

GAO

Report to the Chairman, Committee on
Appropriations, House of
Representatives

February 2007

DATA MINING

Early Attention to Privacy in Developing a Key DHS Program Could Reduce Risks



GAO Accountability Integrity Reliability Highlights

Highlights of GAO-07-293, a report to the Chairman, Committee on Appropriations, House of Representatives

DATA MINING

Early Attention to Privacy in Developing a Key DHS Program Could Reduce Risks

Why GAO Did This Study

The government's interest in using technology to detect terrorism and other threats has led to increased use of data mining. A technique for extracting useful information from large volumes of data, data mining offers potential benefits but also raises privacy concerns when the data include personal information.

GAO was asked to review the development by the Department of Homeland Security (DHS) of a data mining tool known as ADVISE (Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement). Specifically, GAO was asked to determine (1) the tool's planned capabilities, uses, and associated benefits and (2) whether potential privacy issues could arise from using it to process personal information and how DHS has addressed any such issues. GAO reviewed program documentation and discussed these issues with DHS officials.

What GAO Recommends

To ensure that privacy protections are in place, GAO is recommending that the Secretary of Homeland Security immediately conduct a privacy impact assessment of the ADVISE tool and implement privacy controls, as needed, to mitigate any identified risks.

DHS generally agreed with the content of this report and described actions initiated to address GAO's recommendations.

www.gao.gov/cgi-bin/getrpt?GAO-07-293.

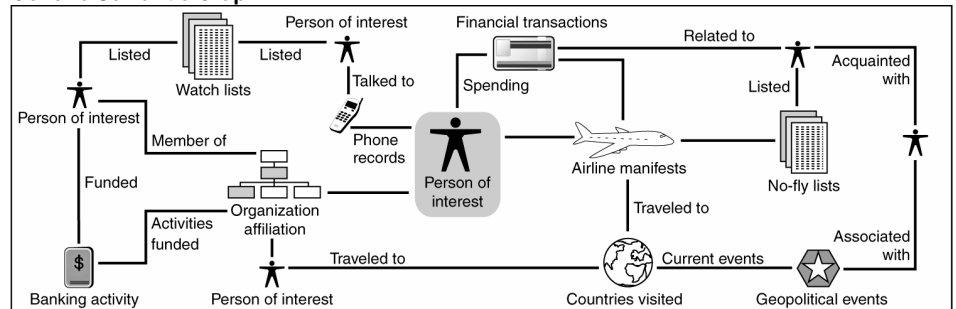
To view the full product, including the scope and methodology, click on the link above. For more information, contact Linda Koontz at (202) 512-6240 or koontzl@gao.gov.

What GAO Found

ADVISE is a data mining tool under development intended to help DHS analyze large amounts of information. It is designed to allow an analyst to search for patterns in data—such as relationships among people, organizations, and events—and to produce visual representations of these patterns, referred to as semantic graphs (see fig.). None of the three planned DHS implementations of ADVISE that GAO reviewed are fully operational. (GAO did not review uses of the tool by the DHS Office of Intelligence and Analysis.) The intended benefit of the ADVISE tool is to help detect threatening activities by facilitating the analysis of large amounts of data. DHS is currently in the process of testing the tool's effectiveness.

Use of the ADVISE tool raises a number of privacy concerns. DHS has added security controls to the tool; however, it has not assessed privacy risks. Privacy risks that could apply to ADVISE include the potential for erroneous association of individuals with crime or terrorism and the misidentification of individuals with similar names. A privacy impact assessment would identify specific privacy risks and help officials determine what controls are needed to mitigate those risks. ADVISE has not undergone such an assessment because DHS officials believe it is not needed given that the tool itself does not contain personal data. However, the tool's intended uses include applications involving personal data, and the E-Government Act and related guidance emphasize the need to assess privacy risks early in systems development. Further, if an assessment were conducted and privacy risks identified, a number of controls could be built into the tool to mitigate those risks. For example, controls could be implemented to ensure that personal information is used only for a specified purpose or compatible purposes, and they could provide the capability to distinguish among individuals that have similar names to address the risk of misidentification. Because privacy has not been assessed and mitigating controls have not been implemented, DHS faces the risk that ADVISE-based system implementations containing personal information may require costly and potentially duplicative retrofitting at a later date to add the needed controls.

Generic Semantic Graph



Source: GAO.

Contents

Letter		1
	Results in Brief	2
	Background	4
	ADVISE Is Intended to Help Identify Patterns of Interest to Homeland Security Analysts	13
	DHS Has Not Yet Addressed Key Privacy Risks Associated with Expected Uses of the ADVISE Tool	18
	Conclusions	23
	Recommendations for Executive Action	23
	Agency Comments and Our Evaluation	24
Appendix I	Objectives, Scope, and Methodology	26
Appendix II	Comments from the Department of Homeland Security	27
Appendix III	GAO Contact and Staff Acknowledgments	30
Table		
	Table 1: Fair Information Practices	12
Figures		
	Figure 1: An Overview of the Data Mining Process	5
	Figure 2: Major Elements and Functions of ADVISE	14
	Figure 3: Typical Semantic Graph	16

Abbreviations

ADVISE	Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement
DHS	Department of Homeland Security
ICAHST	Interagency Center for Applied Homeland Security Technology
OECD	Organization for Economic Cooperation and Development
OMB	Office of Management and Budget
PIA	privacy impact assessment

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, DC 20548

February 28, 2007

The Honorable David R. Obey
Chairman, Committee on Appropriations
House of Representatives

Dear Mr. Chairman:

Since the terrorist attacks of September 11, 2001, there has been an increasing focus on the need to prevent and detect terrorist threats through technological means. Data mining—a technique for extracting useful information from large volumes of data—is one type of analysis that has been used increasingly by the government to help detect terrorist threats. While data mining offers a number of promising benefits, its use also raises privacy concerns when the data include personal information.¹

Federal agency use of personal information is governed primarily by the Privacy Act of 1974 and the E-Government Act of 2002, which prescribe specific activities that agencies must perform to protect privacy, such as (1) ensuring that personal information is used only for a specified purpose, or related purposes, and that it is accurate for those purposes and (2) conducting assessments of privacy risks associated with information technology used to process personal information, known as privacy impact assessments.² Agencies that wish to reap the potential benefits of data mining are faced with the challenge of implementing adequate privacy controls for the systems that they use to perform these analyses.

You asked us to review the Department of Homeland Security's (DHS) development of an analytical tool known as Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE). Specifically, we agreed with your staff that our objectives were to determine (1) the planned capabilities, uses, and associated benefits of the ADVISE tool and (2) whether potential privacy issues could arise from using ADVISE to

¹For purposes of this report, the term *personal information* encompasses all information associated with an individual, including both identifying and nonidentifying information. *Personally identifying information*, which can be used to locate or identify an individual, includes things such as names, aliases, and agency-assigned case numbers.

²A privacy impact assessment is an analysis of how personal information is collected, stored, shared, and managed in a federal system to ensure that privacy requirements are addressed.

process personal information and how DHS has addressed any such issues. Our review did not include intelligence applications, such as uses of the tool by the DHS Office of Intelligence and Analysis.

To address our first objective, we identified and analyzed the ADVISE tool's planned capabilities, uses, and associated benefits. We reviewed program documentation, including annual program execution plans, and interviewed agency officials responsible for managing and implementing the program. We also interviewed officials at DHS components that have begun to implement the tool³ in order to identify their current or planned uses, the progress of their implementation, and the benefits they hope to gain.

