

Date of Approval: **November 20, 2020**

PIA ID Number: **5408**

## **SYSTEM DESCRIPTION**

*Enter the full name and acronym for the system, project, application and/or database.*

Big Data Analytics, BDA

*Is this a new system?*

No

*Is there a PCLIA for this system?*

Yes

*What is the full name, acronym and milestone of the most recent PCLIA?*

2786 Big Data Analytics, BDA, approved

*What is the approval date of the most recent PCLIA?*

11/3/2017

*Changes that occurred to require this update:*

Expiring PCLIA

*Were there other system changes not listed above?*

No

*What governance board or Executive Steering Committee (ESC) does this system report to? Full name and acronym.*

ES Governance Board (ESGB) and Infrastructure Executive Steering Committee (IESC)

*Current ELC (Enterprise Life Cycle) Milestones:*

Operations & Maintenance (i.e. system is currently operational)

*Is this a Federal Information Security Management Act (FISMA) reportable system?*

Yes

## GENERAL BUSINESS PURPOSE

*What is the general business purpose of this system? Provide a clear, concise description of the system, application or database, the reason for the system, and the benefits to the IRS to use the information, and how the information will be used.*

Big Data Analytics (BDA) is an appliance-based infrastructure that provides the ability to conduct advanced analytics, low latency data processing, and in-depth analysis of data. BDA performs high-speed data processing that facilitates audit selections, analyzing taxpayer filings, and more. BDA is a Technology Shared Service (TSS). Projects share resources on the BDA platform to satisfy business requirements. BDA manages the Infrastructure (hardware and software) while projects manage their input sources, data loads, verification and validation processes and methods and processes for dissemination. BDA does not process data. BDA does not manage project work. BDA provides a secure Infrastructure platform where projects can load, process and distribute SBU/PII and business data. Each project manages their data and data needs independently. Projects on the BDA platform include Return Review Program (RRP), Integrated Production Model (IPM), Customer Account Data Engine 2 Data Store (CADE2 ODS), Business Master File Case Creation Nonfiler Creation Database (BMF CCNIP) and Information Sharing and Reporting (IS&R).

## PII DETAILS

*Does the system use, collect, receive, display, store, maintain, or disseminate IR Code 6103 taxpayer information; or any other type of Sensitive but Unclassified (SBU) information or PII such as information about IRS employees or outside stakeholders?*

Yes

*Does the system use, collect, receive, display, store, maintain, or disseminate Social Security Numbers (SSN's) or tax identification numbers (i.e. last 4 digits, etc.)?*

Yes

*What types of tax identification numbers (TIN) apply to this system?*

Social Security Number (SSN)

*List the approved Treasury uses of the SSN:*

Security Background Investigations

When there is no reasonable alternative means for meeting business requirements

Statistical and other research purposes

*Explain why the authorized use(s) above support the new or continued use of SSNs (or tax identification numbers).*

BDA does not use, collect, receive, display, maintain or disseminate SSNs. BDA provides a Technology Shared Service for projects to store SSNs. The projects on the BDA platform are responsible for their usage and management of SSNs.

*Describe the planned mitigation strategy and forecasted implementation date to mitigate or eliminate the use of SSN's (or tax identification numbers).*

BDA projects (Ex: Integrated Production Module (IPM), Return Review Program (RRP), Customer Account Data Engine (CADE2), Operational Data Store (ODS), Information Sharing and Reporting (IS&R) are responsible for mitigating or eliminating the use of SSNs in their data.

Employer Identification Number

Other Taxpayer Identification Number

*Does this system use, collect, receive, display, store, maintain or disseminate other (non-SSN) PII (i.e. names, addresses, etc.)?*

Yes

*Specify the PII Elements:*

Name

Mailing address

Phone Numbers

Date of Birth

Protection Personal Identification Numbers (IP PIN)

Financial Account Numbers

Tax Account Information

*Does this system use, collect, receive, display, store, maintain, or disseminate SBU information that is not PII?*

Yes

*Specify the types of SBU from the SBU Types List.*

**Agency Sensitive Information** Information which if improperly used or disclosed could adversely affect the ability of the agency to accomplish its mission

**Official Use Only (OUO) or Limited Official Use (LOU)** Information designated as OUO or LOU is information that: is exempt under one of the statutory Freedom of Information Act exemptions; is prohibited by other laws or regulations; would significantly impede the agency in carrying out a responsibility or function; or would constitute an unwarranted invasion of privacy.

