 2004. I look forward to working with this Subcommittee in the months ahead as we develop CAPPS II, to realize the efficiency and effectiveness it offers in improving security while protecting the privacy interests of travelers.

Thank you for the opportunity to appear before your Subcommittee. I would be pleased to answer any questions you may have.

Mr. PUTNAM. As with our previous witness, our ranking member will be recognized first for questions.

Mr. CLAY. Thank you, Mr. Chairman. Thank you Admiral for being here today. I want to salute you for your leadership in the area of airline security, over the TSA and being willing to serve this Nation in that capacity.

Admiral, airline security has a troubled history of racial profiling, even before the attack on the World Trade Towers. During the 1991 Gulf war, individuals with Middle Eastern names were forced off their flights despite the fact they were American citizens. One gentleman, an American citizen, whose parents were from Bangladesh, was told he should carry his passport to prove his citizenship.

Last year the ACLU testified before Congress of dozens of such incidents, individuals discriminated against in airports or on airplanes based on race and heritage. The same people who oversaw the private contractors who provided discriminatory security are now designing new systems. What is TSA doing to prevent racial profiling from continuing in our air transportation?

Admiral LOY. Mr. Clay, the design work associated with CAPPS II as the replacement for the existent CAPPS program in place today has a very clear specification: there will be no racial or gender profiling. We frankly don't believe there's any value in going in that direction at the other end of a security risk assessment. What one has to do with the other is simply unknown to us. So the design work here is to keep such things totally out of the picture by specifying in the contract that we don't go there. There is no reason for it either, and it's certainly totally out of both my personal and our organizational ethic.

Mr. CLAY. You mentioned in your testimony the random checks that occur, I was leaving Orlando a couple weeks ago, from the Chair's area, I had a one-way ticket back here to Washington and I had four "S″s on my ticket. Quickly routed into special security, take off the shoes and all of that.

Mr. PUTNAM. It's not racial, man, I was there with you.

Mr. CLAY. You did it, too. I mean, so now under CAPPS II it will be less and less of that?

Admiral LOY. Absolutely, sir.

Mr. CLAY. How do you get picked?

Admiral LOY. Today CAPPS is a rule-based system. It has been that way with no changes to those rules for a rather lengthy period of time. In the immediate wake of the tragedies of September 11, 2001, those rules were actually reinforced. I would rather not go into a public listing of those rules, but I can tell you that they are recognizable; they are compromisable; they are broken, sir. That is exactly the reason why CAPPS, as a system in place today, needs to be replaced.

We will not have rules associated with the manner in which CAPPS II will do its work. We will take advantage of added pieces of information that we will ask of travelers. A condition of a reservation is that you no longer just give your name; you give your name, address, phone number and date of birth, and allow the extraordinary technology of today to give us a risk score associated with authenticating that identification.

Mr. CLAY. Thank you. Admiral, 2 weeks ago the Wall Street Journal ran an article on the problems created by the no-fly list. That article began with the story of Larry Mussara. Mr. Mussara is a retired Coast Guard Commander, the father of three, a local hero in Alaska for his daring helicopter rescues of stranded fishermen and mountaineers.

But, every time Mr. Mussara flies Alaska Airlines, which is about once a month, he gets stopped, often missing his plane. This kind of error occurs because Mr. Mussara has a name similar to one or more names on the no-fly list.

In CAPPS II, the TSA is going to use a number of private data bases to make these same kind of comparisons. Will that increase the chance of a missmatch like the one Mr. Mussara faces?

Admiral LOY. No, sir. It will actually radically decrease the chance of a mismatch. I want to make it clear here, Mr. Mussara will not be singled out in the days when CAPPS II is an active program, though he is focused on time and time again, unfortunately, under the program that is in place today.

Mr. CLAY. Just to wrap up with you. Can you share with us how many names are on the no-fly list? Is that available?

Admiral LOY. Again, sir, I think that I would prefer to tell you that in confidence or in private. I will be happy to do that. I can offer that CAPPS II compared to CAPPS as it is in place today is quantum levels better in both identification scrutiny, and in risk assessment across a watch list created predominately by the Justice Department in the Terrorist Tracking Task Force over the course of this last year. So we are talking hundreds on one hand, tens of thousands on the other.

Mr. CLAY. Under CAPPS II the kind of mistakes that Mr. Mussara encounters will not occur?

Admiral LOY. That is exactly right, sir.

Mr. CLAY. Well, thank you very much for that.

Thank you, Mr. Chairman.

Mr. PUTNAM. The gentlelady from Michigan, Mrs. Miller.

Mrs. MILLER OF MICHIGAN. Thank you, Mr. Chairman. And to my colleagues, I have also been under selection numerous times. It seems, to go through the whole check.

I have to tell you that I was on a flight on Friday night with six Members. I noticed all of the Republicans were being selected, but not the Democrats. So I don't think it is racial, it is partisan. That is my observation.

Admiral, I appreciate your testimony and certainly your service to our Nation as well. I would like to ask a question as you talked to categorizing it as a risk assessment score, perhaps we can call it a threat score, what have you.

Will it only be the TSA that would have that kind of information of a threat score? Will you be utilizing or sharing, or sharing any of this information with commercial entities? If you do intend to share any of it with commercial entities, how can your organization ensure that the commercial entities are not sharing this information when they should not be?

Admiral LOY. Ma'am, we will not be sharing the risk scores, either with respect to identification or with respect to final risk, with

anyone outside that firewall I described as being the break point between inside government and outside government.

We will be enormously concerned about four or five privacy parameters that I believe are the framing elements of what we wanted to build our privacy strategy around. But particularly to your point, we will not be getting into data itself; we will be designing arithmetic algorithms that will be able to search those data bases.

Our first effort will be to take those traveler-initiated pieces of information—PNR data—aggregate them, and send them to commercial data bases for the manipulation, if you will, that offers us back an identification authentication.

Goal one is to be able to look travelers in the eye and have great confidence that they are who they claim to be. That is job one, to get us from where we are today, with a name-based system only, with its potential for challenge, and this goes directly to Mr. Clay's question as well, and toward a system where we have great confidence that the person who is asking for this reservation is the person he claims to be.

Second, armed with that, we will run that authenticated name against the government data bases. That will reflect for us a final risk score determination. That will only be shared with another law enforcement organization if the purpose for which CAPPS II has been met, should they or should they not be allowed to board that plane. And there is a recognition in that risk score that law enforcement attention is actually in order.

Mrs. MILLER OF MICHIGAN. Admiral, you said, "we want to look them in the eye and make sure that they are the person that they are claiming to be." So, let me ask if you have any comment on using technology, the retinal scans, looking them in the eye. That is the best technology that we really have available today.

I will tell you, as a frequent traveler, I would be happy to have presecurity clearance and look me in the eye and make sure that you do the retinal scan and let me through the lines rather than standing there forever.

Do you have any comment about whether we ever get to that point? I recognize the privacy advocates are talking about that. But, I mean, I think it makes a lot of sense.

