PRIVACY IMPACT ASSESSMENT

Organization: Office of Economic Opportunity, State of Arizona

Information System: Integrated Data System (IDS)

System Owner: Vignesh Sukumaran

Date: February 24, 2022

1. INTRODUCTION

1.1. Describe the system including the name, acronym, and a brief description of the program or purpose for the system.

OEO operates an integrated data system (IDS) that integrates administrative data from various state agencies and educational institutions for evidence-based policymaking. OEO does not collect new data directly from the public for the IDS. It uses existing data that is collected by state agencies and educational institutions as part of their routine administrative operations.

The IDS is used exclusively for "statistical purposes." A "statistical purpose" is defined as "the description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups; and includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support such purposes." Project proposals requesting access to the IDS are approved by data-contributing agencies and institutions.

In the IDS, the personal identifiers in the data are only used for record linkage, which is the process of identifying and matching records that correspond to the same individual from disparate administrative data sources. Before the system of linked records created from record linkage operations is used in approved projects, it is de-identified by removing personal identifiers such as names, social security numbers and addresses.

1.2. Describe the purpose for which the personally identifiable information (PII) is collected, used, maintained or shared.

The IDS is used exclusively for statistical purposes—that is, the analysis of data to generate inferences about groups. Project proposals requesting access to the IDS are approved by data-contributing agencies and institutions.

In the IDS, the direct identifiers in the data are only used for record linkage, which is the process of identifying and matching records that correspond to the same individual from disparate administrative data sources. Before the system of linked records created from record linkage operations is used in approved projects, it is de-identified by removing direct identifiers such as names, social security numbers and addresses.

1.3. Is this a new system, or one that is currently in operation?

This is a new system.

1.4. Is this Privacy Impact Assessment (PIA) new, or is it updating a previous version?

This is a new PIA for a new information system.

1.5. Is the system operated by the agency or by a contractor?

The system is operated by the agency.

1.5.1. If the system is operated by a contractor, does the contract or other acquisition related documents include privacy requirements?

N/A

2. LEGAL AUTHORITIES AND OTHER REQUIREMENTS

2.1. What specific legal authorities and/or agreements permit and regulate the collection and use of data by the system? Please include the name and citation of the authority.

The following data are collected and used in the IDS: (1) Education records from the Maricopa County Community College District; (2) Education records from the Pima County Community College District; (3) Participant records from the Adult Education program in the Arizona Department of Education (ADE); (4) Participant records from the Workforce Innovation & Opportunity Act (WIOA) programs in the Arizona Department of Economic Security (DES); and (5) Wage records from the Unemployment Insurance (UI) program in DES.

Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 C.F.R. Part 99) is a Federal law that protects the privacy of student **education records** and affords parents and eligible students certain rights with respect to these education records. The general rule is that a parent or eligible student must provide a signed and dated written consent before an educational agency or institution discloses PII from the student's education records.

There are exceptions to FERPA's general requirement of consent that allow educational agencies and institutions to disclose PII from education records without consent. Educational agencies and institutions that are participating in the IDS use FERPA's *audit and evaluation exception* to share PII from their education records with OEO for the explicit purpose of operating and maintaining the IDS; and use either FERPA's *audit and evaluation exception* or FERPA's *studies exception* to permit redisclosure of PII from their education records in the IDS to their authorized representatives for carrying out projects in the IDS.

Both the audit and evaluation exception and the studies exception require written agreements with mandatory elements in the agreements. Participating educational agencies and institutions have entered into written agreements with OEO for operating the IDS. They will also enter into written agreements with their authorized representatives using their data in the IDS on a project-by-project basis.

