

Dependent Database (DEPDB) – Privacy Impact Assessment

PIA Approval Date – October 17, 2011

System Overview

Dependent Database (DEPDB) addresses non-compliance relevant to the Earned Income Tax Credit (EITC) and other tax benefits related to the dependency and residency of children. The DEPDB project fulfills and addresses two congressional mandates. The first legislative mandate is the Kohl amendment to the Taxpayer Relief Act of 1997, Section 1090 which authorized the IRS to access data from the Federal Case Registry (FCR). The second mandate is Section 1085 of the Taxpayer Relief Act of 1997, which was established to reduce non-compliance relative to the Earned Income Tax Credit. This special appropriation provided additional resources to eliminate fraud and abuse relevant to EITC. To consistently apply the tax laws to a return claiming EITC, other tax issues had to be addressed at the same time, such as, dependent exemptions, filing status, Child and Dependent Care Credit, Child Tax Credit, and education benefits. The Dependent Database Project examines EITC tax returns and applies a set of rules, Fair Isaac Model and SRA Clementine model to determine residency and relationship issues. Tax returns are examined in a pre-refund environment. This means the money is stopped before going out the door. In the past, the IRS would send out the refund and then work the cases in a post-refund environment. In many cases, once the money is gone it becomes very difficult to get it back. The DEPDB system incorporates data (Department of Health and Human Services (HHS), Social Security Administration (SSA) & IRS) to more accurately validate refunds entitled to a taxpayer; thus allowing the IRS to enforce laws passed by Congress more effectively.

Systems of Records Notice (SORN):

- IRS 42.021--Compliance Returns and Project Files
- IRS 34.037--IRS Audit trail and security records system

Data in the System

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

- A. Taxpayer – via the Individual Returns Transaction File On-line (IRTFOL) (subsystem of Individual Masterfile (IMF)) system which includes, but is not limited to the following information contained on tax returns:
 - Names (e.g., taxpayer, dependent spouse)
 - Social Security Numbers (SSNs) (e.g., taxpayer, dependent, spouse)
 - Dependents date of birth (DOB)
 - Address
 - Prior and current tax year tax return schedule
 - Various income deductions and credits
 - Filing status
- B. Audit Trail Information – DEPDB relies on Resources Access Control Facility (RACF) auditing capabilities. At a minimum, an event record shall specify the following:
 - Logon and logoffs of the system
 - Password changes

- All systems administrator (SA) actions while logged on to the system as an SA and as a user
- Switching accounts or running privileged actions from another account
- Stored and Ad hoc queries
- Data files opened and closed
- Specific actions, such as reading, editing
- Deleting records or fields, and printing reports

C. Other –Tax Practitioner’s SSN/Employer Identification Number (EIN)

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

A. IRS – DEPDB obtains the following data elements from IRTFOL (Subsystem of IMF), **Integrated Data Retrieval System (IDRS), Individual Master File On–Line (IMFOL) (Subsystem of IMF) and Generalized Mainline Framework (GMF)** that includes, but is not limited to:

- Names (e.g., taxpayer, dependent spouse, tax practitioner)
- Social Security Numbers (SSNs) (e.g., taxpayer, dependent, spouse tax practitioner)
- Dependents date of birth (DOB)
- Employer Identification Number (EIN)
- Address
- Prior and current tax year tax return schedule
- Tax return schedule
- Various income deductions and credits
- Filing status

Duplicate Tax Identification Number (TIN) Database (DUP TIN): Dependent’s primary and secondary SSN

IBM Masterfile Platform (MITS–21GSS): Department of Health and Human Services (HHS) information, such as:

- Custodial orders for children receiving public assistance
- Custody orders involving private divorce cases,
- Name and SSN of custodial party
- Name and SSN of non–custodial party or putative father

The data is stored on an Automated Tape Library (ATL) Storage device via CONNECT:Direct on the MITS–21 GSS mainframe. The Office of Child Support Enforcement (OCSE) transmits to the IRS data from the Federal Case Registry (FCR) by CONNECT:Direct once a year in December. DEPDB connects to the ATL storage device via the mainframe. This information is then extracted from the mainframe during its daily scheduled routine.

National Account Profile (NAP): DEPDB receives the Data Master–1 (DM–1) file, which contains data obtained from the Social Security Administration (SSA). DEPDB obtains the following SSA information from DM–1:

- SSN
- Date of Birth (DOB)
- Date of Death (DOD)
- Citizenship indicators

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. DEPDB use the information to identify potential erroneous Earned Income Tax Credit claims and duplicate dependency exemptions claimed on current year filed tax returns. Returns are scored and selected for examination.

4. How will each data item be verified for accuracy, timeliness, and completeness?

The data received by DEPDB is verified by the various applications as being complete and accurate prior to being transmitted to DEPDB. Additionally, DEPDB schema is configured in accordance with its data sources; thus the data, when it is retrieved via batch processing, will automatically load in the right format. Data transferred via flat files to the IBM system are processed and verified by the 793–01 router run before the data are retrieved by the DEPDB system. DEPDB also perform Log Analysis and Reporting Services (LARS) Counters to verify that all records coming into a particular DEPDB batch job are processed/ accounted for. For instance, the record count of the GMF file that is received on the IBM mainframe is verified against the record count of the file that was sent from the Unisys system.

5. Is there another source for the data? Explain how that source is or is not used.

No. There is no other source for the data.

6. Generally, how will data be retrieved by the user?

The data is retrieved by actual selection of the returns for examination. The IDRS end user access the data via the following secured command codes from the IDRS screen:

- Dependent Database Kid (DDBKD)
- Dependent Database On–line (DDBOL)

Once the user entered the DDBKD or DDBOL command code, they will then input the child’s SSN information to obtain further information.

7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?

