DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, D.C. 20224

NOV 25 2003

MEMORANDUM FOR BRENT HILL
DIRECTOR, CUSTOMER APPLICATIONS
DEVELOPMENT MANAGEMENT DIVISION

FROM: Maya A. Bernstein
Privacy Advocate

SUBJECT: Management Information System Data Warehouse
Privacy Impact Assessment

The Office of the Privacy Advocate has reviewed the Privacy Impact Assessment (PIA) for the Management Information System Data Warehouse. Based on the information provided, we do not have any privacy concerns that would preclude this system from operating. However, a revised PIA is required when considering any future upgrades or modifications to the system, or before the project progresses to the next milestone.

We will forward a copy of the PIA to the Director, Modernization and System Security, to be included in the Certification and Accreditation package for formal acceptance for operation. That office may request information concerning the statements contained in the PIA to ascertain compliance with applicable security requirements. If you have any questions please contact me at 202-927-5170, or your staff may contact Gino Talbot at 202-622-2302.

Date June 30, 2003

MEMORANDUM FOR MAYA A. BERNSTEIN
PRIVACY ADVOCATE

FROM: H. Phillip Griner, Chief IMAS Section
Rapid Applications Development Division
M:I:SD:RS:MD:AM

SUBJECT: Request for Privacy Impact Assessment (PIA) –
MIS Data Warehouse

Name of Request Contact:

Name: Sharon J West
Organization Name & Symbols: Information Management Applications Section (M:I:B:CA:AP:AM)
Mailing Address: 3651 S. IH-35, Austin, TX 78741
Phone Number (with area code): (512) 460-2647

Name of Business System Owner:

Name: H. Phillip Griner
Organization Name & Symbols: Information Management Applications Section
(M:I: B:CA:AP:AM)
Mailing Address: 3651 S. IH-35, Austin, TX 78741
Phone Number (with area code): (512) 460-2647

Requested Operational Date:

Category:

New: _____X____
Existing System:
Recertification, (with no change, date of last certification)_____
With Significant Modification:_____
National Standard Application (NSA): _____
Modernization Project or System:
Check Milestone: one___ two___ three___ four ___ five ___

System of Record Number(s) (SORN) #:

- Treasury/IRS 24.013 Combined Account Number File
- Treasury/IRS 24.029 Individual Account Number File
- Treasury/IRS 24.030 CADE Individual Master File
- Treasury/IRS 34.037 IRS Audit Trail and Security Records System
- Treasury/IRS 24.046 CADE Business Master File
- IRS 42.008, Audit Information Management System (AIMS)

The above SORN was approved by David Silverman on July 8, 2003.

Purpose of the System:

The Austin IMAS Staff has been tasked to gather Management Information System (MIS) data for the past several years. The MIS Data Warehouse (MISDW) will consolidate these efforts into one authoritative source of MIS information. The primary purpose of this system is to store MIS type data from various IRS systems such as mainframe output runs or other IRS subsystems such as AUR, CAR, etc.

The lack of enterprise Management Information System (MIS) data is problematic making effective management, performance analysis and accomplishment reporting very difficult for all business units.

It is mission critical that this deficiency be immediately addressed. To that end, the IMAS Staff lead the development of a single source for MIS data for all business unit operations nationwide.

The MIS Data Warehouse Project is designed to be an authoritative source for MIS data nationwide. Upon completion, the MIS Data Warehouse will provide a single authoritative source to capture, archive and provide data for critical management information type applications.