Admiral LOY. Well, we are actually working with the privacy advocates at the table, on design work, both with respect to CAPPS II and with respect to another project that we have underway, the transportation workers identification credential. The notion there is biometrical, such that the identification, so beating the so-called identity theft issue is very much within our grasp.

CAPPS II will not have a biometric base associated with it, but I think we are only months away from having as a foundation block CAPPS II on one hand and the TWIC program on the other. We will build a registered traveler program that will be biometrically based, that will seek those players who are willing to step forward to get the background investigation, and get it accredited in the form of a biometrically-based card, so that we can facilitate a quicker passage through the airport system.

Everyone will always go through the basic screening, and then of course if you trigger an alarm, based on having gone through the magnetometer, whatever screening is required there.

If we can facilitate that in a frequent flyer line, or in some fashion that we can work out, and I am very optimistic that we can do that with the airlines, because they see it as a great value as well, we will end up doing exactly what you described. I would like to call it a registered traveler program, rather than frequent, which is, of course, associated with airlines exclusively, if you will, or trusted, the obverse of which I am not too keen on, labeling people as untrusted travelers.

So we are right there. That is exactly the design work that we are following.

Mrs. MILLER OF MICHIGAN. Well, call it whatever you want to. I would like to sign up when you get it available, please.

You mentioned that the purpose of CAPPS is really, a primary purpose is obviously to improve identification processes and those kinds of things. You heard me mention earlier I had been a person that was in charge of DMVs.

And, of course, the driver's license really has become sort of the critical foundation of establishing anyones identity, whether they are utilizing driver's licenses, whether they are utilizing State identification cards, what have you.

Let me ask you to comment on something that Secretary Ridge has, I know, made some comments on. There has been a lot of discussion amongst all of the States about the possibility, the potential of having a nationwide driver's license, because currently the type of primary documents that are required by the individual States to establish identity have such a huge fluctuation it is unbelievable. We take a lot of pride in Michigan, we think that we have some of the more stringent standards in the Nation. Minnesota also has very stringent standards.

But I have never been able to figure out why it matters whether you get a driver's license in Minnesota or Michigan or Tennessee, or what have you, and that you have all of these different requirements. It must be an unbelievable challenge for yourself, the FBI and others, looking at these driver's licenses that are often times issued with erroneous documentation or very little kind of primary identification requirements.

I think people are sometimes startled to know that it is the rule rather than the exception that almost every State in our Nation must issue a driver's license or a State identification card to people that we know are here illegally, illegal aliens are getting these driver's licenses. By most of our State laws we are required to give those out.

So I can't imagine what kind of impact that is having on the CAPPS program, and some of these others as you try to identify trusted travelers.

Admiral LOY. Yes, ma'am. I think you are absolutely right on point. We have given up using that base as a means by which we can gain confidence that a person who claims to be whoever they claim to be is really that person.

In the other work of the Transportation Security Administration across all modes of transportation, that is where the transportation workers' identification credential is going, because, among other reasons, we can't have faith in the systems you were just describing.

The notion of a biometrically based transportation workers' credential, both for identification purposes and for access control purposes, is where we believe we need to go in the transportation system at large. And we are working on two prototype projects in that regard, one in the Philadelphia area, and one in the Los Angeles area over the next several months.

So I reinforce your concerns, and let you know that we can't go there with comfort and have to design a better mousetrap. I would also offer, based on your question to our first panelists, that just last Friday, ma'am, we did issue an interim final rule on hazardous materials endorsements on CDLs in conjunction with FMCSA, as well as RSPA, the Research and Special Programs Administration in DOT, with Justice alongside in terms of making sure we have met that requirement that you described in the Patriot Act.

And it does require a BI, which is exactly, I think, the question you asked.

Mrs. MILLER OF MICHIGAN. Thank you, Admiral. Thank you, Mr. Chairman.

Mr. PUTNAM. You are very welcome.

Admiral, several things about the mechanics of CAPPS II program. First of all, the international terrorist organizations have shown a remarkable agility in selecting a variety of different targets. And in response to September 11th, I think that we have disappropriately focused our efforts on protecting airline safety at the expense of rail, passenger cruise ships and other potential threats. Will this same technology be deployed for rail and passenger cruise lines?

Admiral LOY. The potential is very much there, sir. My notion is that CAPPS II is a phased kind of project. Our first goal needs to be to construct; I have analogized it to our closets at home. And our first challenge is to build the rail. And then how we develop and use the multiple applications that might come from the risk assessment engines that will be designed as the rail. I liken those to multiple hangers sequentially being put on that rod over time.

And with not only the knowledge of, but the consent of oversight-responsible organizations, not the least of which, of course, is the Congress.

So, yes, my charge from Secretary Ridge is to build a national transportation system security plan, not an aviation security plan. And it does go to aviation, it does go to maritime, and it is about rail, transit systems, highways, and pipelines. All of those, plus maritime and aviation, compose our national transportation system.

My goal is to make sure that Secretary Ridge is not found with a weak link among any or all of those aspects of our system. And, of course, that is just one of the puzzle pieces he has to fit into his much bigger challenge across the rest of our homeland.

Mr. PUTNAM. Do you currently have congressional authorization to deploy that beyond air travel?

Admiral LOY. We do not. That is exactly why I said it would be enormously important for us to come back and think it through carefully, not only with authorizing committees, but of course to seek the appropriations necessary to make it happen.

Mr. PUTNAM. And how would the technology detect or review or assign a threat score to suspect domestic terrorists?

Admiral LOY. The process would be absolutely similar, sir. Armed with those four pieces of information, the system would first of all build that identification score. That score then, as part of a review of those government data bases, would allow us to assign that final risk score. It frankly doesn't matter whether it is a foreign terrorist, although that is what we are looking for. Our challenge to allowing a U.S. person on that plane would simply be based on the fact that their score had elevated beyond the threshold of going from yellow to red, and we would then allow an investigative effort to take place by the right law enforcement organization.

Mr. PUTNAM. So it would then also detect persons who are not necessarily a threat to that airline, but who are wanted for some other crime?

Admiral LOY. As you heard me say earlier, sir, we are not searching NCIC as part of the data that we are looking at. This is a very focused tool, designed not without potential to do other things, if authorized and challenged by the Congress to do so, but at the moment, we are charged with finding, in the aviation sector, foreign terrorists or those associated with foreign terrorists and keep them off airplanes. That is our very limited goal at the moment.

Mr. PUTNAM. Certainly the additional hangers that would be in your closet would inevitably lead to the technological ability in detecting anyone on rail, seacraft, aircraft, who would be then reviewed in data bases, that would include any number of warrants outstanding for any number of crimes.

Admiral LOY. The potential there is very real, sir. And frankly at the other end of the day, even as heinous as it sounds, the ax murderer that gets on the airplane with a clean record in New Orleans and goes to Los Angeles and commits his or her crime, is not the person we are trying to keep off that airplane at the moment.

Mr. PUTNAM. Today.

Admiral LOY. Yes, sir.

Mr. PUTNAM. But, clearly circumstances could change?

Admiral LOY. As I indicated, there are several issues here.