To address our second objective, we searched for potential privacy concerns by reviewing relevant reports, including prior GAO reports and the DHS Privacy Office 2006 report on data mining.⁴ We identified and analyzed actions to comply with the Privacy Act of 1974 and the E-Government Act of 2002. We also interviewed technical experts within the DHS Science and Technology Directorate and personnel responsible for implementing ADVISE at DHS components to assess privacy controls included in the ADVISE tool, as well as the quality assurance processes for data analyzed using ADVISE. We performed our work from June 2006 to December 2006 in the Washington, D.C., metropolitan area and Laurel, Maryland. Our work was performed in accordance with generally accepted government auditing standards. Our objectives, scope, and methodology are discussed in more detail in appendix I.

Results in Brief

ADVISE is a data mining tool under development that is intended to facilitate the analysis of large amounts of data. It is designed to accommodate both structured data (such as information in a database) and unstructured data (such as e-mail texts, reports, and news articles) and to allow an analyst to search for patterns in data, including relationships among entities (such as people, organizations, and events),

³These DHS components include Immigration and Customs Enforcement and other components. We also interviewed officials from the Interagency Center of Applied Homeland Security Technology, who are responsible for testing the tool's capabilities. ADVISE is also being used by the Office of Intelligence and Analysis. We did not review that application.

⁴DHS, *Data Mining Report: DHS Privacy Office Response to House Report 108-774* (July 6, 2006).

and to produce visual representations of these patterns, referred to as semantic graphs. Although none are fully operational, DHS's planned uses of this tool include implementations at four departmental components (including Immigration and Customs Enforcement and other components).⁵ DHS is also considering further deployments of ADVISE. The intended benefit of the ADVISE tool is to help detect activities that threaten the United States by facilitating the analysis of large amounts of data that otherwise would be very difficult to review. DHS is currently in the process of testing the tool's effectiveness.

Use of the ADVISE tool raises a number of privacy concerns. DHS has added security controls to the ADVISE tool, including access restrictions, authentication procedures, and security auditing capability. However, it has not assessed privacy risks. Privacy risks that could apply to ADVISE include the potential for erroneous association of individuals with crime or terrorism, the misidentification of individuals with similar names, and the use of data that were collected for other purposes. A privacy impact assessment would determine the specific privacy risks associated with ADVISE and help officials determine what controls are needed to mitigate those risks. Although DHS officials are considering conducting a modified version of such an assessment, the ADVISE tool has not yet been assessed because department officials believe it is not needed given that the ADVISE tool itself does not contain personal data. However, the tool's intended uses include applications involving personal information, and the E-Government Act, as well as related Office of Management and Budget and DHS guidance, emphasize the need to assess privacy risks early in systems development. Further, if a privacy impact assessment were conducted now and privacy risks identified, a number of controls exist that could be built into the tool to mitigate those risks. For example, controls could be implemented to ensure that personal information is used only for a specified purpose or compatible purposes, or they could provide the capability to distinguish among individuals that have similar names (a process known as disambiguation) to address the risk of misidentification. Because privacy risks such as these have not been assessed and decisions about mitigating controls have not been made, DHS faces the likelihood that ADVISE-based system implementations containing personal information may require costly and potentially duplicative retrofitting at a later date to add the needed privacy controls.

⁵ADVISE is also being used by the Office of Intelligence and Analysis. We did not review that application.

To ensure that privacy protections are in place before DHS proceeds with implementations of systems based on ADVISE, we are recommending that the Secretary of Homeland Security immediately conduct a privacy impact assessment of the ADVISE tool to identify privacy risks and implement privacy controls to mitigate those risks.

We obtained oral and written comments on a draft of this report from DHS. In its comments DHS generally agreed with the content of this report and described actions initiated to address our recommendations.

Background

As defined in a report that we issued in May 2004,⁶ data mining is the application of database technology and techniques—such as statistical analysis and modeling—to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the prediction of future results. This definition is based on the most commonly used terms found in a survey of the technical literature.

Data mining has been used successfully for a number of years in the private and public sectors in a broad range of applications. In the private sector, these applications include customer relationship management, market research, retail and supply chain analysis, medical analysis and diagnostics, financial analysis, and fraud detection. In the government, data mining has been used to detect financial fraud and abuse. For example, we used data mining to identify fraud and abuse in expedited assistance and other disbursements to Hurricane Katrina victims.⁷

Although the characteristics of data mining efforts can vary greatly, data mining generally incorporates three processes: data input, data analysis, and results output. In *data input*, data are collected in a central data “warehouse,” validated, and formatted for use in data mining. In the *data analysis* phase, data are typically queried to find records that match topics of interest. The two most common types of queries are pattern-based queries and subject-based queries:

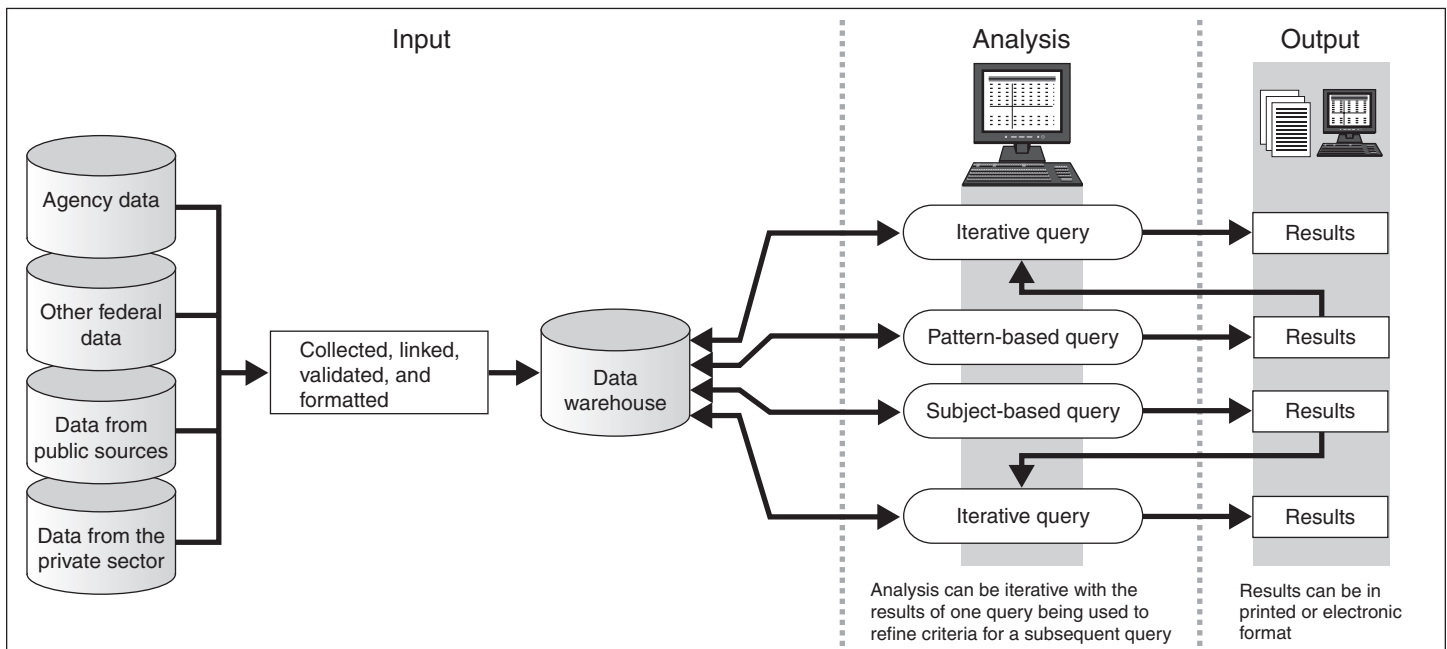
⁶GAO, *Data Mining: Federal Efforts Cover a Wide Range of Uses*, [GAO-04-548](#) (Washington, D.C.: May 4, 2004).

