



# Homeland Security

April 14, 2008

Ms. Marcia Hofmann  
Electronic Frontier Foundation  
1875 Connecticut Avenue, N.W.  
Suite 650  
Washington, D.C. 20009

Re: DHS/OS/PRIV 07-197/Hofmann request

Dear Ms. Hofmann:

This is the final letter in response to your Freedom of Information Act (FOIA) request to the Department of Homeland Security (DHS), dated November 20, 2006 and referred to this office by the Privacy Office on March 7, 2008. We were asked to review 11 documents, consisting of 27 pages, to determine if your requested documents can be released or if the documents are exempt from release.

Of those pages, we have determined that 25 pages of the records are releasable in their entirety, one page is partially releasable and we are withholding one page in their entirety pursuant to Title 5 U.S.C. § 552. I have withheld these documents under FOIA Exemption 5.

**FOIA Exemption 5** protects from disclosure those inter- or intra-agency documents that are normally privileged in the civil discovery context. The three most frequently invoked privileges are the deliberative process privilege, the attorney work-product privilege, and the attorney-client privilege. After carefully reviewing the responsive documents, I determined that the responsive documents qualify for protection under the deliberative process privilege. The deliberative process privilege protects the integrity of the deliberative or decision-making processes within the agency by exempting from mandatory disclosure opinions, conclusions, and recommendations included within inter-agency or intra-agency memoranda or letters. The release of this internal information would discourage the expression of candid opinions and inhibit the free and frank exchange of information among agency personnel.

You have a right to appeal the above withholding determination. Should you wish to do so, you must send your appeal and a copy of this letter, within 60 days of the date of this letter, to: Associate General Counsel (General Law), U.S. Department of Homeland Security, Washington, D.C. 20528, following the procedures outlined in the DHS regulations at 6 C.F.R. § 5.9. Your envelope and letter should be marked "FOIA Appeal." Copies of the FOIA and DHS regulations are available at [www.dhs.gov/foia](http://www.dhs.gov/foia).

If you need to contact our office again about this matter, please refer to S&T 08-0003.13/Hofmann request. This office can be reached at (202) 254-6819.

Sincerely,

*Nicole Marum*  
*Acting AGC for S+T*

Mark E. Rosen

Associate General Counsel for Science & Technology

Enclosure: a/s

FOR OFFICIAL USE ONLY

ADVISE

**Specific Open Questions  
March 23, 2007**

The following is a list of outstanding questions related to DHS use of ADVISE. Some of these questions have already been asked of various individuals. PRIV would like S&T to confirm the by providing updated/confirmed answers.

In the event S&T would like to refer specific questions to other DHS components, please also send the contact information for that individual to PRIV.

**The Interagency Center for Applied Homeland Security Technology (ICAHST)**

The following are general questions about the overall capability of ADVISE tools:

1. Please provide a short summary of the results of the testing to date: what works, what does not work? Any overall views, based on tests, for the future of ADVISE?
2. What is the overall value of ADVISE, what can be said about it that fits with the current investment?
3. Can ADVISE identify "unknown" patterns and/or predict?
4. Can ADVISE automatically identify items in data or automatically establish relationships?

From S&T descriptions, PRIV understands that ADVISE can only work with entities and relationships established by a human – that the advantage of ADVISE is its capability to work with all entities and relationships already established in the data.

*Please clarify S&T's responses in the discussion with Herb Engle from March 12, 2006 – reproduced at the bottom "ICAHST Capability."*

5. Does the Ontology create a limit on the data that can be loaded into a deployment of ADVISE technology? What is the interaction between a data load scenario and an ontology if the ontology is more limited than the available data? Is the ontology updated or the data load limited?
6. Has ADVISE ever been used to make any operational decisions?
7. Did any individual or DHS component outside S&T operate an ADVISE pilot?

S&T reported that S&T operated all the pilots (loading data, identifying relationships, demonstrating results) and that all others only watched.

**All-Weapons of Mass Effect (All-WME)**

The following are specific questions directly related to the All-WME pilot:

8. A draft Privacy Threshold Analysis states that this pilot originally started on October 1, 2006 and was last updated on October 1, 2006. Please describe the history of this pilot with specific references to those dates.
9. Was any data loaded into the pilot?  
S&T reported that data was loaded from FIBIS (opensource.gov) and CNS (cns.miss.edu).
10. If data was loaded, what are the data elements?
11. If data was loaded, what was the source and range of data?
12. If data was loaded, did any information relate to individuals?
13. Is the pilot intended to relate to groups or individuals?
14. Is the pilot intended to relate to individuals, will the individuals be US Persons?