Attachment: PIA

Data in the System

1. Describe the information (data elements and fields) available in the system in the following categories: A. Taxpayer B. Employee C. Audit Trail Information (including employee log-in info) D. Other (Describe)	A. Taxpayer: Information provided by existing IRS systems and/or sub-systems. Please refer to Attachment 1. B. Employee: None at this time C. Audit Trail Information: The system the user uses to access the Intranet collects the employee log-in information. The Web Interface will have the capability to record and audit employee activity via their NT Login. However, the MIS Data Warehouse will not have the functionality to provide the identity of a specific NT Login. D. Other: The MIS Data Warehouse contains data from existing IRS legacy systems or other existing balanced measures and workload indicator MIS information. Normally, this information is neither taxpayer nor employee specific. Please refer to Attachment 1.
2. Describe/identify which data elements are obtained from files, databases, individuals, or any other sources. A. IRS B. Taxpayer C. Employee D. Other Federal Agencies (List agency) E. State and Local Agencies (List agency) F. Other third party sources (Describe)	A. The MIS Data Warehouse utilizes existing reports or data files generated from a number of existing Tier I, Tier II and Tier III systems. These include: Audit Information Management System (AIMS), Summary Examination Time Transmission System (SETTS), Automated Substitute for Return (ASFR), Collection Activity Reports (CAR), and Quality Review Database (QRDB). Specific data elements from the reports and/or data file is described in the respective Data Definition Document for each data source. Please refer to Attachment 1. B. Taxpayer information is limited to the data available in existing reports or data files generated by the Tier I, II, or III systems described in Section A. C. The system the user uses to access the Intranet collects the employee log-in information. MIS Data Warehouse will not have access to this data. D. None - No Federal Agencies, outside IRS, are providing data. E. None – No State or Local Agencies provide data.

	F. None – No other third party sources are providing data.
3. Is each data item required for the business purpose of the system? Explain.	All of the data elements identified in the systems outlined in the response to #2 above, are needed, for management information purposes. The MIS Data Warehouse contains relevant data from the systems listed above as well as data from future data sources as needed. Specific data elements are described in the Data Definition Document for each data source.
4. How will each data item be verified for accuracy, timeliness, and completeness?	The MIS Data Warehouse will contain extracts from existing IRS legacy systems (see 2. above), each of which have their own internal validation and verification processes. This provides the ability to compare the actual data extract to the data contained in the MIS Data Warehouse for verification when required.
5. Is there another source for the data? Explain how that source is or is not used.	No. All sources have been identified in this document. There are no other sources for the information being warehoused, however, this does not exclude the need for other future data sources. If future sources of data are introduced in the Data Warehouse an Updated or New PIA will be submitted to the Privacy Advocate Office.
6. How will the data generally be retrieved by the user?	The MIS Data Warehouse accumulates and stores the data into two databases. The MIS_DW database (non-sensitive data) and the MIS_DW_L3 database (sensitive data). Individual users will not have access to actual data stored in the MIS Data Warehouse and access will be limited to a Web Interface that only provides high-level information about the data sources, status of data loads, range of information available, data schemas, etc. This high level data will allow developers or analysts to determine if a particular data source is needed for a Data Mart or other IRS Internal project. The following business rules and design features are specific to the MIS Data Warehouse • Outside access to data is not permitted. • All authorized data transfers will be accomplished via a data "push" initiated by the MIS Data Warehouse. Data "pulls" are not permitted. • A "user" of data from the MIS Data Warehouse is

	referred to as a "subscriber". • A subscriber is an authorized IRS system, Data Mart, IRS Developed Application, etc., never an individual or IRS employee. • Current MISDW Subscribers and their certification status are; This system will use the e-authentication schema to determine the subscriber system's access level: L1 – lowest access concern L2 – next highest access designation, Sensitive but Unclassified (SBU) data. L3 – Highly Sensitive, Classified (Need to know basis) data L4 – Highest level of access control (national security issues, etc) Business Measures Data Mart (BMDM)- L1 complete Business Performance management System (BPMS)- L2 complete Submission Processing Measures Analysis Reporting Tool (SMART)- L1 complete Program for Automated Work Plan System (PAWS)- L1 completed E-File Reports-L1 certification almost complete Integrated Financial System (IFS) - Application and Certification in progress but RIS requesting data has been received. • Subscribers can only receive data equal to or below the level of security approved for the subscriber. For example, a Data Mart with a Level 2 security certification may not receive sensitive data unless all sensitive information is first removed during the data transfer process. • Subscribers of sensitive data must submit a signed copy of their Level 3 or higher security certification. This document will be reviewed and the expiration date will be validated and used by the MISDW to automatically terminate the subscription process; therefore, it will be the responsibility of the subscriber to receive timely re-certifications to prevent the disruption of their data transmission.
7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?	Only a subscriber with the appropriate security certification can receive this type of information if it is available within a data source stored. The current RIS process will be used to request and approve requests

	for sensitive data. See Attachment 2, which is a live example of a response to a RIS requesting sensitive data.