⁷GAO, *Expedited Assistance for Victims of Hurricane Katrina and Rita: FEMA’s Control Weaknesses Exposed the Government to Significant Fraud and Abuse*, [GAO-06-403T](#) (Washington, D.C.: Feb. 13, 2006).

- Pattern-based queries search for data elements that match or depart from a predetermined pattern (e.g., unusual claim patterns in an insurance program).
- Subject-based queries search for any available information on a predetermined subject using a specific identifier. This could be personal information such as an individual identifier (e.g., an individual's name or Social Security number) or an identifier for a specific object or location. For example, the Navy uses subject-based data mining to identify trends in the failure rate of parts used in its ships.

The data analysis phase can be iterative, with the results of one query being used to refine criteria for a subsequent query. The *output* phase can produce results in printed or electronic format. These reports can be accessed by agency personnel and can also be shared with personnel from other agencies. Figure 1 depicts a generic data mining process.

Figure 1: An Overview of the Data Mining Process



Sources: GAO, adapted from Vipin Kumar and Mohammed J. Zaki.

In recent years, data mining has emerged as a prevalent government mechanism for processing and analyzing large amounts of data. In our May 2004 report, we noted that 52 agencies were using or were planning to use

data mining in 199 cases, of which 68 were planned, and 131 were operational. Additionally, following the terrorist attacks of September 11, 2001, data mining has been used increasingly as a tool to help detect terrorist threats through the collection and analysis of public and private sector data. This may include tracking terrorist activities, including money transfers and communications, and tracking terrorists themselves through travel and immigration records. According to an August 2006 DHS Office of Inspector General survey of departmental data mining initiatives,⁸ DHS is using or developing 12 data mining programs, 9 of which are fully operational and 3 of which are still under development.

One such effort is the ADVISE technology program. Managed by the DHS Science and Technology Directorate,⁹ the ADVISE program is primarily responsible for (1) continuing to develop the ADVISE data mining tool and (2) promoting and supporting its implementation throughout DHS. According to program officials, it has spent approximately \$40 million to develop the tool since 2003.

To promote the possible implementation of the tool within DHS component organizations, program officials have made demonstrations (using unclassified data) to interested officials, highlighting the tool's planned capabilities and expected benefits. Program officials have established working relationships with component organizations that are considering adopting the tool, including detailing them staff (typically contractor-provided) to assist in the setup and customization of their ADVISE implementation and providing training for the analysts who are to use it.

Program officials project that implementation of the tool at a component organization should generally consist of six main phases and take approximately 12 to 18 months to complete. The six phases are as follows:

- preparing infrastructure and installing hardware and software;
- modeling information sources and loading data;

⁸DHS Office of Inspector General, *Survey of DHS Data Mining Activities* (August 2006).

⁹The mission of the Science and Technology Directorate is to act as the primary research and development arm of DHS, providing federal, state, and local officials with the technology and capabilities to protect the United States homeland.

-
- verifying and validating that loaded data are accurate and accessible;
 - training and familiarizing analysts and assisting in the development of initial research activities using visualization tools;
 - supporting analysts in identifying the best ways to use ADVISE for their problems, obtaining data, and developing ideas for further improvements; and
 - turning over deployment to the component organizations to maintain the system and its associated data feeds.

The program has also provided initial funding for the setup, customization, and pilot testing of implementations within components, under the assumption that when an implementation achieves operational status, the respective component will take over operations and maintenance costs. Program officials estimate that the tool's operations and maintenance costs will be approximately \$100,000 per year, per analyst. The program has also offered additional support to components implementing the tool, such as helping them develop privacy compliance documentation. According to DHS officials, the program has spent \$12.15 million of its \$40 million in support of several pilot projects and test implementations throughout the department.

Currently, the department's Interagency Center for Applied Homeland Security Technologies (ICAHST) group within the Science and Technology Directorate is testing the tool's effectiveness, adequacy, and cost-effectiveness as a data mining technology. ICAHST has completed preliminary testing of basic functionality and is currently in the process of testing the system's effectiveness, using mock data to test how well ADVISE identifies specified patterns of interest.

Privacy Concerns Have Been Raised Regarding Data Mining

The impact of computer systems on the ability of organizations to protect personal information was recognized as early as 1973, when a federal advisory committee on automated personal data systems observed that "The computer enables organizations to enlarge their data processing capacity substantially, while greatly facilitating access to recorded data, both within organizations and across boundaries that separate them." In addition, the committee concluded that "The net effect of computerization

is that it is becoming much easier for record-keeping systems to affect people than for people to affect record-keeping systems.”¹⁰

In May 2004, we reported that mining government and private databases containing personal information creates a range of privacy concerns.¹¹ Through data mining, agencies can quickly and efficiently obtain information on individuals or groups by searching large databases containing personal information aggregated from public and private records. Information can be developed about a specific individual or a group of individuals whose behavior or characteristics fit a specific pattern. The ease with which organizations can use automated systems to gather and analyze large amounts of previously isolated information raises concerns about the impact on personal privacy.

Further, we reported in August 2005¹² that although agencies responsible for certain data mining efforts took many of the key steps required by federal law and executive branch guidance for the protection of personal information, none followed all key procedures. Specifically, while three of the four agencies we reviewed had prepared privacy impact assessments (PIA)—assessments of privacy risks associated with information technology used to process personal information—for their data mining systems, none of them had completed a PIA that adequately addressed all applicable statutory requirements. We recommended that four agencies complete or revise PIAs for their systems to fully comply with applicable guidance. As of December 2006, three of the four agencies reported that they had taken action to complete or revise their PIAs.

Federal Laws and Guidance Define Steps to Protect Privacy of Personal Information

Federal law includes a number of separate statutes that provide privacy protections for information used for specific purposes or maintained by specific types of entities. The major requirements for the protection of personal privacy by federal agencies come from two laws, the Privacy Act of 1974 and the privacy provisions of the E-Government Act of 2002. The

¹⁰U.S. Department of Health, Education, and Welfare, *Records, Computers and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems* (Washington, D.C.: July 1973).

¹¹[GAO-04-548](#).

¹²GAO, *Data Mining: Agencies Have Taken Key Steps to Protect Privacy in Selected Efforts, but Significant Compliance Issues Remain*, [GAO-05-866](#) (Washington, D.C.: Aug. 15, 2005).

Office of Management and Budget (OMB) is tasked with providing guidance to agencies on how to implement the provisions of both laws and has done so, beginning with guidance on the Privacy Act, issued in 1975.

The Privacy Act places limitations on agencies' collection, disclosure, and use of personal information maintained in systems of records. The act describes a "record" as any item, collection, or grouping of information about an individual that is maintained by an agency and contains his or her name or another personal identifier. It also defines "system of records" as a group of records under the control of any agency from which information is retrieved by the name of the individual or by an individual identifier. The Privacy Act requires that when agencies establish or make changes to a system of records, they must notify the public through a "system of records notice": that is, a notice in the *Federal Register* identifying, among other things, the type of data collected, the types of individuals about whom information is collected, the intended "routine" uses of data, and procedures that individuals can use to review and correct personal information.¹³ In addition, the act requires agencies to publish in the *Federal Register* notice of any new or intended use of the information in the system, and provide an opportunity for interested persons to submit written data, views, or arguments to the agency.

Several provisions of the act require agencies to define and limit themselves to specific predefined purposes. For example, the act requires that to the greatest extent practicable, personal information should be collected directly from the subject individual when it may affect an individual's rights or benefits under a federal program. The act also requires that an agency inform individuals whom it asks to supply information of (1) the authority for soliciting the information and whether disclosure of such information is mandatory or voluntary; (2) the principal purposes for which the information is intended to be used; (3) the routine uses that may be made of the information; and (4) the effects on the individual, if any, of not providing the information. In addition, the act requires that each agency that maintains a system of records store only such information about an individual as is relevant and necessary to accomplish a purpose of the agency.