**Remote Thread Alerting System (RTAS)**

The following are specific questions directly related to the RTAS pilot:

15. What was the exact start and end dates for this pilot?  
S&T reports the pilot started in 2004 and ended in 2006.
16. What does "decommissioned" mean?  
S&T reports RTAS "decommissioned" at the end of the pilot period.
17. What happened to the data once the pilot ended?
18. Is there a way to determine whether the data supplied actually contained personally identifiable information?  
S&T reports the PIERS data included data fields for name and that the data in this field could be either a business or an individual.
19. Please confirm that there is no personally identifiable information in the census data used with this pilot.
20. During the Intellectual property scenario of the demonstration of RTAS, were any searches conducted regarding the shippers, consignees or notify parties?  
S&T states that the intellectual property search was conducted but that S&T does not recall whether the demonstration searched the details of the shipper, consignee or notify parties.
21. At any point, was the RTAS pilot used to search for information about an individual? (Note question 18 re: the determination of whether personally identifiable information could be included in the data sets.)

**ICE Demonstration (ICE Demo)**

The following are specific questions directly related to the ICE Demo pilot:

22. What was the exact start and end dates for this pilot? When did the pilot actually start and when it actually end?

S&T reports the demonstration of the pilot occurred on July 28, 2005.

23. What specific data sources were used in this pilot, what was S&T's source for each data set, and for each data set, what range of data was actually used?

S&T reports that a small sample of data from up to eight different data sources were provided by ICE to S&T. S&T reports that it does not recall whether data from all of the data sources was used, nor does it recall how much data was loaded from the data sources that were used. Without this specific information it is impossible to identify whether personally identifiable information was in fact used during the pilot and thus it is impossible to accurately determine whether there was a privacy violation.

The following are the data sets currently identified related to this pilot. For each data set, please:

- Confirm that it was in fact used;
- Please describe each data set;
- The range of data used from each data set;
- S&T's source for the data set (with specific contact information); and
- Which System of Records Notice covers each data source.

- SEVIS
- LESC.
- No fly List
- Selectee List
- NORA
- NSEERS
- Unconfirmed Overstays
- SITSDATA

24. Were all the data identified in question 24 were fused (combined) together? Or were different sources fused in different combinations?
25. If the data was combined, did the ontology limit the actual data that was actually combined? If so, what data was actually combined?

**Threat Vulnerability Integration System (TVIS)**

26. The June 27, 2006 Privacy Threshold Analysis states that this pilot is a new development effort. When did it start and what were the dates of any substantial updates to the pilot?

27. Was all personally identifiable information removed from this pilot?

S&T reports that all PII was removed from the pilot and all activity related to this pilot has stopped.

28. Please confirm these data source were the actual data sources used. If any other data sources were used please identify those other data sources.

The following are the data sets currently identified related to this pilot. For each data set, please:

- Confirm that it was in fact used;
  - Please describe each data set;
  - The range of data used from each data set;
  - S&T's source for the data set (with specific contact information); and
- No Fly List
  - Selectee List
  - TSC Daily Summaries
  - Intelligence Community Message Traffic
  - NTIDB
  - Patriot Reporting
  - SEVIS

FOR OFFICIAL USE ONLY

Specific Open Questions  
ADVISE

March 23, 2007  
Page 5 of 6

ICAHST CAPABILITY

-----Original Message-----

From: Sand, Peter  
Sent: Monday, March 12, 2007 7:59 AM  
To: Engle, Herbert  
Cc: Hoyt, John; Baicar, Bruce; Jorgensen, Bruce <CTR>  
Subject: RE: ADVISE - ICAHST Questions

Herb,

In terms of the actual function of the ADVISE tools (separate from the experience of using it), can you describe what else it can do - in addition to analyzing the relationships between linked nodes?

Does ADVISE have the capacity to identify new patterns itself? Note, this is different from the manually-created scenarios described in the below quote:

"Linking both nodes gives you a pattern for which you can query from... first we will create a pattern and find it, then we will modify the ontology then find the same pattern again. Results should look like pattern."

Can ADVISE create its own patterns?

Thanks,

Pete

-----Original Message-----

From: Engle, Herbert  
Sent: Monday, March 12, 2007 8:26 AM  
To: Sand, Peter  
Subject: RE: ADVISE - ICAHST Questions

Peter,

If the question is does ADVISE it's self seek out patterns in data and then notify an analyst of a pattern then that would be no. ADVISE is a vary powerful visualization tool. It will graphically depict pattern that are in the data that has been loaded the system but it requires an analyst to identify patterns. An Analyst will query the system asking to see the relationships in the data. ADVISE then take the "raw data" and presents it in the from of links and patterns. The test statement refers to the process of placing a pattern into the system (How A relates to B and how A and B relate to C). Then that pattern is modified and those results are compared to the first pattern. This was part of the Phase 1 testing and was used to verify basic operational functionality.