First of all, Mission Creep, if you will, is one of those absolute parameters that the privacy community is enormously concerned about, and I am enormously concerned about. We will build such concerns into the privacy strategy that we will have for CAPPS II.

On the other hand, over the course of time, with an airing, clearly with the oversight associated with not only the Congress but our continued collaboration with our privacy colleagues, there are changes that can be made. There can be additional hangers hung on the closet rod.

Mr. PUTNAM. Let me—you mentioned that the score—the criteria that determine the score change, depending on different circumstances.

Admiral LOY. Yes, sir.

Mr. PUTNAM. Who sets the criteria? Who makes those changes based on other intelligence or other circumstances?

Admiral LOY. Well, I would offer, sir, that it is the identification score that comes back first across that firewall from its mix in the commercial data bases that will be searched. Armed with that score, we then assess the risk and produce the final risk score, and that score is going to be probably the same, regardless of what is also happening at the same time.

That process will run its course, and we will end up with a risk score for that traveler. In the meantime, if we are at alert condition blue or yellow or orange or red, there may be enormously different attention being paid to one thing or another.

If the intelligence drift over the course of that past month or week or day or hour is being focused on an airline or an airport or a flight or such things as that, then we would have the ability to adjust that rheostat in such a fashion, if there is a score of—pretend it is 100 max, if it was a score of 94, if 1 day would find you in a yellow capacity as opposed to red, based on the focus of intelligence that day, the security environment, if you will, that is associated with the world in which we are living in.

We believe it is enormously important for this system to recognize adjustments in the flow of intelligence across our daily desks and be able to do something about it. At the moment, nothing like that exists in the CAPPS system that is on the books today.

Mr. PUTNAM. And all of that is based simply on name, address, phone number, and date of birth?

Admiral LOY. That is correct. That is the only data; that is only four pieces of data that, A, the traveler will offer, that, B, TSA will aggregate, that, C, will go to the commercial review process to produce the identification authentication score.

And when it comes then back into across that firewall to be assessed against our government data bases for the final risk score, there are no other pieces of data. We will only see scores, not data beyond what the traveler offers us.

Mr. PUTNAM. It will be sensitive enough that if Mr. Clay purchases a ticket from Washington and uses his Washington address and phone number, and returns back from Missiouri using his Missouri address and phone number, the discrepancy alone will not flag him red?

Admiral LOY. That is exactly the case. Because, first of all, if it ever occurs, his opportunity for redress is, again, one of those parameters in the privacy strategy that we will have for CAPPS II that offers an appeal system to challenge decisions that have been made.

Now, to the degree he goes red on that flight or even yellow, he has every right to call us. We are going to establish an ombudsman—we are not going to call it that, passenger advocate, I think, is the phrase that we are going to use—whose purpose in life is going to be to take calls from people who feel that they have been misread by the system, and adjust accordingly.

That person will be able to search the data and come to the right answer.

Mr. PUTNAM. In a timely manner for him to make his flight?

Admiral LOY. I can't say that it will be in a timely manner for him to make his flight. If it requires research that the person, that

the passenger advocate has to do for us, that is not a promise that I can make today.

Mr. PUTNAM. I have overstepped the bounds of time that I set for everyone else. I apologize. We will excuse this panel temporarily and swear in the third panel and then bring all three of you back forward.

So at this time we will excuse you, Admiral. Thank you for your testimony. And we will welcome panel three.

Admiral LOY. Thank you, Mr. Chairman.

Mr. PUTNAM. Thank you, sir.

[Witness sworn.]

Mr. PUTNAM. Note for the record Dr. Tether responded in the affirmative.

Dr. Anthony J. Tether was appointed as Director of the Defense Advanced Research Projects Agency on June 18, 2001. DARPA is the principal agency within the Department of Defense for research, development and demonstration of concepts, devices and systems that provide highly advanced military capabilities.

As Director, Dr. Tether is responsible for management of the agency's projects for high-payoff innovative research and development. Prior to his appointment as Director, DARPA, Dr. Tether held the position of chief executive officer and president of the Sequoia Group, which he founded in 1996.

He has served as chief executive officer for Dynamics Technology Inc., vice president of Science Applications International Corp.'s Advanced Technology Sector, and then vice president and general manager for Range Systems at SAIC.

He spent 6 years as vice president for technology and advanced development at Ford Aerospace and has also held positions in the Department of Defense serving as Director of DARPA's Strategic Technology Office from 1982 through 1986, and as Director of the National Intelligence Office in the Office of the Secretary of Defense from 1978 to 1982.

Prior to entering government service, he served as executive vice president of Systems Control from 1969 to 1978, where he applied estimation and control theory to military and commercial problems with particular concentration on development and specifications of algorithms to perform realtime resource allocation and control.

Your mother must be very proud.

Mr. TETHER. Well, I haven't been able to hold a job for very long.

Mr. PUTNAM. You do move around a lot. But you are not exactly slumming. We welcome you to the subcommittee, and look forward to your testimony, if we can understand it.

STATEMENT OF TONY TETHER, DIRECTOR, DEFENSE ADVANCE RESEARCH PROJECT AGENCY, DEPARTMENT OF DEFENSE

Mr. TETHER. Well, thank you very much. I would like to offer my written testimony for the record if I can.

I am not going to really go through much of the written testimony, since you have had it, and have had a chance to review it. So I just want to make a few points.

First of all, my testimony today is less on TIA, the Total Information Awareness Program, as a program and more addresses just a very small part of that program, the data mining part of it.

The TIA itself is a much larger program dealing with collaborative technology, language translation, biometrics identification and so forth and so on. Now, on the other hand, you all will be getting a major report on May 20th, assuming that I can get it through all of the coordination that still has to go on, which will describe in great detail for you the Total Information Awareness Program, as a program. So you will have that available to you very shortly.

Data mining, as sometimes used, more commonly used, refers to the clever statistical techniques which basically seek to comb through large amounts of data looking for previously unknown but useful possible patterns. And, as you know, it has been used commercially by pharmaceutical companies and so forth and so on.

The problem is that this approach, while useful for coming up with correlations, does lead to many false positives and so forth and so on. Also, it typically requires that all of the data be centralized in one place for those algorithms to work. We are really not pursuing that technique. I want to really make that clear, primarily, because if you have done nothing but read the papers about TIA, you are thinking that we doing nothing at DARPA but just piling through tons and tons of data about people in the United States looking for possible wrongdoings, and nothing could really be further from the truth.

Our approach basically starts with a hypothesis about attack scenarios. Given an attack scenario, we create a model, a model which basically says, if this is the attack scenario being carried out, these are the observables, these are the questions that if we asked them, that came up positive, would indicate to us that this attack scenario was underway.

So basically what we end up doing, we spend a lot of effort and time, basically in creating a model, which ends up with a pattern, a pattern that indicates that model or attack scenario is true.

We then take that pattern to the data base, and look in the data base to see if that pattern exists. Now, this allows us basically to really cut down, to narrow to scope on the data bases on the answers. Also, it also allows us to let the data bases remain distributed. We don't have to bring the data to a central location with this approach.

