Date of Submission: November 29, 2015 PIA ID Number: **1496**

A. SYSTEM DESCRIPTION

1. Enter the full name and acronym for the system, project, application and/or database. W-2 Verification Code Pilot Validation Tool, W2VC

2. Is this a new system? Yes

3. Check the current ELC (Enterprise Life Cycle) Milestones (select all that apply)

Yes	Vision & Strategy/Milestone 0
Yes	Project Initiation/Milestone 1
Yes	Domain Architecture/Milestone 2
Yes	Preliminary Design/Milestone 3
Yes	Detailed Design/Milestone 4A
Yes	System Development/Milestone 4B
No	System Deployment/Milestone 5
No	Operations & Maintenance (i.e., system is currently operational)

4. Is this a Federal Information Security Management Act (FISMA) reportable system? No

A.1 General Business Purpose

5. 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.

The IRS has been working with the tax industry on a series of steps to combat tax-related identity theft and refund fraud. One area we're exploring with tax industry is a new way to verify wages and withholding claimed by taxpayers on e-filed 1040s. For filing season 2016, we are planning a pilot to test a capability to verify the authenticity of Form W-2 data. Here's how it will work. A small number of payroll firms and employers partnering with the IRS for this pilot will originate W-2s with a new Verification Code field on the Employee Copy of the W-2 issued by that payroll firm. We hope to soon announce the names of the payroll service providers that will be participating in this pilot for tax-year 2015. The W-2 Verification Code will have a length of 16 alphanumeric characters. It will be displayed on the W-2 in four groups of 4 alphanumeric characters, separated by hyphens. These Verification Codes will be calculated by W-2 Processors based on the output of a secure hash code algorithm, and in accordance with guidance we are providing through a Technical Specification Document. In other words, a calculation will be performed for each W-2 form based on select data elements on that form, and thus these values are not predetermined and distributed by the IRS ahead of time. When taxpayers with a Verification Code on their W-2s prepare to electronically file their 1040s, they would be requested to input the associated Verification Code value into their software product. The IRS will analyze this pilot data in a "test-and-learn" review using the W@VC tool to see if it is useful in evaluating the integrity of W-2 information submitted by taxpayers. The IRS will partner with four payroll firms to originate W-2s with a new Verification Code field on the employee copies of their Form W-2. This Verification Code value will be derived by the algorithm – essentially, a formula – using (1) a secret key securely provided to these W-2 Processors by the IRS, and (2) data from select W-2 boxes. When taxpayers with hash Verification Code W-2s file their 1040s, they would be requested by partnering e-File providers to input the Verification Code value now appearing on the employee copies of their W-2. The IRS will analyze

this pilot data using the Verification Tool in a “test-and-learn” review to see if it is useful in evaluating the integrity of W-2 information submitted by taxpayers.This tool will validate the W-2 Verification Code submitted on the relevant W-2s. The validation tool will take in two tab delimited input files: EIN/Key file and relevant W-2 data including W-2 Verification Code. The EIN/Key file will be supplied by W&I W-2 Verification Code Pilot Team and W-2 data will be extracted through BOE RRP. The tool will read in the two files and apply HMAC algorithm to regenerate the W-2 Verification Code. It will compare the submitted Verification Code with the tool generated Verification Code and provide results in an output file.

B. PII DETAIL

6. Does the system use, collect, receive, display, store, maintain, or disseminate IR Code 6103 taxpayer information; or any type of Sensitive but Unclassified (SBU) or Personally Identifiable Information (PII)? Yes

6a. If **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

If **yes**, check who the SSN (or tax identification number) is collected on.

Yes On Primary Yes On Spouse No On Dependent

If **yes**, check all types SSN s (or tax identification numbers) that apply to this system:

Yes	Social Security Number (SSN)
Yes	Employer Identification Number (EIN)
Yes	Individual Taxpayer Identification Number (ITIN)
No	Taxpayer Identification Number for Pending U.S. Adoptions (ATIN)
No	Practitioner Tax Identification Number (PTIN)

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

The Office of Management and Budget memorandum M-07-16 requires that federal agencies develop a mitigation or elimination strategy for systems that use SSNs. An exception to that requirement is when the SSN is uniquely needed to identify a user’s record. The W-2VC tool requires the use of SSN’s because no other identifier can be used to uniquely identify a taxpayer. SSNs are permissible from Internal Revenue Code (IRC) 6109, which requires individual taxpayers to include their SSNs on their income tax returns.

6b. Does this system contain other (non-SSN) PII that it uses, collects, receives, displays, stores, maintains, or disseminates? (i.e. Names, addresses, etc.) Yes

If **yes**, specify the information.

