



## Treasury Inspector General for Tax Administration Office of Audit

### **CUSTOMER ACCOUNT DATA ENGINE 2 DATABASE VALIDATION IS PROGRESSING; HOWEVER, DATA COVERAGE, DATA DEFECT REPORTING, AND DOCUMENTATION NEED IMPROVEMENT**

Issued on September 29, 2014

## Highlights

Highlights of Report Number: 2014-20-063 to the Internal Revenue Service Chief Technology Officer.

### **IMPACT ON TAXPAYERS**

There is significant effort underway to ensure the accuracy of individual taxpayer account data on the Customer Account Data Engine 2 (CADE 2) database. This effort is an important part of its implementation because inaccurate data could delay this database from becoming the authoritative source of data, thereby increasing the cost of implementation.

### **WHY TIGTA DID THE AUDIT**

This review was part of our Fiscal Year 2014 Annual Audit Plan and addresses the major management challenge of Modernization. The overall audit objective was to evaluate IRS efforts to ensure that the data in the CADE 2 database are accurate and complete.

The IRS requested that TIGTA evaluate the new data validation testing methodology. TIGTA performed this audit during the data validation testing process and provided the IRS with recommendations for continuous improvement.

### **WHAT TIGTA FOUND**

Data validation efforts were efficiently performed due to adequate planning and resource coordination. For example, detailed data validation plans ensured that test activities were on track and a new process ensured that data defects were effectively managed.

The IRS identified the data fields to be verified and how each would be validated. While a large percentage of the data fields are validated with automated data compare tools, there is no documented plan to ensure that data fields validated using other means are validated periodically. The data sampling methodology for validating CADE 2 data is sound. The IRS developed a data sampling methodology to enable maximum data validation coverage by using a statistical sample, but key activities were not documented. After discussing the

need to document the data sampling methodology, the IRS began development of the documentation. Several in-progress documents were provided for our review.

The IRS developed a Data Quality Scorecard to track progress in meeting data quality success criteria. However, the processes needed to effectively perform these activities were not sufficiently documented. As a result, some of the metrics were initially incorrectly reported.

TIGTA also found that data discrepancy reports needed improvement. Our analysis found that 10 data field identifiers were missing from a discrepancy report. As a result, the IRS is using another automated tool to validate the 10 data field identifiers until the main tool is corrected.

### **WHAT TIGTA RECOMMENDED**

TIGTA recommended that the Chief Technology Officer ensure that: 1) data validation test results are maintained and available for data fields not validated by automated data compare tools; 2) data validation plans include periodically validating the data fields that are not validated with automated data compare tools; 3) all data sampling processes are completely documented; 4) details needed for determining the Data Quality Scorecard metrics are completely documented; 5) all documentation needed to verify the data in the Data Quality Scorecard is stored for future reference; 6) automated data compare tools identify and report on data fields, not field identifier numbers; and 7) automated data compare tool reports clearly identify counters and align with data validation metrics.

The IRS agreed with six of the report's seven recommendations. The IRS plans to maintain results for manual data validation activities, validate changes to the data fields that are not validated with automated data compare tools, develop documentation on the procedures to collect and maintain data used to support data validation metrics and the Scorecard development process, and store Scorecard source documentation.

### **READ THE FULL REPORT**

To view the report, including the scope, methodology, and full IRS response, go to:

<http://www.treas.gov/tigta/auditreports/2014reports/201420063fr.pdf>.

E-mail Address: [TIGTACommunications@tigta.treas.gov](mailto:TIGTACommunications@tigta.treas.gov)

Phone Number: 202-622-6500

Website: <http://www.treasury.gov/tigta>