Basically, this is not a new approach, one of the questions was, well, how do you know that your approaches are going to provide any security? This is really an approach that has been under development by DARPA and other agencies for many years. One of the examples is in image processing.

When we receive an image, we can either have an analyst go through and look at every little pixel to see if a target is there, or we can develop a model that says, this is what a target looks like in this picture, and then have the algorithms go through the image to find if that target is there.

And this is called automatic target recognition, very successful. We have used it for years, and quite frankly, have just recently used it in the last Iraq war very successfully.

Privacy, however, is really a major concern to us. And from the outset of the TIA program, we have really worried about privacy. Now, we worried about privacy for perhaps a different reason. But it does transcend into the public. As you can tell from reading my resume, I have been around for a long time. And every time there has been an intelligence failure, it has never been because we didn't have the data. We have always had the data.

And, I think, as you see from the hearings on September 11, it turns out we always had the data. The problem was that the data was distributed. The problem is that some of data was held by CIA. Some of the data was held by NSA. The problem was is that there was no method to have everybody collaborate to work on a problem to bring that data to be able to answer questions.

One of the reasons that there is difficulty in doing that, is that the agencies try to protect their sources and methods, for good reasons.

I mean, this is for really good reasons. SIGINT, signals intelligence, is a special case. It has special rules and regulations because signals intelligence is gathering information overseas, but it can accidentally pull in what is known as data about U.S. persons, which is not just people, but also corporations. So there is a regulation that prevents people from really automatically sharing raw data.

So we really are worried about the privacy concerns from how do we develop a system which would allow this collaboration to take place with distributed data bases for everyone to get together virtually, work on a problem, yet be comfortable that the privacy of their data bases was going to be held.

And we have spent a lot of time and a lot of money on that. One of the major things we also believe is that an audit technique has to be developed. In other words, in order for people to feel comfortable about somebody seeing a piece of data, they want to know, in my world, the DOD world, there is a technique known as ORCON, originator controlled data.

This is an attempt by the originator to make sure the data is not not used for other than the purpose for which it was granted to the individual. We believe that you need to have an audit technology which attaches itself to the data so that everyone, from that moment on, will know who is looking at that data and where has that data gone. We are spending a great deal of money and effort trying to do that too.

I think I will stop, because time is getting late and I ask you for your questions.

Mr. PUTNAM. Thank you, Dr. Tether.

[The prepared statement of Mr. Tether follows:]

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.

Mr. PUTNAM. We will lead off with Mr. Clay, again.

Mr. CLAY. Thank you, Mr. Chairman. And thank you, Dr. Tether for being here.

You indicate that TIA will use transaction information held by private companies. What will you do to assess the accuracy of those information systems?

Mr. TETHER. Well, actually, I don't think we ever really said that TIA will use transaction information held by private companies. I know it has been said in the press. Our emphasis really has been on the data that is currently legally collected by our intelligence community, and also the counterintelligence community.

Now, what we have done, I, personally, believe that all of the data that we really need to have in order to detect these attacks already exists in that legally gathered data, because of my experience.

On the other hand, we have looked at, "Is there other data that perhaps would accelerate the process of coming to a conclusion about whether an attack is going on?" So we have had research ongoing to see if there was other data besides data that is normally collected by the intelligence community that could be used.

Now, we have a lot of researchers. DARPA itself does not have any internal capability. We contract out. We have researchers and industry that really do the work. So we have to describe a problem to them to work on. And in some of the descriptions of the problems we talked about transaction data. And I am afraid that they probably took it maybe a little more literally than we really meant it. We didn't mean credit card data. We really were looking for people to research data as to what might be extra data.

Now, we have found in that research that transportation data is probably data that makes a lot of sense to be included into the mix. Why? Because these terrorists have to travel. And they don't have their own means of communication, which means they are going to have to take commercial means of communication like airlines, trains, rental cars, rental trucks and so forth and so on. So it looks like, if we were to add data that looks like it must be very advantageous, we would say transportation data really looks like it would be advantageous.

But, again as to how do we know the data is good, and I don't mean this as a cop-out, but at DARPA we develop the tools. You have heard about these other two organizations who want to have this capability, but they can't have the capability if someone doesn't develop the tools.

We are the ones that basically develop the tools for them to use. And presumably we don't collect any data, but we are developing tools for them to use on the data. And the accuracy of the data is really on their nickel.

Mr. CLAY. Doctor, along these lines, you indicate in your testimony that you want to operate a system that is consistent with the spirit not the letter of the Privacy Act.

Mr. TETHER. That is correct.

Mr. CLAY. I don't quite understand how you are going to do that. Fundamental to the Privacy Act is the right of individuals to see the information held about them, and to correct that information.

Since DARPA will be using data held in the private sector, it cannot give individuals the right of access and correction, to say that you are going to comply with the Privacy Act seems misleading.

How will you address this fundamental conflict?

Mr. TETHER. Well, the fundamental conflict again is going back to what we do. We don't collect any data. We are not the people that collect data. We are the people that supply the analytical tools to those who collect the data. Now, we do worry about privacy. And, therefore, we are providing technology so that the people in the different agencies, as they share data with each other, can be comfortable that their sources and methods are not being compromised.

We are developing new technology for doing that. For example, one of the major efforts we have in this privacy technology, we have Oracle as a contractor. Why Oracle? Well, as you know they are a great data base company. And, in fact, it is the only way we will get the transitioning of that technology is to have a company like that develop the technology and be comfortable with its use, to pull it in.

But we really are not the ones that collect the data. And I really want to make that very clear. We are the ones that supply the tools to the people that collect the data, who by the way are operating in full cognizance of all of the regulations.

Mr. CLAY. Thank you for that.

You, know, you talked in your testimony about us knowing the hijackers, knowing they were here 2 days after September 11, having known Attorney General Ashcroft for the last 20 years as my State AG, Governor and U.S. Senator, he is a guy that can't keep a job either, but he told me that we knew that the hijackers were in this country, we just had no real good system of tracking them.

Now, we'll be able to track them under this new technology?

Mr. TETHER. We are developing a technology, which will allow the agencies involved to easily collaborate with each other and to be able to ask questions and create a model.

For example, if one of the models was, and we knew that using an aircraft as a weapon was something on their minds, because of the Philippines in 1995. So that was a known technique. Now, imagine if we had a model created about how would somebody go and use an airplane as a weapon? And we had created all of the indicators that had to be true.

Well, one would say, well, it is going to be hard to have somebody have a pilot who is being paid to fly that airplane actually fly it into a building. So that means that someone is going to probably have to take over the plane and learn how to fly it.

So one question would be, do we have any unusual people learning aviation? OK. Imagine if we had a collaborative system like we are talking about, and there was a set of questions up there, and one of them is, has anybody come across anybody wanting to learn how to fly aircraft under unusual circumstances, like maybe they don't care if they land?

If that system had been in effect, that question in that memo that the FBI talked about, would have popped up, and people would have looked at it and looked further into it. I really believe

that. What we are doing is providing the technology to allow that to happen, not the collection of the data itself.