**Protected Information** Information which if modified, destroyed or disclosed in an unauthorized manner could cause: loss of life, loss of property or funds by unlawful means, violation of personal privacy or civil rights, gaining of an unfair procurement advantage by contractors bidding on government contracts, or disclosure of proprietary information entrusted to the Government

*Are there other types of SBU/PII used in the system?*

No

*Cite the authority for collecting SBU/PII (including SSN if relevant).*

PII for federal tax administration is generally Internal Revenue Code Sections 6001, 6011, & 6012e(a)

SSN for tax returns and return information is Internal Revenue Code Section 6109

*Has the authority been verified with the system owner?*

Yes

## **BUSINESS NEEDS AND ACCURACY**

*Explain the detailed business needs and uses for the SBU/ PII, and how the SBU / PII is limited only to that which is relevant and necessary to meet the mission requirements of the system. If SSNs (or tax identification numbers) are used, explicitly explain why use of SSNs meets this criteria. Be specific.*

BDA Projects (Ex: IPM, RRP, CADE2 ODS, IS&R) utilize SBU/PII to support case identification, selection, prioritization, delivery, compliance analysis and decision analytics. IPM provides a centralized database, populated with current and historical return data from a variety of data sources. As a result of the analysis, case workloads are selected, notices may be generated, or transactions may be required to post back to the individual's tax account. IPM client-project applications performing analysis, and receiving data downloaded from the IPM BDA Greenplum database, use the taxpayer SSN as a major search and account matching key.

*How is the SBU/PII verified for accuracy, timeliness and completion?*

BDA projects (Ex: IPM, RRP, CADE2 ODS, IS&R) are responsible for managing the accuracy timeliness and completeness of SBU/PII data utilized. The client projects, such as IPM, have their own verification process for data accuracy, timeliness, completeness. BDA assumes that the data is accurate, timely, and complete when the client projects store the data on the Greenplum database.

## **PRIVACY ACT AND SYSTEM OF RECORDS**

The Privacy Act requires Federal agencies that maintain a system of records to publish systems of records notices (SORNs) in the Federal Register for records from which information is retrieved by any personal identifier for an individual who is a US citizen or an alien lawfully admitted for permanent residence. The Privacy Act also provides for criminal penalties for intentional noncompliance.

*Does your application or this PCLIA system pertain to a group of any record from which information is retrieved by any personal identifier for an individual who is a US citizen, or an alien lawfully admitted for permanent residence? An identifier may be a symbol, voiceprint, SEID, or other personal identifier that is used to retrieve information.*

Yes

*Identify the Privacy Act SORN(s) that cover these records.*

IRS 24.030 Customer Account Data Engine Individual Master File

IRS 42.021 Compliance Programs and Projects Files

IRS 24.046 Customer Account Data Engine Business Master File

IRS 34.037 Audit Trail and Security Records

IRS 22.061 Information Return Master File

IRS 22.026 Form 1042S Index by Name of Recipient

## **RESPONSIBLE PARTIES**

*Identify the individuals for the following system roles:*

## Official Use Only

## **INCOMING PII INTERFACES**

*Does the system receive SBU/PII from other systems or agencies?*

No

## **DISSEMINATION OF PII**

*Does this system disseminate SBU/PII?*

No

## **PRIVACY SENSITIVE TECHNOLOGY**

*Does this system use social media channels?*

No

*Does this system use privacy-sensitive technologies such as mobile, global position system (GPS), biometrics, RFID, etc.?*

No

*Does the system use cloud computing?*

No

*Does this system/application interact with the public?*

No

## **INDIVIDUAL NOTICE AND CONSENT**

*Was/is notice provided to the individual prior to collection of information?*

Yes

*How is notice provided? Was the individual notified about the authority to collect the information, whether disclosure is mandatory or voluntary, the purpose for which the information will be used, with whom the information may be shared, and the effects on the individual, if any, if they decide not to provide all or any of the requested information?*

The projects on the shared platform provide the notice. Notice is provided by the projects that are on this shared platform. The IRS notifies all individuals who file tax returns of such collection via the Privacy Act Notice in tax return instructions.