Selected	PII Element	On Primary	On Spouse	On Dependent
No	Name	No	No	No
No	Mailing address	No	No	No
No	Phone Numbers	No	No	No
No	Date of Birth	No	No	No
No	Place of Birth	No	No	No
No	SEID	No	No	No
No	Mother's Maiden Name	No	No	No

No	Protection Personal Identification Numbers (IP PIN)	No	No	No
No	Internet Protocol Address (IP Address)	No	No	No
No	Criminal History	No	No	No
No	Medical Information	No	No	No
No	Certificate or License Numbers	No	No	No
No	Vehicle Identifiers	No	No	No
No	Passport Number	No	No	No
No	Alien (A-) Number	No	No	No
No	Financial Account Numbers	No	No	No
No	Photographic Identifiers	No	No	No
No	Biometric Identifiers	No	No	No
Yes	Employment (HR) Information	No	No	No
No	Tax Account Information	No	No	No

6c. Does this system contain SBU information that is not PII, it uses, collects, receives, displays, stores, maintains, or disseminates? Yes

If **yes**, select the types of SBU

Selected	SBU Name	SBU Description
Yes	Agency Sensitive Information	Information which if improperly used or disclosed could adversely affect the ability of the agency to accomplish its mission
No	Procurement sensitive data	Contract proposals, bids, etc.
No	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.
No	Proprietary data	Business information that does not belong to the IRS
No	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
No	Physical Security Information	Security information containing details of serious weaknesses and vulnerabilities associated with specific systems and facilities
No	Criminal Investigation Information	Information concerning IRS criminal investigations or the agents conducting the investigations.

6d. Are there other types of SBU/PII used in the system? No

6e. Cite the authority for collecting SBU/PII (including SSN if relevant)

No	PII for federal tax administration is generally Internal Revenue Code Sections 6001, 6011, 6012e(a
Yes	SSN for tax returns and return information is Internal Revenue Code Section 6109
No	SSN for personnel administration (IRS Employees) is 5 USC & Executive Order 9397
No	PII for personnel administration is 5 USC
No	PII about individuals for Bank Secrecy Act compliance 31 USC
No	Information by CI for certain money laundering cases may be 18 USC

6f. Has the authority been verified with the system owner? Yes

B.1 BUSINESS NEEDS AND ACCURACY

7. 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.

The SSN on participating Forms W-2 is a data element used to generate the Verification Code. The SSN is required information for the tool to validate the Verification Code recieved by the IRS, and if continued toward production, will be used by IRS to verify the legitimacy of data said to come from a Form W-2.

8. How is the SBU/PII verified for accuracy, timeliness, and completeness? Explain how steps are taken to ensure that all information maintained by the system that is used by IRS to make any adverse determination about an individual's rights, benefits, and/or privileges is maintained with such accuracy, relevance, timeliness, and completeness as is reasonably necessary to assure fairness to the individual in the determination

The project team will use a validation tool developed by IT to recalculate the VC values for any Form W-2 impacted by the pilot. More specifically, the respective TP-submitted and recalculated values will be compared, and if the values are equivalent then the W-2 data will considered authentic. In cases where there is a data discrepancy, analysts will assess whether the value mismatch is likely related to error, or potentially refund fraud. To verify data authentication, hash functions are combined with Keyed-Hash Message Authentication Code (HMACs) or digital signatures. For HMACs, a secret key is shared between the third-party submitting the data and the verifying party, and this key becomes part of the input data used in the hash function. Thus HMACs verify the source of the submitted information because only intended users would possess the secret key. The software shall identify which symmetric secret key to use to generate a given hash-based message authentication code (HMAC). The software shall create an HMAC digest using assigned shared secret key and select W-2 data elements: • Social Security Number (SSN) • Employer Identification Number (EIN) • Federal Wages • Federal Withholding This is to recalculate hash digest with data from tax returns. Calculated digest is to be written to an output file along with other output.User interface shall be provided via a web interface which allows the user to upload input files 1 and 2, and then download an output file of the processed data. Once the user downloads the output file or a timer limit is reached the uploaded data and output file must be deleted from the system. As a precaution, a configurable timer should be used to ensure files are cleaned up on the system in the event a user processes data, but does not download it. Software shall compare calculated hash digest with taxpayer supplied digest and compute Exact Match Rate and Adjusted Match Rate, and assign Reason Code(s). Then write data to output file.

C. PRIVACY ACT AND SYSTEM OF RECORDS

9. Are 10 or more records containing SBU/PII maintained, stored, and/or transmitted by or through this system? Yes

9a. If **yes**, are records in the system retrieved by any personal identifier (e.g., name, SSN, Photograph, IP Address) for an individual? Yes

If **yes**, is there a System of Records Notice(s) or SORNs that addresses the PII records in this system? Yes

If **yes**, enter the SORN number(s) and the complete the name of the SORN.