Mr. CLAY. Thank you, Doctor, for your answers. Thank you.

Mr. PUTNAM. The gentlelady from Michigan.

Mrs. MILLER OF MICHIGAN. Thank you, Mr. Chairman.

Dr. Tether, I think that you may have just answered my question here. I am listening to you talk about modeling and models, and so as opposed to doing sort of broad surveillance, as you are doing the construct for your models and using some of those things for investigative work, then you say you do a model and you find a pattern, then you match the pattern against the data bases.

I was going to ask you for an example, a pattern of what. Could you give us a pattern of what? And perhaps you just answered that, by talking about someone going to a flight school and learn how to fly an airplane, didn't care whether or not he landed. That is a pattern of odd behavior.

Mr. TETHER. That is a pattern of odd behavior that would come out if you had a center which people from different agencies were able to collaborate and actually address various questions. Now, we used to do this in the cold war in Germany in Stuttgart, there was a center where there was like 250 questions that were developed over the years for what would the state of the Soviet Union be if it were going to attack us?

And there were a whole bunch of questions. Some of them were, where is the Soviet leadership? Are they still publicly visible? The CIA could go over with their HUMINT people and find out the answer to that question and come back and say they are visible, yes we know where they are, and never have to disclose their source. And that is what I meant about having the technique to protect the sources and methods.

But, we have had that system in the past. And it has worked really well. It is just never been applied here. We firmly believe that, first of all, all of our agencies have these stovepipes. These are really based on culture. We don't want to get rid of the stovepipes. I don't think anybody wants to either, because there is a certain value to having a culture.

But, what we are trying to do is develop the technology so those stovepipes can be punctured full of holes. While the culture remains, it allows people to cross communicate in a very easy way. That is really all TIA is about. The data mining part is really the easiest part. It is all of the going before, the collaboration technology and coming up with the models, which you then take the pattern to the data base.

That part, while really not trivial, is really the easy part. It is really the before part that is what we are trying to do. I hope that helps.

Mrs. MILLER OF MICHIGAN. Yes. Thank you.

Thank you, Mr. Chairman.

Mr. PUTNAM. You mentioned that there was no transactional data contemplated, that was in the media, but that actually was not contemplated by DARPA?

Mr. TETHER. Except in a research way to assess what might be a sweet spot of new data that is not currently being collected by our intelligence and counterintelligence organizations.

Mr. PUTNAM. You identified transportation data.

Mr. TETHER. I probably will regret that. But that seems to be a major part to the puzzle, because of the fact that these terrorists don't have their own Air Forces or Navy and have to rely upon commercial capabilities.

Mr. PUTNAM. But you are stating here, on the record, that you don't contemplate monitoring credit card transactions, library card check-outs, video rentals.

Mr. TETHER. Contemplating is the word I am having trouble with, because we have a lot of researchers who are out there looking at what data might be useful.

I personally would be extraordinarily surprised if video rentals, and such, would be some of that data. The only person I know of who has said that credit card data monitoring would be a good thing to do, actually was President Clinton at an address at UC Davis in the spring of 2002, because he talked about some facts about the hijackers with their credit cards being maxed out, and so forth and so on, which might have tipped you off that something is wrong about this person who has six credit cards that are all maxed out.

But that is sort of a hindsight type of input, that I don't really know if it would be valuable in doing the predictive something.

Mr. PUTNAM. I am not trying to back you in a corner. I am trying to help you guys fix a little bit of your PR problem.

Mr. TETHER. I am trying to contemplate.

Mr. PUTNAM. We pay you to contemplate.

Mr. TETHER. You pay us to contemplate.

Mr. PUTNAM. So there is no plans to deploy?

Mr. TETHER. There are no plans to deploy. Contemplate, is there a researcher someplace that we are paying at some university who is thinking.

Mr. PUTNAM. Again, I am not trying to back you in a corner. I am trying to get out there what is not occurring.

Mr. TETHER. Thank you.

Mr. PUTNAM. I will stop there in the interests of time, and let's go ahead and reseat everyone. Let's take a 2-minute recess and allow everyone to come forward.

[Recess.]

Mr. PUTNAM. Let's go ahead and reconvene. I want to thank Admiral Loy and Mr. McCraw for remaining with the committee. We deliberately gave everyone their own panel to focus on the unique aspects of each of those systems, which I think is an important distinction to draw, because they are unique and have different purposes.

But, under the umbrella of general factual data analysis or data mining, I think it also makes some sense to bring everyone back and talk about collaboration and some of the other issues that are common to all three departments or agencies.

So we will, in keeping with our tradition, allow Mr. Clay to lead off with questions for the entire panel.

Mr. CLAY. Thank you, Mr. Chairman.

There is a dark cloud of secrecy that hangs over this administration. And the programs we are discussing today are a part of the reason that cloud is so dark. These agencies are creating surveil-

lance systems, and they don't want to tell the American people how they work. Just this week U.S. News reported that there are a number of children among the detainees at Guantanamo. How many we don't know, but we do know that the Secretary of State has objected to the situation.

TSA wants to develop a profiling system and doesn't want to allow the public to see or correct the information that is used to profile them. The FBI has decided that it no longer has to worry about the accuracy of the information it holds on people. DARPA wants to build yet another system to profile the American public.

Security at any cost is not what either Congress or the American public wants. Systems that skirt the Privacy Act or try to wiggle under the fourth amendment are not in our best interest. As we saw, when the public learned about the TIA program at DARPA, they were outraged, and Congress put a hold on the funding for that program.

I would like to ask each of you to explain to the subcommittee what you are doing to explain to the public what you are doing, and why you must keep it a secret.

Admiral LOY. Mr. Clay, thank you very much for the opportunity. First of all, TSA has filed a Federal Register notice, probably with as wide a set of opportunities for comment as has ever been published. We have reached out deliberately to a large spectrum of the privacy community, those folks who would be the first to support the fourth amendment.

We intend to continue to hold that ongoing conversation by holding public meetings about exactly what CAPPS II is all about. We are absolutely not putting together a profiling system that we will be ashamed of or fear offering to the public for comment. In fact, we are going desperately quickly in the other direction, to give every opportunity to the public, through public meetings, and through publicly stated records, with every intention of refiling a public notice when we have heard that conversation carefully and forged the privacy strategy associated with CAPPS II.

So I take issue with the kind of question that seems shaped around a conclusion, when, in fact, the conversation has not even been continuously held yet.

Mr. CLAY. Well, Admiral, I shape it that way because in our last hearing, Mark Forman, the Federal CIO, indicated that TSA was not meeting its deadlines for providing information to OMB. That was the end of March. Has TSA provided OMB with all of the information that OMB has requested about CAPPS II since that hearing?

Admiral LOY. Indeed we have, sir. What Mr. Forman is responsible for is making sure that the business case end of CAPPS II complies, not only with the procedures appropriate in any instance, but those that are sensitive in a project like you and I are speaking about.

Indeed we have provided that information to OMB.

Mr. CLAY. Thank you. Dr. Tether.