¹³Under the Privacy Act of 1974, the term "routine use" means (with respect to the disclosure of a record) the use of such a record for a purpose that is compatible with the purpose for which it was collected. 5 U.S.C. § 552a(a)(7).

Agencies are allowed to claim exemptions from some of the provisions of the Privacy Act if the records are used for certain purposes. For example, records compiled for criminal law enforcement purposes can be exempt from a number of provisions, including (1) the requirement to notify individuals of the purposes and uses of the information at the time of collection and (2) the requirement to ensure the accuracy, relevance, timeliness, and completeness of records. In general, the exemptions for law enforcement purposes are intended to prevent the disclosure of information collected as part of an ongoing investigation that could impair the investigation or allow those under investigation to change their behavior or take other actions to escape prosecution.

The E-Government Act of 2002 strives to enhance protection for personal information in government information systems or information collections by requiring that agencies conduct PIAs. As described earlier, a PIA is an analysis of how personal information is collected, stored, shared, and managed in a federal system. More specifically, according to OMB guidance,¹⁴ a PIA is an analysis of how

...information is handled: (i) to ensure handling conforms to applicable legal, regulatory, and policy requirements regarding privacy; (ii) to determine the risks and effects of collecting, maintaining, and disseminating information in identifiable form in an electronic information system; and (iii) to examine and evaluate protections and alternative processes for handling information to mitigate potential privacy risks.

Agencies must conduct PIAs before (1) developing or procuring information technology that collects, maintains, or disseminates information that is in a personally identifiable form or (2) initiating any new data collections involving personal information that will be collected, maintained, or disseminated using information technology if the same questions are asked of 10 or more people. OMB guidance also requires agencies to conduct PIAs in two specific types of situations: (1) when, as a result of the adoption or alteration of business processes, government databases holding information in personally identifiable form are merged, centralized, matched with other databases, or otherwise significantly manipulated and (2) when agencies work together on shared functions

¹⁴OMB, *Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*, Memorandum M-03-22 (Washington, D.C.: Sept. 26, 2003).

involving significant new uses or exchanges of information in personally identifiable form.¹⁵

DHS has also developed its own guidance¹⁶ requiring PIAs to be performed when one of its offices is developing or procuring any new technologies or systems, including classified systems, that handle or collect personally identifiable information. It also requires that PIAs be performed before pilot tests are begun for these systems or when significant modifications are made to them. Furthermore, DHS has prescribed detailed requirements for PIAs. For example, PIAs must describe all uses of the information, and whether the system analyzes data in order to identify previously unknown patterns or areas of note or concern.

Fair Information Practices

The Privacy Act of 1974 is largely based on a set of internationally recognized principles for protecting the privacy and security of personal information known as the Fair Information Practices. A U.S. government advisory committee first proposed the practices in 1973 to address what it termed a poor level of protection afforded to privacy under contemporary law.¹⁷ The Organization for Economic Cooperation and Development (OECD)¹⁸ developed a revised version of the Fair Information Practices in 1980 that has, with some variation, formed the basis of privacy laws and related policies in many countries, including the United States, Germany,

¹⁵A PIA may not be required for all systems. For example, no assessment is required when the information collected relates to internal government operations, when the information has been previously assessed under an evaluation similar to a PIA, or when privacy issues are unchanged.

¹⁶DHS Privacy Office, *PIA Official Guidance* (March 2006).

¹⁷U.S. Department of Health, Education, and Welfare, *Records, Computers and the Rights of Citizens: Report of the Secretary's Advisory Committee on Automated Personal Data Systems* (Washington, D.C.: July 1973).

¹⁸OECD, *Guidelines on the Protection of Privacy and Transborder Flow of Personal Data* (Sept. 23, 1980). The OECD plays a prominent role in fostering good governance in the public service and in corporate activity among its 30 member countries. It produces internationally agreed-upon instruments, decisions, and recommendations to promote rules in areas where multilateral agreement is necessary for individual countries to make progress in the global economy.

Sweden, Australia, New Zealand, and the European Union.¹⁹ The eight principles of the OECD Fair Information Practices are shown in table 1.

Table 1: Fair Information Practices

Principle	Description
Collection limitation	The collection of personal information should be limited, should be obtained by lawful and fair means, and, where appropriate, with the knowledge or consent of the individual.
Data quality	Personal information should be relevant to the purpose for which it is collected, and should be accurate, complete, and current as needed for that purpose.
Purpose specification	The purposes for the collection of personal information should be disclosed before collection and upon any change to that purpose, and its use should be limited to those purposes and compatible purposes.
Use limitation	Personal information should not be disclosed or otherwise used for other than a specified purpose without consent of the individual or legal authority.
Security safeguards	Personal information should be protected with reasonable security safeguards against risks such as loss or unauthorized access, destruction, use, modification, or disclosure.
Openness	The public should be informed about privacy policies and practices, and individuals should have ready means of learning about the use of personal information.
Individual participation	Individuals should have the following rights: to know about the collection of personal information, to access that information, to request correction, and to challenge the denial of those rights.
Accountability	Individuals controlling the collection or use of personal information should be accountable for taking steps to ensure the implementation of these principles.

Source: OECD.

The Fair Information Practices are not precise legal requirements. Rather, they provide a framework of principles for balancing the need for privacy with other public policy interests, such as national security, law enforcement, and administrative efficiency. Ways to strike that balance vary among countries and according to the type of information under consideration.

¹⁹European Union Data Protection Directive (“Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and the Free Movement of Such Data”) (1995).

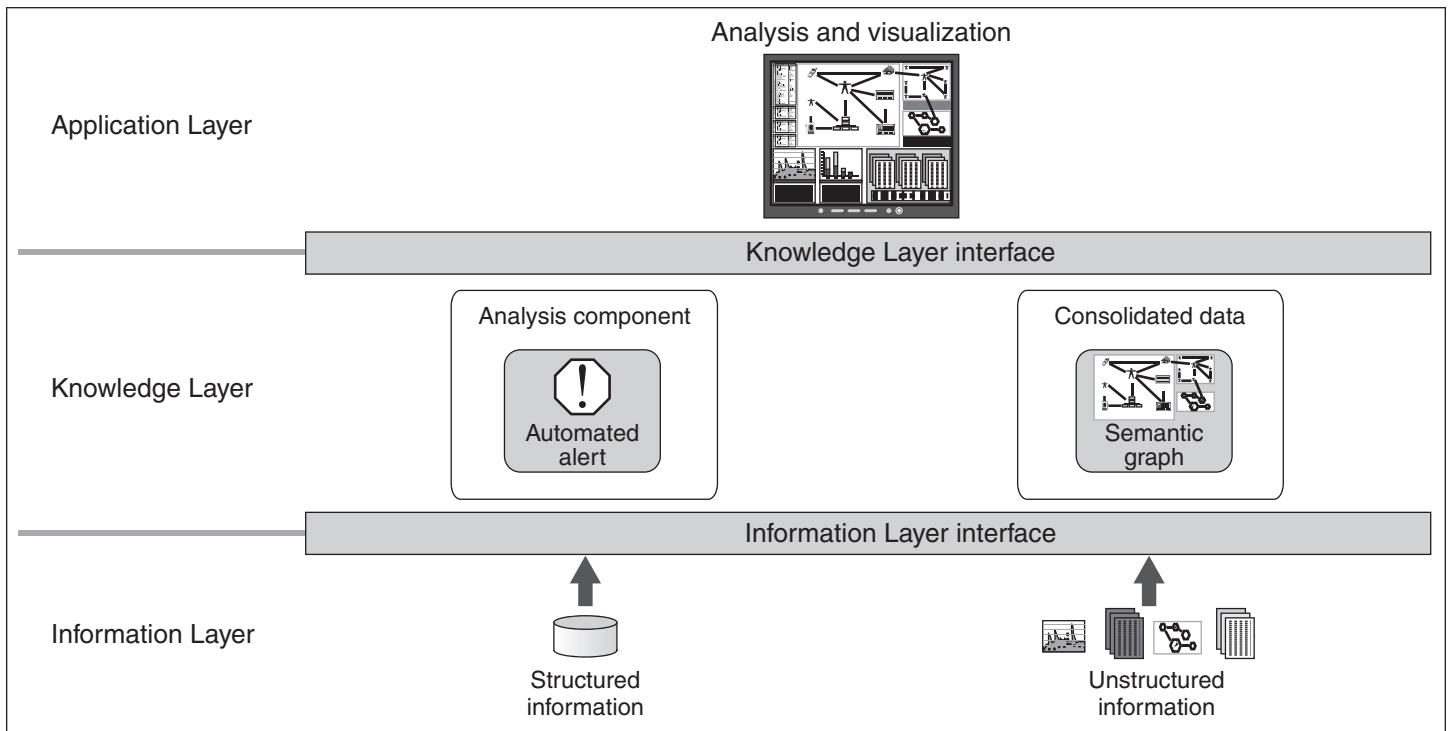
ADVISE Is Intended to Help Identify Patterns of Interest to Homeland Security Analysts