Access to the Data

8. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Others)?	System Administrators will have access to the data in the MIS Data Warehouse. Authorized subscribers will receive data via a data "push" from the MISDW. Other IRS users will be able to see high-level information about the data stored in the MISDW via a Web Interface. See item #6 above.
9. How is access to the data by a user determined and by whom?	A Subscriber can receive data from the MISDW after applying for and receiving approval. See # 6 above. Subscribers can receive approval for non-sensitive data via a request for subscription that contains a valid business reason justifying the need. This request must be made via the official RIS process. In order to receive sensitive information, subscribers must provide proof of certification equal to or greater than Level 3 in order to receive "unscrubbed" data. A subscriber that does not have a Level 3 certification may only receive "scrubbed" data (data in which all sensitive information types have been removed) from a sensitive data source if a valid business reason is submitted and approved. See item # 6 above for additional information concerning requirements for subscribers of sensitive data.
10. Do other IRS systems provide, receive, or share data in the system? If YES, list the system(s) and describe which data is shared. If NO, continue to Question 12.	Yes - Currently, the MIS Data Warehouse loads available data, using output files, from existing IRS and legacy systems. IRS systems that provide data to the MIS Data Warehouse are: • Audit Information Management System (AIMS) • Summary Examination Time Transmission System (SETTS) • Work Planning and Control (WP&C) • Automated Substitute for Return (ASFR) • Collection Activity Reports (CAR) • Quality Review Database (QRDB)

11. Have the IRS systems described in Item 10 received an approved Security Certification and Privacy Impact Assessment?	• AIMS is Actually • Summary Examination Time Transmission System (SETTS) is certified under the Examination Return Central System (ERCS) • WP&C is actually National Work Planning and Control System (NWP&CS), and has been certified. • Automated Substitute for Return (ASFR) is certified. • Quality Review Database (QRDB)is certified under the Balance Measures Applications System. • Collection Activity Reports (CARS) certification under the IDRS certification.
12. Will other agencies provide, receive, or share data in any form with this system?	No – Other agencies will not share or have access to the MIS Data Warehouse.

Administrative Controls of Data

13. What are the procedures for eliminating the data at the end of the retention period?	The procedures vary according to data type (i.e., records, backup sets, printouts, etc.). The guidelines are contained in IRM 1.15.1 and IRM 1.15.2. (See attachment3). The MISDW contains meta-data that specifies the retention period for each data source. On a monthly basis, an automated process is run which references the retention period and automatically deletes data which is beyond this period. The MISDW will retain MIS Type data for a period of 10 years.
14. Will this system use technology in a new way? If "YES" describe. If "NO" go to Question 15.	No. The system is not using technologies in ways that the IRS has not previously employed. The MIS Data Warehouse accumulates and stores data from pre-existing IRS and legacy systems.
15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.	No - This system does not have the capability to identify or locate individuals or groups of people. However, it is possible that a subscriber of MISDW data can perform this type of activity or other Data Mining type activities. The MIS Data Warehouse is a data repository only and can not mandate nor specify how data will be utilized. This would need to be addressed in the subscribers own security certification documentation and is beyond the scope of this application.

16. Will this system provide the capability to monitor individuals or groups? If yes, describe the business purpose for this capability and the controls established to prevent unauthorized monitoring.	No - This system does not have the capability to monitor individuals or groups of people. The MIS Data Warehouse is a data repository.
17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently? Explain.	No – This system does not allow IRS to treat taxpayers, employees or others differently. The MIS Data Warehouse will not affect the equitable treatment of taxpayers/employees.
18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?	The MIS Data Warehouse has no negative effects on the due process rights of taxpayers or employees. The MIS Data Warehouse is a data repository for MIS information, not a case processing or case management tool.
19. If the system is web-based, does it use persistent cookies or other tracking devices to identify web visitors?	No – The Web interface does not utilize persistent cookies or other tracking devices to identify web visitors. The MIS Data Warehouse is an internal (Intranet) resource available only within the Treasury Firewall for IRS employees (See #7 above). The Web Interface will have the capability to record and audit employee activity via their NT Login, however, the MIS Data Warehouse will not have the functionality to provide the identity of a specific NT Login. This type of logging is required for a Level 3 system containing sensitive information. Any type of auditing or logging will meet all appropriate requirements for security and internal IRS Web applications.