Mr. TETHER. Well, first of all, DARPA is not developing a system to profile the American public. And I hope, maybe I have convinced you in our previous conversation and the testimony.

Nothing could really be further from the truth. I mean, we aren't doing that at all. As far as secrecy, I can't think of anybody that has been more open about what we are doing. It is because of the nature of us. In order to develop this technology, we reach out for everybody that we can. Consequently, we advertise what we are doing. In this particular area, all of what we are doing is really unclassified.

Now, the application of it and the use of it on the data, the data becomes classified. But the actual technology and all is unclassified. We had a conference last year. We had 2,000 people there. We disclosed fully what we were doing. The press was there. We even here in the House last year, in the House Authorization Report, what we were doing was lauded by the report. As, "Hey, this is really great stuff."

Now, so it hasn't been kept quiet. On the other hand, because of the deluge of comments that happened last fall, I knew what was going on at DARPA, I was reading the papers. I mean, my God, I finally called up my guys, and asked if they were doing what I thought they were doing? I really became worried myself.

I can imagine that you all, with the inputs that you were getting from your constituents, I know about them because a lot of them have been forwarded to me to answer, that you must have been worried too.

Some of them were so outrageous we were stunned, quite frankly. The one mistake that we at DARPA made is that we were so stunned by the outrageous comments that we didn't do anything about it for some time. We just watched it. And finally we woke up and got over here and tried to get the truth out.

Now, on the other hand the Department of Defense, the Honorable Pete Aldridge, who I work for, said, "Look, I know that what you are doing is OK, Tony. But I am going to do two things to make sure that everybody else believes that too." First of all, two boards have been created in the Department of Defense, an internal board and an external board. The internal board is really set up to really just review the DARPA TIA project. OK, to make sure that the DARPA TIA project is truly following all of the rules and regulations and so forth and so on.

And we have certain outputs already from our internal Inspector General that says what we are doing is in compliance with all of the laws and regulations. On the other hand, they have also established an external board. Now, this external board is comprised of people from the outside. I didn't give you a list in my testimony, but I would be happy to do that, of people on the outside, some are privacy advocates, but all are well known, and understand the Constitution and the laws and regulations.

And the purpose of that is not just for us. But you see in the Department, if you want to do experiments on animals, you go to an animal board. You say, look this is an experiment using animals. Can I do it? And the board says yes or no. If you want to use humans, same thing. What we don't have is a board that really does privacy things.

If you go to the board and say look, here is a new information technique that I want to use. You know, can I do it? Is it violating any laws and regulations? This external board is supposed to be a

board that people like us can go to and get someone to say yes or no on what we are trying to do. But we really have been open.

We are not developing a system to profile the American public.

Mr. CLAY. Thank you.

Mr. McCraw.

Mr. MCCRAW. Yes, sir, Congressman.

First, I would like to say, you know, a core value of the FBI is a rigorous obedience to the Constitution and the rule of law. And that there is no justification at all when the FBI is in search of or investigating or trying to enforce the law, violating the law. We must adhere to all proscribed laws by Congress, the attorney general guidelines.

In maintaining, you know, the data in our systems, we need to apply to those laws and we will. And whatever laws and whatever rules that are applicable, the internal guidelines, the FBI will abide by those.

Mr. CLAY. Let me say in summary, Mr. Chairman, to the witnesses, that TSA won't tell us how CAPPS II works or what data goes into it.

DARPA wants to keep the same kinds of secrets. The FBI issued a rule saying that it doesn't have to keep its records on Americans accurate and up to date. I would like each of you to go back to your respective agencies and figure out what you can do to help build public confidence in your activities through openness and then report back to this committee and tell us what you think would make this process a little easier on the American public.

Mr. MCCRAW. If I may, I would like to go back and research the rule that you are referring to, because, clearly, I need to. I will do that. And I will get back to you.

Mr. CLAY. Get back to us. I would appreciate it. Thank you.

Mr. PUTNAM. Thank you, Mr. Clay. And if you have additional questions at any time, just let me know.

I really sympathize with the position that all three of you are in. Prior to September 11th, we would have town hall meetings and letters and phone calls where people were unhappy about red light cameras. And then people were unhappy when Tampa hosted the Super Bowl and they deployed a new face recognition technology that scanned the crowd and detected people who may be terrorists or who may be criminals. And people were up in arms. It was a very hot constituent issue.

On September 12th, people were up in arms that we had not done more to surveil potential bad people, to track potentially bad people, to keep tabs on them or to prevent their entry into this country and so forth and so on.

We are plowing new ground here. We are taking advantage of wonderful new technologies that offer the hope of greater national security, and the threat of greater intrusiveness into innocent lives. And we hold Mr. McCraw and Admiral Loy on a regular basis responsible for all a lot of those things, for the safety of Americans, and yet all of us as policymakers in the executive branch and the legislative branch and the judiciary, are all trying to find this new line to go along with this new technology.

So we have this obligation to find this balance and to take the promise of the technology, without having inaccuracies, cause good

people to be held, detained or held in suspicion or cause to disrupt their lives, whether it is as simple as missing a flight or as serious as being detained and questioned and perhaps even held for some charge that is false.

So that is what we are attempting to do with this Phase 1 of these hearings is exercise our responsibility to weigh in on where these lines are. All of these events, all of these terrorist incidents have highlighted new weaknesses and new gaps in our ability as a country to detect or prevent them.

And one of them is really a lack of collaboration, it is not a technological challenge, it is a human capital challenge, and a couple of you have referred to that. You are a creature of the FBI culture, a very proud, rich tradition in law enforcement.

You are a DOD animal, a creature of that community. And Admiral Loy, you and your guys are still figuring out your culture.

Admiral LOY. Well, I came with one from 200 years of culture.

Mr. PUTNAM. You certainly have a rich tradition from the Coast Guard, and, hopefully, that will influence the spotted zebra that we have created at Homeland Security. But, how can each of you address the challenge of overcoming the cultural obstacles to utilize this technology, to share the data, to break down the mistrust, and have this system, these systems when they are deployed in an efficient way, have them truly collaborative across the Federal Government and including State and local law enforcement, beginning with Admiral Loy.

Admiral LOY. Sir, it is an enormous challenge. If you go back to even a founding father, the balance was being discussed even then. Franklin said, "He who would give up even a moment of liberty for hours of safety deserves neither."

Our challenge is to build the culture you were just describing. It is to take a set of core values, I use the phrase that was just used a moment ago, and in our case at TSA, to bring together a composite of people from so many different walks of life, not only the Federal Government, but from throughout America as well, and compose a culture. That culture would be based on the Constitution that we all hold so dear, and we would be reminded on a daily basis that what was attacked on September 11, 2001 was not just the World Trade Center, it was not just the Pentagon, but it was the whole notion, idea and ethic of what America is all about. We should be in the business in a post-September 11 environment of designing systems that recognize that very Constitutional foundation that we are so enormously proud of as Americans.