ADVISE is a data mining tool under development that is intended to facilitate the analysis of large amounts of data. It is designed to accommodate both structured data (such as information in a database) and unstructured data (such as e-mail texts, reports, and news articles) and to allow an analyst to search for patterns in data, including relationships among entities (such as people, organizations, and events) and to produce visual representations of these patterns, referred to as semantic graphs. Although none are fully operational, DHS's planned uses of this tool include implementations at several departmental components, including Immigration and Customs Enforcement and other components. DHS is also considering further deployments of ADVISE. The intended benefit of the ADVISE tool is to help detect activities that threaten the United States by facilitating the analysis of large amounts of data that otherwise would be prohibitively difficult to review. DHS is currently in the process of testing the tool's effectiveness.

The ADVISE Tool Provides Analytical Capabilities Intended to Identify Patterns of Interest to DHS Analysts

ADVISE provides several capabilities that help to find and track relationships in data. These include graphically displaying the results of searches and providing automated alerts when predefined patterns of interest emerge in the data. The tool consists of three main elements—the Information Layer, Knowledge Layer, and Application Layer (depicted in fig. 2).

Figure 2: Major Elements and Functions of ADVISE



Source: DHS.

Information Layer

At the Information Layer, disparate data are brought into the tool from various sources. These data sources can be both structured (such as computerized databases and watch lists) and unstructured (such as news feeds and text reports). For structured data, ADVISE contains software applications that load the data into the Information Layer and format it to conform to a specific predefined data structure, known as an ontology. Generally speaking, ontologies define entities (such as a person or place), attributes (such as name and address), and the relationships among them.

For unstructured data, ADVISE includes several tools that extract information about entities and attributes. As with structured data, the output of these analyses is formatted and structured according to an ontology. Tagging information as specific entities and attributes is more difficult with unstructured data, and ADVISE includes tools that allow analysts to manually identify entities, attributes, and relationships among them. According to DHS officials, research is continuing on developing

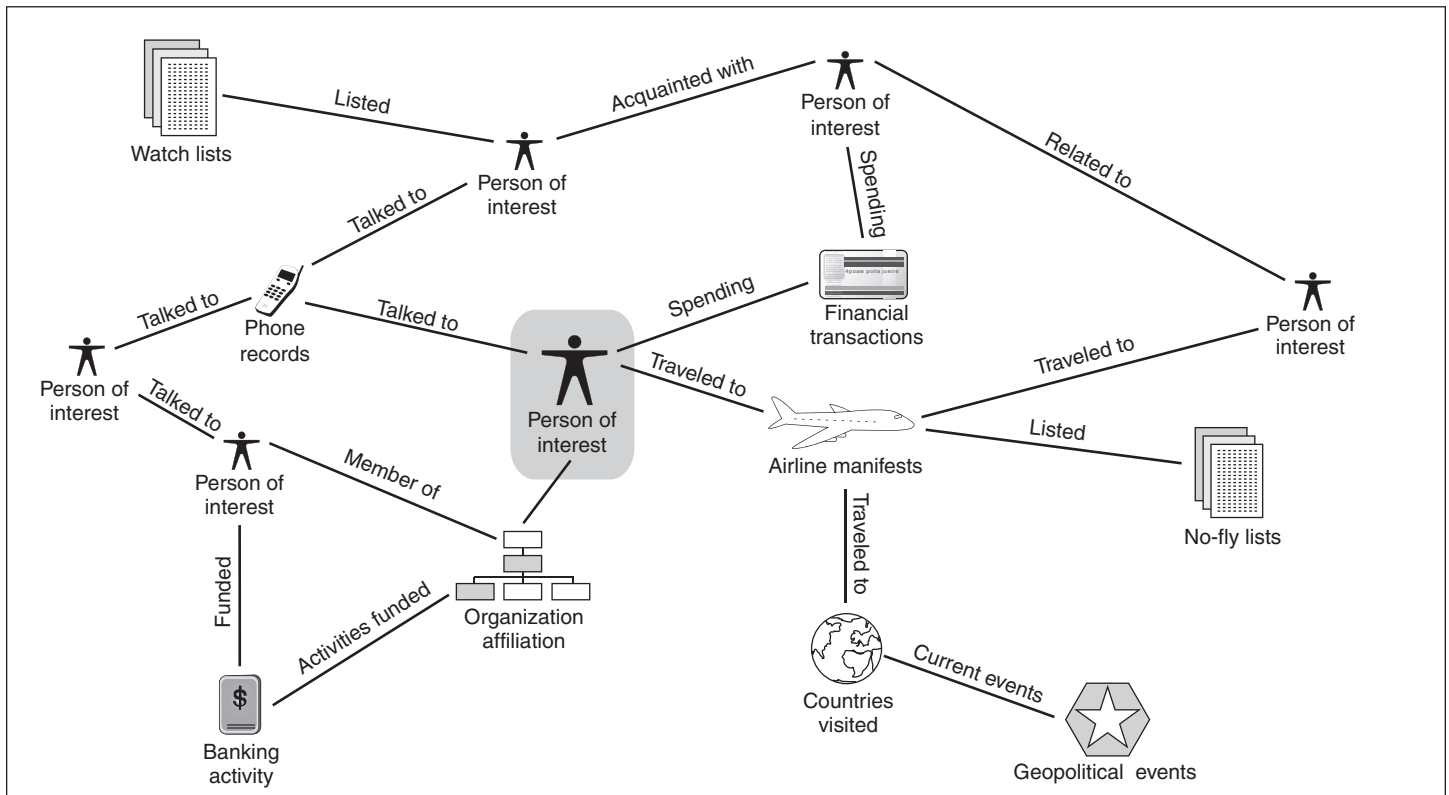
efficient and effective mechanisms for inputting different forms of unstructured data.

ADVISE can also include information about the data—known as “metadata”—such as the time period to which the data pertain and whether the data refer to a U.S. person. ADVISE metadata also include confidence attributes, ranging from 1 to -1, which represent subjective assessments of the accuracy of the data. Each data source has a predefined confidence attribute. Analysts can change the confidence attribute of specific data, but changes to confidence levels are tracked and linked to the analysts making the changes.

