Risk Based Scoring System Version 2.0 (RBSS v 2.0) – Privacy Impact Assessment

PIA Date – February 18, 2009

<u>**System Overview**</u>
The Risk Based Scoring System Version 2.0 (RBSS v2.0) is a minor application which is owned and operated by the Internal Revenue Service (IRS) Wage and Investment (W&I) Division. RBSS v2.0 is a collection of Commercial-off-the-shelf (COTS) applications designed to aid Reporting Compliance (RC) in scoring and ranking tax returns based upon the information (e.g., name, Social Security Number (SSN), deductions and all relevant tax return information) provided. The purpose of this ranking and scoring is to match potentially improperly filed returns to the most appropriate treatment and to identify those returns (cases) that can be handled through less burdensome and lower-cost activities. RBSS v2.0 utilizes Negative Taxpayer Identification Number (TIN) checking to ensure that there is no improper viewing of ones own tax return information or related data in the application database.

<u>**Systems of Records Notice (SORN):**</u>

- IRS 42.021--Compliance Programs and Projects Files
- IRS 34.037--IRS Audit Trail and Security Records

<u>**Data in the System**</u>

1. Describe the information (data elements and fields) available in the system in the following categories:

   A. Taxpayer – The following taxpayer data is available in the system:
- Mortgage Insurance Premiums;
- Mortgage Amount; Points Paid;
- Refund Overpaid Amount;
- Student Loan Amount; Qualified Dividends;
- Capitol Gains;
- Foreign Taxes Paid;
- Investment Expenditures;
- Interest;
- Savings Bonds; Attorney Fees; Excess Golden Parachute;
- Other Income;
- Rents;
- Royalties;
- Section 409A Referrals;
- Section 409A Income;
- Dependent Care;
- Medicare Wages;
- Taxable Federal Insurance Contributions Act (FICA) Tips;
- Taxable FICA Wages;
- Type of Employment;
- Spouse TIN;
- Combat;
- Tax Year;
- TIN;

- Name Control;
- Master File Tax (MFT) Code;
- Zip Code;
- Date of Birth;
- Dependents SSNs;
- Casualty;
- Routing Transit Numbers;
- Adjusted Income;
- Deposit Amounts;
- Distribution Date;
- Extension Date;
- Transaction Codes;
- Preparer Code;
- Transaction Date;
- Tax Class;
- Estimated Tax;
- Repairs;
- Risk;
- Travel;
- Gas Fuel;
- Tax Payer First Name;
- Tax Payer Second Name; Pension;
- Adjusted Gross;
- Adjusted Income;
- Child Care Credit;
- Child Tax;
- Taxable Income;
- Total Itemized Deductions; and
- Description of Business or Profession.

B. Employee – Data which will be collected from employees during authentication includes:
- Standard Employee Identifier (SEID)
- Data base Standard User Name Convention
- Unix Standard User Name Convention.

C. Audit Trail Information – Includes:
- User Name;
- Specific Audit Trail Information includes the following:
- User actions – (Generated By Business Objects Server (Central Management Server (CMS));
  - A named user logon succeeds;
  - A user logon fails;
  - A user's password is changed;
  - User logs off.

2. Describe/identify which data elements are obtained from files, databases, individuals, or any other sources.

A. IRS – RBSS v2.0 obtains data from the following systems:
- Integrated Production Model (IPM),

- Earned Income Tax Credit (EITC),
- Data Master File 1 (DM1).

RBSS v2.0 obtains data from IPM and the following tables within IPM:
- Information Returns Master File (IRMF)
- Individual Returns Transaction File (IRTF).

Data elements that are obtained from the above sources (systems and tables within systems) include:
- Mortgage Insurance Premiums; Mortgage Amount;
- Points Paid;
- Refund Overpaid Amount;
- Student Loan Amount;
- Qualified Dividends;
- Capitol Gains;
- Foreign Taxes Paid;
- Investment Expenditures;
- Interest;
- Savings Bonds;
- Attorney Fees;
- Excess Golden Parachute;
- Other Income;
- Rents; Royalties;
- Section 409A Referrals;
- Section 409A Income;
- Dependent Care;
- Medicare Wages;
- Taxable FICA Tips;
- Taxable FICA Wages;
- Type of Employment;
- Spouse TIN;
- Combat;
- Tax Year;
- TIN;
- Name Control;
- MFT Code;
- Zip Code;
- Date of Birth;
- Dependents SSNs;
- Casualty;
- Routing Transit Numbers;
- Adjusted Income;
- Deposit Amounts;
- Distribution Date;
- Extension Date;
- Transaction Codes;
- Preparer Code;
- Transaction Date;
- Tax Class;

- Estimated Tax;
- Repairs;
- Risk; Travel;
- Gas Fuel;
- Tax Payer First Name;
- Tax Payer Second Name; Pension;
- Adjusted Gross;
- Adjusted Income;
- Child Care Credit;
- Child Tax;
- Taxable Income;
- Total Itemized Deductions; and
- Description of Business or Profession.