I also, at this particular point in my 1-year tutelage of this organization, have come to conclude that the Founding Fathers, also had mobility as one of the inalienable rights they were talking about. My challenge is to protect that particular notion for Americans who want to get up and go, wherever they want to go, with some security and comfort. This includes not only being able to get to their destination, but any system that will be imposed on them for either their benefit for security or for their benefit for safety will be such that they don't feel they have given up something in the way of privacy in order to be comfortable in the way of security.

That is the challenge that we have on a daily basis; you're absolutely right. It's the execution of those things, day after day after

day, and the aggregation of that which will represent the answer at the other end of the day as to whether we, in our generation, had our chance to get it right and did so. That calls for a commitment to the very basic precepts that have always guided legislators, judges and operators in the executive branch, to meet the needs of Americans.

Mr. TETHER. First of all, I think you're absolutely right, collaboration is the key. And that is really the technology that really needs to be developed that will allow people like the three of us to be able to remain physically where our particular offices are, but yet be able to go into a "room" and have no constraints on being able to talk to each other like we're at this table right now. We're working hard on developing that collaboration technology which we believe in and we're experimenting with it. We have experimental nodes at several places within the DOD, INSCOM, overseas, SOCOM, STRATCOM, JFCOM, these are all DOD locations. We are putting in these experimental nodes where the collaboration technology and other tools are being tried. I really agree with you.

Now how do we make people comfortable with what we're doing? Well, hearings like this. This is really the right way to do it. I don't know how else to do it except talk about what we're doing. To talk about it where people will listen. I'm sure there's press here in the room. And you know, this is the way to do it. Hearings like this, congressional oversight.

As you know, we had a major flap, over the last year, and a lot of good came out of that. Because we weren't doing anything wrong. But now I believe that with the report that's coming out and the conversations that we've had, a lot of other people are understanding what we're doing. I think they're starting to be more comfortable with it. At least I'm not getting as many letters as I was. But collaboration is the tool, the coin of the realm. And meetings like this and hearings like this are the way we're going to make people in the United States comfortable. That's all I have.

Mr. PUTNAM. Mr. McCraw.

Mr. MCCRAW. Yes, sir. First it gives me another chance an opportunity to thank you. For example, the Patriot Act, now enables us to share data with other agencies that we could not up until the passing of the Patriot Act, as well as certain modifications in terms of the Attorney General guidelines. I know it's been talked about a great deal about the cultures within agencies not sharing information, protecting information. And I'm not going to say that the FBI has had the best reputation in the past in terms of sharing information, but we've all worked with the American public and we've all taken an oath. And one core value we all share is we're all willing to sacrifice for our country. Like you, we all love our country.

We cannot afford to have something happen like September 11th, because we didn't share information as a result of parochial reasons. And, frankly, with technology and the type of things that the good doctor is working on, clearly it enables us, I think, collectively as a community, to leverage that technology so that we can find those links, relationships, and do what all of us want to do, and every American citizen wants to do and that's protect the United States from a future terrorist attack.

Mr. PUTNAM. Let me change gears just a little bit with you, Mr. McCraw, because Admiral Loy and I had sort of an extended dialog over the long-term potential of the CAPPS technology. It is difficult for me to believe that at some point in the future, when hopefully the national tension has declined a bit over international terror and it has returned to good old fashioned crime, it's difficult to believe that we wouldn't take that technology and use it as a tool to capture someone who had kidnapped a minor, was attempting to bring them back across different States on an airline, essentially an Amber Alert plus, it's difficult to believe that you wouldn't take that same technology that's already paid for, that's already in place, that's already omnipresent and use it to capture felons, drug dealers, perhaps even a step further, people who owe child support, it's limitless when you already have that type of technology in place.

So as a law enforcement officer and former field agent, wouldn't the temptation be you would want to go to the parents of someone whose child has been kidnapped and say we've sent out the alert; there's no station, no airport, no cruise terminal anywhere, no rental car agency, that they could go without being picked up?

Mr. MCCRAW. Well, you know, clearly, we're excited about the technology because it doesn't just make us more effective in terrorism, you're right, it also makes us more effective with our own data in the internal data we collect when we're trying to investigate corporate fraud, gang-related violence, the sniper case, all of these particular things will enable us, whatever the agent is investigating, and certainly to the degree that there's things in place, technology that's out there that we can employ over other data sets, as long as it's legally permissible by law, by statute and attorney general guidelines then clearly we would love to leverage that and will leverage that not being unlike we use some of the technologies out there in the private sector today and information when we are using some of the public source data base that we talked about.

Mr. PUTNAM. Admiral Loy, is there anything you want to add to that?

Admiral LOY. Yes, sir. I will just add that is what the notion of oversight is all about, both the general congressional oversight of what the executive branch is up to, checks and balances in the classic sense, but more appropriately, an oversight group not only about the Congress but to include folks who have as a raison d'etre in life, the fourth amendment, to have them be part of that system. So Mission Creep is not part of what occurs in a program without the attention very carefully and very publicly being brought to the attention of whoever is responsible for that program.

So if it's me responsible for CAPPS II 5 years from now when what you just described has occurred, I want my oversight panel to be composed in such a fashion that as soon as there is the appearance of Mission Creep in the CAPPS II system, someone will very clearly bring that to my attention, and, as appropriate, bring it to your attention.

Mr. PUTNAM. Dr. Tether, is it your opinion that the appropriate deployment of CAPPS II and Trilogy will substantially improve national security?

Mr. TETHER. Boy, that's a good question. The word is substantially. I believe I do know about CAPPS II. We were involved with the source election in CAPPS II, so I know about it. And I also do know something about Trilogy, because we were involved in talking with the FBI also about providing them some capability. We're prohibited by the way from providing either of these folks these tools that we're developing.

And I would say yes. I believe that. Especially if they have the right tools. I'm hesitating because I'm not really sure what tools they're going to have. I really think the fury over what we're doing hasn't been by the people. The privacy people are well-intentioned, but they know that we're not the problem. It's the people who are going to use the tools and the people who are going to authorize the use of the tools that are the problem. But if they can stop the tools from being built, they don't have to worry about the other people. I really believe that's the reason for the fury over what we're doing. But the tools are important. These folks could use the tools. And I believe that they will. If he does his job right, with the right tools, security will definitely be greatly enhanced.

Mr. PUTNAM. Mr. McCraw and Admiral Loy, will both of your programs, and Dr. Tether, as part of your research, does your program envision a mechanism for the public to have redress of inaccurate or incomplete information that would make them whole essentially for whatever harm is done?

Admiral LOY. Absolutely sir. The passenger advocate role that I mentioned in my testimony is being designed to offer that redress opportunity for anyone aggrieved by the system.

Mr. TETHER. Same here. What we're doing is developing this audit technology so that the data, the information can only be used for the purpose for which it was granted and not abused by being taken out of that environment and used elsewhere. That audit technology again is very crucial to all of this. If it works well, the public or anyone can be assured that the data will only be used for what it was authorized to be used for, and not anything else.

Mr. PUTNAM. Is it limited in its effectiveness to TIA, or does it have potential for these other——

Mr. TETHER. Oh, it's very general. In fact, it needs to be general because the data bases are not homogenous, they're heterogenous data bases so it has to be general.