Knowledge Layer

At the Knowledge Layer, facts and relationships from the Information Layer are consolidated into a large-scale semantic graph and various subgraphs. Semantic graphing is a data modeling technique that uses a combination of “nodes,” representing specific entities, and connecting lines, representing the relationships among them. Because they are well-suited to representing data relationships and linkages, semantic graphs have emerged as a key technology for consolidating and organizing disparate data. Figure 3 represents the format that a typical semantic graph could take. The Knowledge Layer contains the semantic graph of all facts reported through the Information Layer interface and organized according to the ontology.

Figure 3: Typical Semantic Graph



Source: GAO.

The Knowledge Layer also includes the capability to provide automatic alerts to analysts when patterns of interest (or partial patterns) are matched by new incoming information.

Application Layer

At the Application Layer, analysts are able to interact with the data that reside in the Knowledge Layer. The Application Layer contains tools that allow analysts to perform both pattern-based and subject-based queries and to search for data that match a specific pattern, as well as data that are connected with a specific entity. For example, analysts could search for all of the individuals who have traveled to a certain destination within a given period of time, or they could search for all information connected with a particular person, place, or organization. The resulting output of these searches is then graphically displayed via semantic graphs.

ADVISE's Application Layer also provides several other capabilities that allow for the further examination and adjustment of its output. An analyst

can pinpoint nodes on a semantic graph to view and examine additional information related to them, including the source from which the information and relationships are derived, the data source's confidence level, and whether the data pertain to U.S. persons.

The ADVISE Application Layer also provides analysts the ability to monitor patterns of interest in the data. Science and Technology Directorate staff work with component staff to define patterns of interest and build an inventory of automated searches. These patterns are continuously being monitored in the data, and an alert is provided whenever there is a match. For example, an analyst could define a pattern of interest as "all individuals traveling from the United States to the Middle East in the next 6 months" and have the ADVISE tool provide an alert whenever this pattern emerges in the data.

ADVISE Is Expected to Benefit DHS by Helping to Detect Potentially Threatening Activities

The current planned uses of the ADVISE tool include implementations at several DHS components that are planning to use it in a variety of homeland security applications to further their respective organizational missions. Currently none of these implementations is fully operational or widely accessible to DHS analysts. Rather, they are all still in various phases of systems development. These applications are expected to use the tool primarily to help analysts detect threats to the United States, such as identifying activities and/or individuals that could be associated with terrorism.

The intended benefit of the ADVISE tool is to consolidate large amounts of structured and unstructured data and permit their analysis and visualization. The tool could thus assist analysts to identify and monitor patterns of interest that could be further investigated and might otherwise have been missed.

None of the DHS components have fully implemented the tool in operational systems and, as discussed earlier, testing of the tool is still under way. Until such testing is complete and component implementations are fully operational, the intended benefit remains largely potential.

DHS Has Not Yet Addressed Key Privacy Risks Associated with Expected Uses of the ADVISE Tool

Use of the ADVISE tool raises a number of privacy concerns. DHS has added security controls to the ADVISE tool, including access restrictions, authentication procedures, and security auditing capability. However, it has not assessed privacy risks. Privacy risks that could apply to ADVISE include the potential for erroneous association of individuals with crime or terrorism through data that are not accurate for that purpose, the misidentification of individuals with similar names, and the use of data that were collected for other purposes. A PIA would determine the privacy risks associated with ADVISE and help officials determine what specific controls are needed to mitigate those risks. Although department officials believe a PIA is not needed given that the ADVISE tool itself does not contain personal data, the E-Government Act of 2002 and related federal guidance require the completion of PIAs from the early stages of development. Further, if a PIA were conducted and privacy risks identified, a number of controls exist that could be built into the tool to mitigate those risks. For example, controls could be implemented to ensure that personal information is used only for a specified purpose or compatible purposes, or they could provide the capability to distinguish among individuals that have similar names (a process known as disambiguation) to address the risk of misidentification. Because privacy risks such as these have not been assessed and decisions about mitigating controls have not been made, DHS faces the likelihood that system implementations based on the tool may require costly and potentially duplicative retrofitting at a later date to add the needed controls.

Potential Privacy Concerns Arise with the Use of the ADVISE Tool to Process Personal Information

Like other data mining applications, the use of the ADVISE tool in conjunction with personal information raises concerns about a number of privacy risks that could potentially have an adverse impact on individuals. As the DHS Privacy Office's July 2006 report on data mining activities notes, "privacy and civil liberties issues potentially arise in every phase of the data mining process."²⁰

Potential privacy risks can be categorized in relation to the Fair Information Practices, which, as discussed earlier, form the basis for privacy laws such as the Privacy Act. For example, the potential for personal information to be improperly accessed or disclosed relates to the *security safeguards* principle, which states that personal information

²⁰DHS, *Data Mining Report: DHS Privacy Office Response to House Report 108-774* (July 6, 2006), p. 12.

should be protected against risks such as loss or unauthorized access, destruction, use, modification, or disclosure. Further, the potential for individuals to be misidentified or erroneously associated with inappropriate activities is inconsistent with the *data quality* principle that personal data should be accurate, complete, and current, as needed for a given purpose. Similarly, the risk that information could be used beyond the scope originally specified is based on the *purpose specification* and *use limitation* principles, which state that, among other things, personal information should only be collected and used for a specific purpose and that such use should be limited to the specified purpose and compatible purposes.

Like other data mining applications, the ADVISE tool could misidentify or erroneously associate an individual with undesirable activity such as fraud, crime, or terrorism—a result known as a false positive. False positives may be the result of poor data quality, or they could result from the inability of the system to distinguish among individuals with similar names. *Data quality*, the principle that data should be accurate, current, and complete as needed for a given purpose, could be particularly difficult to ensure with regard to ADVISE because the tool brings together multiple, disparate data sources, some of which may be more accurate for the analytical purpose at hand than others. If data being analyzed by the tool were never intended for such a purpose or are not accurate for that purpose, then conclusions drawn from such an analysis would also be erroneous.

Another privacy risk is the potential for use of the tool to extend beyond the scope of what it was originally designed to address, a phenomenon commonly referred to as function or mission “creep.” Because it can facilitate a broad range of potential queries and analyses and aggregate large quantities of previously isolated pieces of information, ADVISE could produce aggregated, organized information that organizations could be tempted to use for purposes beyond that which was originally specified when the information was collected. The risks associated with mission creep are relevant to the *purpose specification* and *use limitation* principles.

DHS Has Implemented Security Controls but Has Not Yet Assessed Privacy Risks

To address security, DHS has included several types of controls in ADVISE. These include authentication procedures, access controls, and security auditing capability. For example, an analyst must provide a valid user name and password in order to gain access to the tool. Further, upon gaining access, only users with appropriate security clearances may view

sensitive data sets. Each service requested by a user—such as issuing a query or retrieving a document—is checked against the user’s credentials and access authorization before it is provided. In addition, these user requests and the tool’s responses to them are all recorded in an audit log.

While inclusion of controls such as these is a key step in guarding against unauthorized access, use, disclosure, or modification, such controls alone do not address the full range of potential privacy risks. The need to evaluate such risks early in the development of information technology is consistently reflected in both law (the E-Government Act of 2002) and related federal guidance. The E-Government Act requires that a PIA be performed before an agency develops or procures information technology that collects, maintains, or disseminates information in a personally identifiable form. Further, both OMB and DHS PIA guidance emphasize the need to assess privacy risks from the early stages of development.²¹

However, although DHS officials are considering performing a PIA, no PIA or other privacy risk assessment has yet been conducted. The DHS Privacy Office²² instructed the Science and Technology Directorate that a PIA was not required because the tool alone did not contain personal data.²³ According to the Privacy Office rationale, only specific system implementations based on ADVISE that contained personal data would likely require PIAs, and only at the time they first began to use such data. However, guidance on conducting PIAs makes it clear that they should be

²¹DHS PIA guidance states that “[t]he purpose of a PIA is to demonstrate that system owners and developers have consciously incorporated privacy protections throughout the entire life cycle of a system. This involves making certain that privacy protections are built into the system from the start, not after the fact when they can be far more costly or could affect the viability of the project.” In addition, OMB guidance states that “[a]gencies should commence a PIA when they begin to develop a new or significantly modified IT system.”