SORNS Number	SORNS Name
Treas/IRS 24.030	IMF
Treas/IRS 24.046	BMF
Treas/IRS 34.037	Audit Trail and Security Records system

If **yes**, does the System of Records Notice(s) (SORN) published in the Federal Register adequately describe the records as required by the Privacy Act? Yes

D. RESPONSIBLE PARTIES

10. Identify the individuals for the following system roles. N/A

E. INCOMING PII INTERFACES

11. Does the system receive SBU/PII from other system or agencies? No

F. PII SENT TO EXTERNAL ORGANIZATIONS

12. Does this system disseminate SBU/PII? No

G. PRIVACY SENSITIVE TECHNOLOGY

13. Does this system use social media channels? No

14. Does this system use privacy-sensitive technologies such as mobile, cloud, global position system (GPS), biometrics, RFID, etc.? No

15. Does the system use cloud computing? No

16. Does this system/application interact with the public? No

H. INDIVIDUAL NOTICE AND CONSENT

17. Was/is notice provided to the individual prior to collection of information? Yes

17a. If **yes**, how is notice provided? Was the individual notified about the authority to collect the information, whether such is mandatory or voluntary, the purpose for which the information will be used, with whom the information may be shared, and the effects, if any, if they decide not to provide any of the requested information?
The participating payroll firms will originate substitute Form W-2s with a new Verification Code field on the employee copies B and C, either in a separate box (labeled "VC"), or in the box used for the employee's name and address. Instructions for the VC will be included or appear on the back of the substitute forms. When impacted taxpayers prepare to e-File their 1040s, they will be requested by participating e-File providers to input the associated VC value into their software product along with their other W-2 data.

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

18a. If **yes**, describe the mechanism by which individuals indicate their consent choice(s):
The Verification Code field is voluntary for the Pilot. Providing the code when submitting a return indicates consent.

19. How does the system or business process ensure due process regarding information access, correction and redress?
Participation in the pilot will be voluntary for taxpayers; the IRS will not reject returns based on the Verification Code value, or lack of value, a taxpayer provides on their returns. There is no compliance check associated with the Verification Code Pilot.

I. INFORMATION PROTECTION

20. 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

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

IRS Employees? Yes

IRS Employees?	**Yes/No**	**Access Level(Read Only/Read Write/ Administrator)**
Users	No	
Managers	No	
Sys. Administrators	Yes	Read and Write
Developers	Yes	Read-Only

Contractor Employees? No

Contractor Employees?	**Yes/No**	**Access Level**	**Background Invest.**
Contractor Users			
Contractor Managers			
Contractor Sys. Admin.			
Contractor Developers			

21a. How is access to SBU/PII determined and by whom? The EIN/Key file will be supplied by W&I W-2 Verification Code Pilot Team and W-2 data will be extracted through BOE RRP. Access to RRP BOE is provided through an aproved 5081. The tool is accessed using a unique username and password via a web interface which allows the user to upload input files and then download an output file of the processed data. The web interface will account for authentication and authorization of end user.

21b. If computer matching occurs, can the business owner certify that it meets requirements of IRM 11.3.39 Disclosure of Official Information, Computer Matching & Privacy Protection Act ? Not Applicable

I.1 RECORDS RETENTION SCHEDULE

22. Are these records covered under the General Records Schedule (GRS), or have a National Archives and Records Administration (NARA) archivist approved a Record Control Schedule (RCS) for the retention and destruction of official agency records stored in this system? Yes

22a. If **yes**, how long are the records required to be held under the corresponding RCS and how are they disposed of? In your response, please include the complete IRM number 1.15.XX and specific item number and title.

Once the user downloads the output file or a timer limit is reached the uploaded data and output file must be deleted from the system. As a precaution, a configurable timer should be used to ensure files are cleaned up on the system in the event a user processes data, but does not download it. The method used for sanitization will follow NIST SP 800-88 guidelines. Any new records generated by the system will be managed according to requirements under IRM 1.15.1 and 1.15.6, and will be destroyed using IRS Records Control Schedules (RCS) 29, Item 85 (1b), (2b), and as coordinated with the IRS Records and Information Management (RIM) Program and IRS Records Officer.

I.2 SA&A OR ECM-R

23. Has the system been through SA&A (Security Assessment and Authorization) or ECM-R (Enterprise Continuous Monitoring Reauthorization)? No

23c. If **no**, is the system secured in accordance with all applicable federal, treasury, and IRS security policy, procedures, and requirements? Yes

23.1 Describe in detail the system s audit trail. RRP BOE has the Negative TIN check in place which provides the file that is upload into the tool, the tool itself does not contain or retain TIN data.

J. PRIVACY TESTING

24. Does the system require a System Test Plan? Yes

24b. If **yes**, Is the test plan in process or completed: In Process

If **in process**, when is the test plan scheduled for completion? 1/15/2016

24.3 If **completed/ or in process,** 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?
The AD EST timeline includes testing the W-2 Post Filing Validation tool beginning 9/25/15 to be completed before the production release 1/15/16.

K. SBU Data Use

25. Does this system use, or plan to use SBU Data in Testing? No

L. NUMBER AND CATEGORY OF PII RECORDS

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

26a. IRS Employees: Not Applicable
26b. Contractors: Not Applicable
26c. Members of the Public: More than 1,000,000
26d. Other: No

M. CIVIL LIBERTIES

27. Does the system maintain any information describing how any individual exercises their rights guaranteed by the First Amendment? No

28. Is the system information used to conduct data-mining as defined in the *Implementing the 9/11 Commission Recommendations Act of 2007, Public Law 110-53, Section 804?* No

29. Will this system have the capability to identify, locate, and monitor individuals or groups of people? No

N. ACCOUNTING OF DISCLOSURES

30. 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

End of Report