3. Is each data item required for the business purpose of the system? Explain.
Yes. Each data item is required for the business purpose of the system. RBSS v2.0 is a data mining system. Its purpose is to discover which data is useful in modelling case selection outcomes. All of the data must be in the system in order to determine if it has any predictive value in the data mining process. In addition, the privacy data is used to link all the tables and to push the results to the Transaction Code (TC) 424 Workload Management System (WMS) for the assigning of cases. Furthermore, the data is needed for reporting purposes in order to determine which cases are still open.

4. How will each data item be verified for accuracy, timeliness, and completeness?
Data will be pulled from the Integrated Production Model (IPM) based upon the date the return was posted to Master File. Only the data that falls within the needed time period will be used by RBSS v2.0. RBSS v2.0 will obtain data directly from other IRS systems. RBSS v2.0 will rely on these other systems to verify the data items for accuracy, timeliness, and completeness. Timeliness is required as the initial prototype of the system will operate on pre-refund data, where strict time limits are enforced by policy.

5. Is there another source for the data? Explain how that source is or is not used.
No. RBSS v2.0 has no other sources for the data.

6. Generally, how will data be retrieved by the user?
The IRS user will login onto their IRS machine. The user will apply for a Business Objects username/password via the On Line 5081 process. The user needs to use this username and password to authenticate themselves when launching Business Objects. Business Objects is a web-based reporting tool that will allow a user to generate reports against IPM and RBSS 2.0 databases.

7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?
Yes. RBSS v2.0 data is retrievable by the TIN, which is used as a unique identifier by the system. Retrieval of cases using the TIN is necessary since tax records use the TIN as the principle identifier for accounting purposes. It is the only way to guarantee the proper identification of the taxpayer.

Access to the Data

8. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Others)?

The following IRS personnel will have access to data in the system: Enterprise Operations (EOPS) System Administrators, EOPS Database Administrators, Compliance Analysts, Compliance Managers, Model Developers, Report Developers and Security Auditing and Analysis System (SAAS) Auditors. The table below identifies the role assignments that have been established for RBSS v2.0 users:

Role: Clementine User (Compliance Analysts and Managers)
Permission: Create reports, graphs, statistics, scripts

Role: Clementine User Administrator
Permission: Issues user name, password

Role: Predictive Enterprise Services (PES) Group Administrator
Permission: Create, modify, delete groups; Create, modify, delete roles; Assign members to both Groups and Roles

Role: PES Model Developer
Permission: Creates and schedule jobs, checks models in and out of PES; Creates folders in the content repository

Role: PES User (Compliance Analysts and Managers)
Permission: Can check in and out models, but has read-only access to models; Read only access to schedules and jobs

Role: Business Objects Administrator
Permission: Set-up group permissions, monitor and tune system performance

Role: Business Objects Universe Designer
Permission: Creates universes

Role: Business Objects Report Developer
Permission: Creates new reports for public

Role: Business Objects Power User
Permission: Refresh, modify existing reports; Update existing reports in Personal area only

Role: Business Objects User (Compliance Analysts and Managers)
Permission: View only; Read-only access to existing reports

Role: Oracle Model Developer
Permission: Select, Insert, Update, Delete and Execute

Role: Oracle Report Developer
Permission: Select, Insert, Update, Execute

Role: SAAS Auditors
Permission: Generate and review application system audit reports
Only IRS personnel will hold roles and access to data in the system. Contractors do not hold any roles or have access to the data in the system.

9. How is access to the data by a user determined and by whom?
Access to the data is determined by the manager based on a user's position and need-to-know. The manager will request a user be added. They must fill out an On Line 5081, Information System User Registration/Change Request, to request access to the application. A user's access to the data terminates when it is no longer required. The user will validate the need of access during their annual OL5081 review. The administrative will perform quarterly reviews of users to determine if user requires access to the system.

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.
Yes. RBSS v2.0 will receive data from the IPM, DM1 and EITC. RBSS v2.0 receives data from IPM and the IRMF and IRTF tables within IPM.