²²The DHS Privacy Office was created in response to the Homeland Security Act of 2002, Pub. L. No. 107-296, § 222, 116 Stat. 2155 (Nov. 25, 2002). The Privacy Officer is responsible for, among other things, “assuring that the use of technologies sustain[s], and do[es] not erode privacy protections relating to the use, collection, and disclosure of personal information.”

²³It is important to note the distinction between the PIA requirement, based on the E-Government Act, and the requirements of the Privacy Act. Because the ADVISE tool itself does not contain any data, it is not considered a system of records for purposes of the Privacy Act and thus is not subject to the requirements of that law. As ADVISE implementations move from development to operations, they may lead to the creation or modification of systems of records, which would require the development of appropriate privacy notices to be published in the *Federal Register* and other actions to protect privacy.

performed at the early stages of development. OMB's PIA guidance requires PIAs at the IT development stage, stating that they "should address the impact the system will have on an individual's privacy, specifically identifying and evaluating potential threats relating to elements identified [such as the nature, source, and intended uses of the information] to the extent these elements are known at the initial stages of development." Regarding ADVISE, the tool's intended uses include applications containing personal information. Thus the requirement to conduct a PIA from the early stages of development applies.

As of November 2006, the ADVISE program office and DHS Privacy Office were in discussions regarding the possibility of conducting a privacy assessment similar to a PIA but modified to address the development of a technological tool. No final decision has yet been made on whether or how to proceed with a PIA. However, until such an assessment is performed, DHS cannot be assured that privacy risks have been identified or will be mitigated for system implementations based on the tool.

Privacy Protection Controls to Mitigate Identified Risks Exist and Could Be Built into ADVISE

A variety of privacy controls can be built into data mining software applications, including the ADVISE tool, to help mitigate risks identified in PIAs and protect the privacy of individuals whose information may be processed. DHS has recognized the importance of implementing such privacy protections when data mining applications are being developed. Specifically, in its July 2006 report, the DHS Privacy Office recommended instituting controls for data mining activities that go beyond conducting PIAs and implementing standard security controls. Such measures could be applied to the development of the ADVISE tool.²⁴ Among other things, the DHS Privacy Office recommended that DHS components use data mining tools principally as investigative tools and not as a means of making automated decisions regarding individuals.²⁵ The report also emphasizes that data mining should produce accurate results and recommends that DHS adopt data quality standards for data used in data mining. Further, the report recommends that data mining projects give

²⁴The Privacy Office's report states that ADVISE is a "technology" and not a data mining program. Accordingly, the report's recommendations ostensibly would not apply to ADVISE. However, the report acknowledges that uses of ADVISE may constitute data mining, in which case the recommendations would apply.

²⁵ADVISE does not provide an automated means for making decisions about individuals. Rather, it is an analysis tool to aid analysts in identifying relationships and patterns of interest.

explicit consideration to using anonymized data when personally identifiable information is involved. Although some of the report's recommendations may apply only to operational data mining activities, many reflect system functionalities that can be addressed during technology development.

Based on privacy risks identified in a PIA, controls exist that could be implemented in ADVISE to mitigate those risks. For example, controls could be implemented to enforce use limitations associated with the purpose specified when the data were originally collected. Specifically, software controls could be implemented that require an analyst to specify an allowable purpose and check that purpose against the specified purposes of the databases being accessed.

Regarding data quality risks, the ADVISE tool currently does not have the capability to distinguish among individuals with similar identifying information, nor does it have a mechanism to assess the accuracy of the relationships it uncovers. To address the risk of misidentification, software could be added to the tool to distinguish among individuals that have similar names, a process known as disambiguation. Disambiguation tools have been developed for other applications. Additionally, although the ADVISE tool includes a feature that allows analysts to designate confidence levels for individual pieces of data, no mechanism has been developed to assess the confidence of relationships identified by the tool. While software specifically to determine data quality would be difficult to develop, other controls exist that could be readily used as part of a strategy for mitigating this risk. For example, anonymization could be used to minimize the exposure of personal data, and operational procedures could be developed to restrict the use of analytical results containing personal information that could have data quality concerns. To implement anonymization, the tool would need the software capability to handle anonymized data or have a built-in data anonymizer. DHS currently does not have plans to build anonymization into the ADVISE tool.²⁶

Until a PIA that identifies the privacy risks of ADVISE is conducted and privacy controls to mitigate those risks are implemented, DHS faces the

²⁶In addition, a feature was to be implemented in January 2007 that would enforce an internal DHS rule regarding how long information about U.S. persons can be maintained in intelligence data bases. However, because this control is designed to respond only to the DHS rule—and not to identified privacy risks—it leaves potential concerns unaddressed about how personal information is used when it is maintained and processed by ADVISE.

risk that privacy concerns will arise during implementation of systems based on ADVISE that may be more difficult to address at that stage and possibly require costly retrofitting.

Conclusions

The ADVISE tool is intended to provide the capability to ingest large amounts of data from multiple sources and to display relationships that can be discerned within the data. Although the ADVISE tool has not yet been fully implemented and its effectiveness is still being evaluated, the chief intended benefit is to help detect activities threatening to the United States by facilitating the analysis of large amounts of data.

The ADVISE tool incorporates security controls intended to protect the information it processes from unauthorized access. However, because ADVISE is intended to be used in ways that are likely to involve personal data, a range of potential privacy risks could be involved in its operational use. Thus, it is important that those risks be assessed—through a PIA—so that additional controls can be established to mitigate them. However, DHS has not yet conducted a PIA, despite the fact that the E-Government Act and related OMB and DHS guidance emphasize the need to assess privacy risks early in systems development. Although DHS officials stated that they believe a PIA is not required because the tool alone does not contain personal data, they also told us they are considering conducting a modified PIA for the tool. Until a PIA is conducted, little assurance exists that privacy risks have been rigorously considered and mitigating controls established. If controls are not addressed now, they may be more difficult and costly to retrofit at a later stage.

Recommendations for Executive Action

To ensure that privacy protections are in place before DHS proceeds with implementations of systems based on ADVISE, we recommend that the Secretary of Homeland Security take the following two actions:

- immediately conduct a privacy impact assessment of the ADVISE tool to identify privacy risks, such as those described in this report, and
- implement privacy controls to mitigate potential privacy risks identified in the PIA.

Agency Comments and Our Evaluation

We received oral and written comments on a draft of this report from the DHS Departmental GAO/Office of Inspector General Liaison Office. (Written comments are reproduced in appendix II.) DHS officials generally agreed with the content of this report and described actions initiated to address our recommendations. DHS also provided technical comments, which have been incorporated in the final report as appropriate.

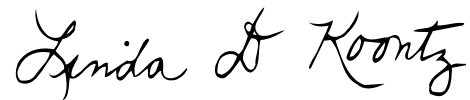
In its comments DHS emphasized the fact that the ADVISE tool itself does not contain personal data and that each deployment of the tool will be reviewed through the department's privacy compliance process, including, as applicable, development of a PIA and a system of records notice. DHS further stated that it is currently developing a "Privacy Technology Implementation Guide" to be used to conduct a PIA for ADVISE. Although we have not reviewed the guide, it appears to be a positive step toward developing a PIA process to address technology tools such as ADVISE.