*Do individuals have the opportunity to decline from providing information and/or from consenting to particular uses of the information?*

Yes

*Describe the mechanism by which individuals indicate their consent choice(s):*

The projects on the shared platform provide this information. Please note that notice, consent and due process are provided in the tax forms instructions filed by the taxpayer, and pursuant to 5 USC. Or if gathered from tax form: The IRS has the legal right to ask for information per Internal Revenue Code sections 6001, 6011, and 6012(a), and their regulations.

*How does the system or business process ensure 'due process' regarding information access, correction and redress?*

The projects that are on this shared platform will allow affected parties the opportunity to clarify or dispute negative information that could be used against them. Due process is provided pursuant to 5 USC.

## **INFORMATION PROTECTION**

*Identify the owner and operator of the system (could be IRS owned and Operated; IRS owned, contractor operated; contractor owned and operated).*

IRS Owned and Operated

*The following people have access to the system with the specified rights:*

*IRS Employees*

Managers: Read Only

System Administrators: Read Write

Developers: Read Write

*IRS Contractor Employees*

Contractor System Administrators: Read Write

Contractor Developers: Read Write

*How is access to SBU/PII determined and by whom?*

Access can be requested through Online 5081. The perspective client applications must present a current SD-ATO/ATO and a Project Engagement Form (PEF) that details the project needs and the connection information. Anyone with a managerial approved Business need, with a current approved Live Data Waiver and managerial approved OL5081 can be granted access to the PII data hosted by projects on the BDA.

## **RECORDS RETENTION SCHEDULE**

*Are these records covered under a General Records Schedule (GRS, IRS Document 12829), or has the National Archives and Records Administration (NARA) approved a Records Control Schedule (RCS, IRS Document 12990) for the retention and destruction of official agency records stored in this system?*

No

*You must work with the IRS Records and Information Management (RIM) Program Office to address records retention requirements before you dispose of any records in this system.*

## **SA&A OR ASCA**

*Has the system been through SA&A (Security Assessment and Authorization) or ASCA (Annual Security Control Assessment)?*

Yes

*What date was it completed?*

5/4/2020

*Describe the system's audit trail.*

Enterprise Security Audit Trails (ESAT) completed and signed the Greenplum Platform Audit Plan April 2014. Audit records will capture all of the required elements contained with the IRS IRMs such as account logon, activity of admin users, as well as failed logon attempts.

## PRIVACY TESTING

*Does the system require a System Test Plan?*

Yes

*Is the test plan completed?*

Yes

*Where are the test results stored (or documentation that validation has occurred confirming that requirements have been met)?*

DocIT

*Were all the Privacy Requirements successfully tested?*

Yes

*Are there any residual system privacy, civil liberties, and/or security risks identified that need to be resolved?*

No

*Describe what testing and validation activities have been conducted or are in progress to verify and validate that the applicable Privacy Requirements (listed in header) have been met?*

Continuous Monitoring (eCM) (now called Annual Security Control Assessment (ASCA)) occurs annually to ensure that controls remain in place to properly safeguard PII.

## SBU DATA USE

*Does this system use, or plan to use SBU Data in Testing?*

Yes

*Was permission granted per the requirements of Form 14664, SBU Data Use Questionnaire or Form 14665, SBU Data Use Request?*

Yes

*Provide the date the permission was granted.*

10/19/2017

*Was testing performed in conformance with IRM 10.8.8 Information Technology (IT) Security, Sensitive But Unclassified (SBU) Data Policy?*

Yes

## **NUMBER AND CATEGORY OF PII RECORDS**

*Identify the number of individual records in the system for each category:*

IRS Employees: 50,000 to 100,000

Contractors: 5,000 to 10,000

Members of the Public: More than 1,000,000

Other: No

## **CIVIL LIBERTIES**

*Does the system maintain any information describing how any individual exercises their rights guaranteed by the First Amendment?*

No

*Is the system information used to conduct 'data-mining' as defined in the Implementing Recommendations of the 9/11 Commission Act of 2007, Public Law 110-53, Section 804?*

No

*Will this system have the capability to identify, locate, and monitor individuals or groups of people?*

No

*Does computer matching occur?*

No

## **ACCOUNTING OF DISCLOSURES**

*Does the system include or require disclosure of tax or employee information to anyone other than IRS employees in the performance of their duties, or to the person to whom the information pertains or to a 3rd party pursuant to a Power of Attorney, tax or Privacy Act consent?*

No