Mr. PUTNAM. Anybody can pull up their credit reports and find mistakes and have an opportunity——

Mr. TETHER. In our case, only if a person is authorized to see the credit report can he pull up the credit report. But then that credit report cannot be used for other purposes, without the audit technology either preventing it or raising an alarm that it's being done. That's what I mean.

Admiral LOY. We're using, sir, a software system called Radiant Mercury, which is an NSA-accredited—literally, the top accreditation, if you will, that you can get in this regard. It's essentially a real-time logging system that will facilitate tracking for everybody that needs to understand access to the system that has been encountered. It will be a literally a real-time logging of who went in, got what, for what purpose. That's one of the auditing systems that we're using in the CAPPS II program.

Mr. PUTNAM. Radiant Mercury?

Admiral LOY. Radiant Mercury, a proprietary system by Lockheed Martin.

Mr. PUTNAM. Only in homeland security?

Admiral LOY. No, sir. That's software that any program manager who needs it for an auditing program or auditing software piece can either purchase or lease as appropriate from its proprietary owner.

Mr. MCCRAW. Auditing is a very important function with us as well, including some of the risk technology that the doctor talked about. In fact, using that, we will definitely improve within the FBI information, who has a right to have access. Are there patterns of activity being done by an agent that would signal a concern in terms of the access? Also, we're even looking at using the same type of technology to provide greater oversight in terms of operations of sources. Because that pattern technology is there and how we can do a better job of using technology, building into Trilogy, things that can serve as triggers, as alarms of activity that we need to take a closer look at.

Mr. PUTNAM. Dr. Tether, are you aware of other nation's approaches to data mining and where they are in their sophistication, and are there other customers in the Federal Government who have this level of data mining technology or greater that may be employed?

Mr. TETHER. You mean the technology that we're trying to develop?

Mr. PUTNAM. Well, you've got three different ones represented. I'm just curious how many others there may be in the rest of the community or in other agencies that we're not as familiar with.

Mr. TETHER. With respect to other nations, I don't know of any. That may be just ignorance on my part. But I don't know of any that have technology that is anything like what we're talking about here. With respect to other Federal agencies, NSA obviously has a concern of people who have access to data that may have audit capability there. They're worried about making sure that data on U.S. persons stays in special compartments. CIA undoubtedly has the same with respect to their human capabilities. Other than that I don't think so. I mean, other than that, I think you've, for better or worse, you've got it here at the table.

Mr. PUTNAM. Admiral Loy, any comments on that?

Admiral LOY. I don't have any to add, simply because I'm not sure of what others might be. There is one point that you asked us to comment on earlier, Mr. Chairman. I would be remiss if I didn't add that the whole purpose of DHS, as I understand the legislation, was to facilitate the notion of information sharing, and that is exactly what's going on under Secretary Ridge. He is wedded clearly to the notion of the value in pulling together all those disparate agencies under one roof.

For example, part of the review process for allowing CAPPS II to go on and making budgetary judgments with respect to building out our facility out at Annapolis Junction was to make absolutely certain that it took into acocunt other programs in the Bureau, the old Customs service, now the BCP Bureau under BTS or BICE, the new Bureau that used to be fundamentally INS under BTS, and

whether there are screening projects going on in any of those organizations. Secretary Ridge and Under Secretary Hutchinson are making absolutely certain that we are present together in the same room and we're not allowed out until the notion of redundancies and overlaps and such are eliminated at the design stage and dealt with in a fashion that was the intention, I think, of the Congress when DHS was put together.

Mr. PUTNAM. And you're satisfied that information sharing is occurring between the intelligence community?

Admiral LOY. I still think we have a way to go. There's no doubt about that. But it is so much better than it was. We all had this quest for connecting the dots after September 11, 2001, and as I think we heard from Dr. Tether the potential to connect those dots is infinitely better today than it was then, sir.

Mr. PUTNAM. With regard to CAPPS II and the TSA, are we ahead of everybody else in the world? Do the Israelis have a model or someone else out there?

Admiral LOY. Actually I'm going over to Israel at the end of the month with several purposes in mind, among which are to check the passenger prescreening process that they use there to make sure that we are learning about whatever they may have that they haven't shared with us already.

Mr. PUTNAM. Very good.

Dr. Tether.

Mr. TETHER. This is ignorance.

Admiral LOY. Is Customs part of TSA?

Mr. TETHER. It's part of homeland security. Customs has a very good system for looking at imports coming in to be able to determine whether a cargo ship or cargo is suspect to know which ships to go and further inspect. And they have a very good system in place.

Admiral LOY. If I may, Mr. Chairman, I have been personally writing about and convinced for several years that the point of origin, the point of destination, with transparency in between is the real key to our understanding the cargo piece, perhaps as well as the passenger piece, in terms of those who would be aboard or even the crews. But properly manifested cargo and people on ships at sea coming toward the United States is an enormous challenge for us to deal with. This goes to your earlier question about the rest of the transportation system not being left behind. As Americans, we have a pretty good penchant to continue to fight last year's war.

Mr. PUTNAM. It's important that the rest of the world join us with all the ports that are now international and the tremendous volume of cargo.

Admiral LOY. Commissioner Bonner has really done an excellent job in his outreach to ports around the world, in his so-called CSI initiative, which basically has convinced about 25 of the major ports of the world who send things to the United States to be in a reciprocity agreement to allowing Customs, U.S. BCP inspectors, portal inspectors to be literally on the pier when things are being loaded on the ship as opposed to when they're already here. It's too late if they're already here.

Mr. PUTNAM. Admiral, do you have any final comments for the subcommittee?

Admiral LOY. Only to thank you for the opportunity as offered in the letter of invitation to offer to the committee some insights as to the intentions of where we're going and how we're going to do it. I appreciate that opportunity, sir, and look forward to working with you in the future.

Mr. PUTNAM. Dr. Tether, anything from DARPA?

Mr. TETHER. No. Thank you very much. I do believe that hearings like this are the way to get the public comfortable, and thank you for having it. Thank you for having me.

Mr. PUTNAM. Director McCraw.

Mr. MCCRAW. There's only one director. If I want to keep my job, I better make that clear. The only thing I know it was to the culture question that has been echoed is that we recognize the FBI's unique collections of data, and we have to be able to share that data to the widest extent possible by law and also leveraging information technology that the good doctor has been working on over at DARPA. So thank you very much for the time of being here.

Mr. PUTNAM. Thank you. I thank all of our distinguished witnesses for their participation today. I'm grateful for their cooperation with Congress. We're particularly grateful at your efforts to get in your written testimony. Your cooperation is key as we continue to deal with these important issues facing the Nation and the Congress. I want to thank the subcommittee members for their participation, particularly our ranking member. I appreciate everything that you've done. And in the event that there may be additional questions we did not reach today, the record shall remain open for 2 weeks for submitted questions and answers. Thank you all. The subcommittee stands adjourned.

[Whereupon, at 5:21 p.m., the subcommittee was adjourned.]