It is not clear from the department's response whether the privacy controls identified based on applying the Privacy Technology Implementation Guide to ADVISE are to be incorporated into the tool itself. We believe that any controls identified by a PIA to mitigate privacy risks should be implemented, to the extent possible, in the tool itself. Specific development efforts that use the tool will then have these integrated controls readily available, thus reducing the potential for added costs and technical risks. The department also requested that we change the wording of our recommendation; however, we have retained the wording in our draft report because it clearly emphasizes the need to incorporate privacy controls into the ADVISE tool itself.

As agreed with your office, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the report date. At that time, we will send copies of this report to the Secretary of Homeland Security and other interested congressional committees. Copies will be made available to others on request. In addition, this report will be available at no charge on our Web site at www.gao.gov.

If you have any questions concerning this report, please call me at (202) 512-6240 or send e-mail to koontzl@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Key contributors to this report are listed in appendix III.

Sincerely yours,

A handwritten signature in black ink that reads "Linda D. Koontz". The signature is written in a cursive, flowing style.

Linda D. Koontz
Director, Information Management Issues

Appendix I: Objectives, Scope, and Methodology

Our objectives were to determine the following:

- the planned capabilities, uses, and associated benefits of the Analysis Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE) tool and
- whether potential privacy issues could arise from using the ADVISE tool to process personal information and how the Department of Homeland Security (DHS) has addressed any such issues.

To address our first objective, we identified and analyzed the tool's capabilities, planned uses, and associated benefits. We reviewed program documentation, including annual program execution plans, and interviewed agency officials responsible for managing and implementing the program, including officials from the DHS Science and Technology Directorate and the Lawrence Livermore and Pacific Northwest National Laboratories. We also viewed a demonstration of the tool's semantic graphing capability. In addition, we interviewed officials at DHS components to identify their current or planned uses of ADVISE, the progress of their implementations, and the benefits they hope to gain from using the tool. These components included Immigrations and Customs Enforcement and other components. We also interviewed officials from the Interagency Center of Applied Homeland Security Technology (ICAHST), who are responsible for conducting testing of the tool's capabilities. We also visited ICAHST at the John Hopkins Applied Physics Laboratory in Laurel, Maryland, to view a demonstration of its testing activities. We did not conduct work or review implementations of ADVISE at the DHS Office of Intelligence and Analysis.

To address our second objective, we identified potential privacy concerns that could arise from using the ADVISE tool by reviewing relevant reports, including prior GAO reports and the DHS Privacy Office 2006 report on data mining. We identified and analyzed DHS actions to comply with the Privacy Act of 1974 and the E-Government Act of 2002. We interviewed technical experts within the DHS Science and Technology Directorate and personnel responsible for implementing ADVISE at DHS components to assess privacy controls included in the ADVISE tool. We also interviewed officials from the DHS Privacy Office. We performed our work from June 2006 to December 2006 in the Washington, D.C., metropolitan area. Our work was performed in accordance with generally accepted government auditing standards.

Appendix II: Comments from the Department of Homeland Security

U.S. Department of Homeland Security
Washington, DC 20528



**Homeland
Security**

February 2, 2007

Ms. Linda D. Koontz
Director, Information Management Issues
U.S. Government Accountability Office
Washington, D. C. 20548

Dear Ms. Koontz:

Thank you for the opportunity to comment on the draft report GAO-07-293 "Datamining: Early Attention to Privacy in Developing a Key DHS Program Could Reduce Risks". In this draft report, GAO recommends that "the Secretary of Homeland Security immediately conduct a privacy impact assessment (PIA) of the ADVISE tool and implement privacy controls as needed to mitigate any identified risks" to ensure that privacy protections are in place

The ADVISE toolset is a set of generic IT tools and does not in itself collect or use any data. The individual ADVISE tools could be combined to create specific systems which would be designed to support specific operational needs. Each of these deployments of the ADVISE toolset would be reviewed and reported through the DHS privacy compliance process and documented in the Privacy Threshold Analysis (PTA) and, as applicable, the Privacy Impact Assessment (PIA) and System of Records Notice (SORN).

Within DHS, the term "Privacy Impact Assessment" refers to two separate and related functions. The first is the *PIA activity* of assessing a technology, program, etc. for potential privacy impacts. The second is the *PIA report* that documents the results of that assessment activity. The distinctions between these two meanings of the term may help to clarify DHS's approach to assessing the potential privacy impacts of the ADVISE toolset.

The DHS Privacy Impact Assessment form (the report) is designed for operational systems and is not well suited to open-ended toolsets like ADVISE. In conducting the privacy impact assessment (the activity) DHS decided that a different type of document would better fit the nature of the ADVISE toolset. Rather than using a descriptive reporting document, DHS is using a proscriptive guidance document that is tailored to the specific nature of the ADVISE toolset. This guidance document is called a "Privacy Technology Implementation Guide" and is currently being developed by the DHS Privacy Office.

www.dhs.gov

The Privacy Technology Implementation Guide is more adaptable to technology frameworks like ADVISE and provides guidance as to how the individual ADVISE tools could be implemented in privacy protective ways. System Developers building specific implementations of ADVISE can use the Guide to build privacy protections into the systems as part of the development process. The flexibility of the Privacy Technology Implementation Guide allows for integrated privacy protection in all uses of the ADVISE tools and will be complemented by Privacy Impact Assessment documents for operational deployments (individual systems).

The current draft of the Privacy Technology Implementation Guide for ADVISE is organized into two sections. The first section identifies privacy protections related to the general nature of ADVISE as a set of tools and recommends that the same privacy protections related to datamining be applied to ADVISE. This first section further recommends that the value and limitations of ADVISE tools be specifically identified and matched to the purpose and success measures of the specific DHS mission the tool would be implemented to support. The second section is organized by the technology architecture (information layer, knowledge layer, security layer, application layer) and offers specific guidance for integrating privacy protection into the use of the tools from each of these architectural components.

The privacy impact assessment activity led to the determination that a new type of document would further assist in building privacy protections into technology. The Privacy Technology Implementation Guide is that new type of documentation and is being developed to accompany the toolset itself. The expected result is that as technology developers decide that the ADVISE toolset could be used to meet a particular DHS need, they receive privacy technology guidance along with the toolset to assist in building privacy protections into the system from the beginning. DHS will continue to use the suite of privacy compliance documents (PTA, PIA, SORN) to report on the potential privacy impacts and integrated privacy protections for each individual systems.

Requested change to GAO recommendation

Based on the above, DHS requests that GAO revise its recommendations to read:

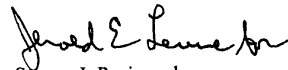
“To ensure that privacy protections are integrated into the development process, the Secretary of Homeland Security should create privacy controls for the ADVISE toolset to guide specific development efforts and conduct a privacy impact assessment for each ADVISE deployment to ensure those controls are implemented and effective.”

Two additional edits

- Page 21, Table 2: Please remove the ICE “implementation.” DHS is only engaged in early discussion and no implementation is currently planned.
- Page 26: The report seems to suggest that ADVISE provides an automated means for making decisions regarding individuals. DHS would like to clarify that ADVISE is an aid for analysts in identifying relationships and patterns of interest. ADVISE is an analysis tool an not a decision-making tool. The tool itself does not make decisions.

DHS appreciates GAO’s work in planning, conducting and issuing this report and for the opportunity to review the draft.

Sincerely,



Steven J. Pecinovsky
Director, Departmental GAO/OIG Liaison Office

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Linda D. Koontz, (202) 512-6240 or koontzl@gao.gov

Staff Acknowledgments

In addition to the individual named above, John de Ferrari, Assistant Director; Idris Adjerid; Nabajyoti Barkakati; Barbara Collier; David Plocher; and Jamie Pressman made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548