Herb

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Specific Open Questions  
ADVISE

March 23, 2007  
Page 6 of 6

-----Original Message-----

From: Sand, Peter  
Sent: Monday, March 12, 2007 8:32 AM  
To: Engle, Herbert  
Subject: RE: ADVISE - ICAHST Questions

Herb,

Just so I am clear, ADVISE as a toolset IS NOT CAPABLE of generating patterns on its own. AN ANALYST must CREATE the LINKS between nodes and a certain SET OF LINKS can be stored as a "PATTERN" and be searched for later and by others.

Did I get it right?

Pete

-----Original Message-----

From: Engle, Herbert  
Sent: Monday, March 12, 2007 8:47 AM  
To: Sand, Peter  
Subject: RE: ADVISE - ICAHST Questions

Peter,

ADVISE can show the link in any data that it has if an analyst creates a query that asks to see that relationship. An analyst does not create a link. He may ask to see what links some node has with another. If you load 1000 data points and ask to see the relationship between all or some of them, then the system will show you. The key is that you can pull data from a number of sources. You might have flight information and phone records. ADVISE will let you see how these two types of information relate by providing a graphical representation of how the data links together.

Herb

FOR OFFICIAL USE ONLY

## **All Weapons of Mass Effect (All-WME)**

The All Weapons of Mass Effect (All-WME) program is currently housed in the Command, Control and Interoperability Division at the Science and Technology Directorate (S&T). The program assesses the capabilities of foreign and domestic terrorist groups to develop and deploy WME threat agents.

DoE's Lawrence Livermore National Laboratory (LLNL) started the All-WME effort in October 2002, prior to the formation of DHS. S&T supplied its initial funding for All-WME in 2003. In these early stages, All-WME activities relied on existing data management and analysis tools developed by LLNL and Los Alamos National Laboratory (LANL) scientists. One such tool was known as the Knowledge Integration Tool (KIT), and used simple Web-like interfaces. Until 2005, all the WME analysts were exclusively DoE analysts at LLNL and LANL. They analyzed classified message traffic collected by the laboratories' Field Intelligence Elements (FIEs). Such message traffic may include personally identifiable information, that is, data that can potentially identify a person, but does not contain data on U.S. persons.

In FY 2005, S&T began funding an effort to explore whether the ADVISE framework could be used to analyze All-WME message traffic data. An internal test and development capability was set up for that purpose early in 2006. In addition, a limited set of message traffic data was entered into a separate, stand-alone ADVISE framework for performance testing and evaluation. No operational decisions were made from this performance test and evaluation. At this time, there has been no further work on the test and development system.

The ultimate relevance of ADVISE to All-WME includes:

- Capturing information and knowledge from high-value documents that would not be available through other means
- Capturing and sharing knowledge of analysts. For example, analysts may annotate or vet documents or information which should then be shared with other analysts in their organization
- Fusing data from multiple sources or organizations

Development of an ADVISE-based, All-WME pilot, which was initially planned for FY 2007, was halted in 2006 as a result of funding priorities. The pilot would have characterized the capabilities of adversaries by creating a comprehensive and current awareness of WME related materials and illicit trafficking.

Privacy status: The All-WME analysts operate as part of the DoE FIEs. As such, they strictly follow DoE rules for protecting privacy. S&T is currently drafting a PTA to reflect the All-WME initiatives prior to FY2007.

**Questions and Answers**  
**For**  
**ADVISE**  
(Analysis, Dissemination, Visualization, and Semantic Enhancement)

- 1. Please thoroughly describe the data-mining tool or activity and the data that is being or will be used.**

ADVISE is a framework of tools to analyze and visually represent relationships between people, places and events. It is being developed to provide analysts with help in quickly retrieving the right information for their current research and reporting needs. Intelligence analysts depend upon information from a variety of sources such as documents in many different forms: email, database records, web pages, spreadsheets, text files, etc. The volume of information available on a daily basis is tremendous and continues to grow beyond the ability of anyone to read and assimilate all of the data contained. The ADVISE system has two primary capabilities to assist analysts: 1) ADVISE shows relations between entities (people, places, things) from disparate data sources that would otherwise go unnoticed using traditional information retrieval approaches; 2) ADVISE provides a fast, accurate analysis of huge quantities of documents to locate the few that are pertinent to an analyst's current research needs.

The basic components of ADVISE are a semantic graph, analysis tools, visualization tools, data loading, text processing and documented application programmer interfaces (APIs) between these components. A typical ADVISE deployment consists of rack-mounted servers and support hardware to enable the semantic graph, analysis engines, and text processing. Client workstations execute software that is loaded on demand to query and visualize the data.

ADVISE is loaded with data selected by the implementing organization per that organization's policies. When data is brought into a functioning ADVISE system, the data loading utilities and text processing utilities extract entities, attributes of entities and relationships between entities from the source documents and data. The extracted entities and relationships are used to construct the semantic graph. Attributes about the entities are stored in an additional data store. A document management system is included to access the documents brought into ADVISE.

- 2. Please describe the goals and plans for the use or development of the data-mining tool or activity. For what purpose(s) is the data-mining tool or activity being developed and deployed?**