Yes. The data are retrievable by entering the command codes: DDBKD or DDBOL from the IDRS users’ screens. The following are the detail on the data that will displayed based on the information entered:

Once the user entered the DDBKD command code, and then input the child’s SSN information, the user be provided with the following information on his screen:

- Date of Birth
- Date of Death
- DM1 update date
- Citizenship code
- Current SSA name control
- Case type
- Custody parent’s SSN
- Non custody parent’s SSN
- Putative father’s SSN
- State code
- Case number
- Birth mother’s SSN and whether SSA verified the SSA
- Birth father’s SSN and whether SSN was verified by SSA
- Child’s current name control

- Kidlink record update date
- The primary SSN of the return that this SSN appeared on
- Use of SSNs on return (as a primary, secondary, dependent and/or qualifying child)

Once the user entered the DDBOL command code, and then input the primary taxpayer's SSN along with the tax processing year, the user be provided with the following information on his screen:

- The DEPDB rule that broke
- Examination information – tax change, EITC change, date closed, etc
- SSN(s) of child/children claimed for Earned Income Tax Credit (EITC), exemption, Child Tax Credit, & Child Care

Access to the Data

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

Only Database Administrators can access the DEPDB system directly. However, Exam Personnel, IDRS users, have access to the information that is used in the DEPDB system through command codes DDBKD and DDBOL. These users have read only access. No contractors have access to the DEPDB system.

Role: Database Administrators

Permission: Read, write, change, delete

Role: Exam Personnel (IDRS users)

Permission: Read

9. How is access to the data by a user determined and by whom?

Access to the DEPDB data is only accessible by IRS employees with IDRS accounts. These IDRS users are EXAM personnel. The employees must have an active IDRS account and have command code capabilities enabled in their profiles in order to access DEPDB data. The managers and/or IDRS Security Representatives enable this feature in the user's profile via the IDRS Online Report Services (IORS). Access to the data is determined by the IDRS user's manager based on a user's position and need-to-know. IDRS users, via the command codes, are restricted to read only access.

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, DEPDB obtains the following data elements from Individual Returns Transaction File On-line (IRTFOL) (subsystem of IMF), Integrated Data Retrieval System (IDRS), Individual Master File On-Line (IMFOL)(subsystem of IMF) and Generalized Mainline Framework (GMF) that includes, but is not limited to:

- Names (e.g., taxpayer, dependent spouse, tax practitioner)
- Social Security Numbers (SSNs) (e.g., taxpayer, dependent, spouse tax practitioner)
- Dependents date of birth (DOB)
- Employer Identification Number (EIN)
- Address
- Prior and current tax year tax return schedule
- Various income deductions and credits
- Filing status

Duplicate TIN On-Line (DUP TIN):

- Dependent's primary and secondary SSN

National Account Profile (NAP)

- Taxpayer name
- SSN
- Address
- Filing status
- DOB
- Social Security Administration (SSA) database (DM-1) data – SSN, date of birth (DOB) DOD and citizenship indicators

DEPDB shares the following data with Report Generation Software (RGS) and IDRS:

- Names (e.g., taxpayer, dependent spouse, tax practitioner)
- Social Security Numbers (SSNs) (e.g., taxpayer, dependent, spouse tax practitioner)
- Dependents date of birth (DOB)
- Employer Identification Number (EIN)
- Address
- Prior and current tax year tax return schedule
- Various income deductions and credits
- Filing status

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

Yes.

Integrated Data Retrieval System (IDRS)

- Authorization to Operate (ATO) – March 20, 2009
- Privacy Impact Assessment (PIA) – November 6, 2008

Individual Returns Transaction File On-line (IRTFOL) (subsystem of IMF)

- Authorization to Operate (ATO) – March 8, 2010
- Privacy Impact Assessment (PIA) – November 10, 2009

Individual Master File On-Line (IMFOL) (subsystem of IMF)

- Authorization to Operate (ATO) – March 8, 2010
- Privacy Impact Assessment (PIA) – November 10, 2009

Duplicate TIN On-Line (DUP TIN):

- ATO and PIA (DUP TIN is a non-reportable application and thus there is no ATO or PIA dates identified)

National Account Profile (NAP)

- Authorization to Operate (ATO) – February 13, 2009
- Privacy Impact Assessment (PIA) – March 23, 2010

Report Generation Software (RGS)

- Authorization to Operate (ATO) – May 7, 2010
- Privacy Impact Assessment (PIA) – April 23, 2009

Generalized Mainline Framework (GMF)

- Authorization to Operate (ATO) – February 18, 2009

- Privacy Impact Assessment (PIA) – October 16, 2008

12. Will other agencies provide, receive, or share data in any form with this system?

DEPDB receives HHS data indirectly. The data is retrieved from the MITS–21 GSS mainframe on a daily scheduled routine. DEPDB also indirectly receives SSA data. The data is retrieved from the DM–1 file located on the NAP system.

Administrative Controls of Data

13. What are the procedures for eliminating the data at the end of the retention period?

The records in DEPDB are scheduled under IRM 1.15.29 Records Control Schedule for Service Center Operations, Item 417 (NARA Job No. N1–58–07–4, approved 5/14/07). System data is approved for destruction, retention periods vary.

14. Will this system use technology in a new way?

No. This system will not use technology in a new way.

15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.

No. DEPDB will not be used to identify or locate individuals or groups.

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. DEPDB will not provide the capability to monitor individuals or groups.

17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently?

No. The use of DEPDB will not allow IRS to treat taxpayers, employees, or others, differently.

18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?

Not applicable. DEPDB is not used to make negative determinations.

19. If the system is web–based, does it use persistent cookies or other tracking devices to identify web visitors?

Not applicable. DEPDB is a batch process and does not interface with the web, therefore no cookies are persistent.

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