A. Data elements that are received from the above sources (systems and tables within systems) include:
   - Mortgage Insurance Premiums;
   - Mortgage Amount;
   - Points Paid;
   - Refund Overpaid Amount;
   - Student Loan Amount;
   - Qualified Dividends;
   - Capitol Gains;
   - Foreign Taxes Paid;
   - Investment Expenditures;
   - Interest;
   - Savings Bonds;
   - Attorney Fees;
   - Excess Golden Parachute;
   - Other Income;
   - Rents;
   - Royalties;
   - Section 409A Referrals;
   - Section 409A IPM will be accessed using a private Database (DB) link between the Oracle database of RBSS and IPM.

B. RBSS v2.0 shares and provides data to the TC 424 WMS. Data shared and provided includes:
   - TIN;
   - TIN Type;
   - Name Control;
   - Tax Period;
   - Preparer ID;
   - Fraud Detection Control (FDC) Number;
   - Document Locator Number (DLN);
   - AGI;
   - Project ID;
   - Service Center;
   - Chargeouts/Labels;
   - Source Code;
   - Project Code;
   - Automated Computer Information System (AIMS) Assignee Code;

- Business Operating Division (BOD);
- Pre/Post Refund Indicator;
- 1040X Claim Amount;
- Gross 10% IRA Amount;
- AIMS Tracking Code; and
- Score

C. RBSS v2.0 will provide data via secure File Transfer Protocol (FTP) using Enterprise File Transfer Utility (EFTU).

11. Have the IRS systems described in Item 10 received an approved Security Certification and Privacy Impact Assessment?
Yes.

TC424 WMS:
- Certification & Accreditation (C&A) – May 31, 2006, expires on May 31, 2009
- Privacy Impact Assessment (PIA) – April 20, 2006, expires on April 30, 2007

IPM:
- Certification & Accreditation (C&A) – November 9, 2007, expires on November 9, 2010
- Privacy Impact Assessment (PIA) – November 16, 2007, expires on November 16, 2008

SAAS:
- Certification & Accreditation (C&A) – June 16, 2007, expires on June 16, 2010
- Privacy Impact Assessment (PIA) – September 6, 2006, expires on September 6, 2007

EAIB:
- Certification & Accreditation (C&A) – July 15, 2007, expires on July 15, 2010
- Privacy Impact Assessment (PIA) – July 25, 2006, expires on July 25, 2007

DEPDB:
- Certification & Accreditation (C&A) – March 17, 2009, expires on March 17, 2012
- Privacy Impact Assessment (PIA) – February 20, 2009, expires on February 20, 2010

CDW:
- Certification & Accreditation (C&A) – May 6, 2009, expires on May 6, 2012
- Privacy Impact Assessment (PIA) – March 24, 2008, expires on March 24, 2009

12. Will other agencies provide, receive, or share data in any form with this system?
No. Other agencies will not provide, receive, or share data in any form with this system.

Administrative Controls of Data

13. What are the procedures for eliminating the data at the end of the retention period?
Data will be retained for five years. When the retention period expires, the procedures in Internal Revenue Manual (IRM) 1.15.3.1–1 in the Record Disposition Handbook will be followed to eliminate the expired data. RBSS will create a record control schedule which will use the Standard Form 115–Request for Record Disposition Authority. This control will be establish to determine if the records should be retired to Federal Record Center or destroyed.

14. Will this system use technology in a new way?
No.

15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.
Yes. RBSS v2.0 is a data mining system. Its purpose is exactly to identify individuals who meet criteria that are shown to result in optimal use of Reporting Compliance resources based on their shared characteristics.

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.
Yes. The business purpose RBSS v2.0 is to assign scores to individuals ranking their appropriateness for various reporting compliance treatment streams. It is crucial that the output of the modelling and data mining process be monitored over time to continuously improve the results. Unauthorized monitoring will be prevented by strong access control restrictions to the system itself, and also through non-technology-based business processes and standard IRS management practices.

17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently?
No.

18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?
Not applicable. RBSS v2.0 does not interface with taxpayers, but supplies information to W&I Compliance.

19. If the system is web-based, does it use persistent cookies or other tracking devices to identify web visitors?
No. RBSS v2.0 is web-based; however, it is not using persistent cookies or other tracking devices to identify web visitors. RBSS v2.0 does not interconnect to any system or application external to the IRS.

[View other PIAs on IRS.gov](#)