Confidential **wage records** in the Unemployment Insurance (UI) program are protected by the confidentiality and disclosure requirements in Federal Unemployment Compensation (UC) law and its implementing regulations set forth at 20 C.F.R. Part 603, and state law. DES, the State UC agency, shares wage records that include PII with OEO for the operation and maintenance of the IDS. Pursuant to 20 C.F.R. Part 603, DES permits OEO to redisclose wage records only to public officials, a term defined in 20 C.F.R. § 603.2, for carrying out projects in the IDS. Additionally, A.R.S. §

23-722.04 allows DES and OEO to use wage records for operating the IDS, and to disclose wage records to ADE, Arizona Board of Regents, Arizona public universities, Arizona community college districts and other program offices in DES when they use the IDS to carry out their projects.

Federal UC regulations require DES to enter into written agreements with mandatory elements in the agreements when disclosing wage records to other public officials. DES has entered into written agreements with OEO, a public official, for operating the IDS using wage records. DES will also enter into written agreements with other public officials before they access wage records in the IDS on a project-by-project basis.

DES is the state agency charged with the administration of the following WIOA programs that share their **participant records** with OEO for use in the IDS: adult, youth and dislocated worker programs, Jobs for Veterans State Grants program, Trade Adjustment Assistance program, Wagner-Peyser Act Employment Services program, and Vocational Rehabilitation (VR) Services program. A.R.S. § 41-1959 protects the confidentiality of participant records in the WIOA programs administered by DES; and it also allows for the use of those participant records by OEO for operating the IDS and disclosure to ADE, Arizona Board of Regents, Arizona public universities and Arizona community college districts when they use the IDS to carry out their projects. PII in the VR program is further protected by federal regulations set forth at 34 C.F.R. § 361.38.

Pursuant to A.R.S. § 41-1959, DES has entered into data sharing agreements with OEO for operating the IDS using WIOA participant records. DES will also enter into written agreements with the aforementioned entities, on a project-by-project basis, before they access DES' WIOA participant records in the IDS.

2.2. Is the information in this system retrieved by an individual's name or personal identifier such as a Social Security Number or other identification?

In the IDS, personal identifiers such as names and Social Security numbers are only used for record linkage, which is the process of identifying and matching records that correspond to the same individual from disparate administrative data sources. Before the system of linked records that was created from automated record linkage operations is used for statistical purposes in approved projects, it is de-identified by removing personal identifiers. This system of records may not be accessed for purposes of determining if the system contains a record pertaining to a particular individual, or for contesting the contents of a record.

2.2.1. If the above answer is YES, this system will need to be covered by the Privacy Act System of Records Notice(s) (SORN(s)). Please provide the SORN name, number, federal register citation and link, or indicate that a SORN is in progress.

N/A

2.2.2. If the above answer is NO, explain why SORN was not necessary. For example, the information is not retrieved by an identifier, the information is not

maintained in a system of records, or the information is not maintained by the Agency, etc.

A SORN is a Privacy Act requirement for federal agencies not for Arizona government agencies.

2.3. Is there a records retention schedule that has been approved by National Archives and Records Administration (NARA) for the records contained in this system? Please provide all relevant NARA schedule numbers and disposition instructions.

The records in the IDS are not covered by NARA. Records retention schedule is determined by data sharing agreements and data governance rules. Retention schedule for the IDS is yet to be established by the Arizona State Library, Archives and Public Records under Schedule Number CS-1184. Records retention is as follows:

(1) Education records from the Maricopa County Community College District - 10 years.

(2) Education records from the Pima County Community College District - 10 years.

(3) Participant records from the Adult Education program in the Arizona Department of Education - 20 years.

(4) Participant records from the Workforce Innovation & Opportunity Act (WIOA) programs in the Arizona Department of Economic Security - 20 years.

(5) Wage records from the Unemployment Insurance program in the Arizona Department of Economic Security - 20 years.

2.4. Is the PII contained in this system disposed of appropriately, and in accordance with the timelines in the records disposition schedule?

The IDS is a new system that has not reached the end of a retention period for records yet.

3. CHARACTERIZATION AND USE OF INFORMATION

3.1. List the specific PII elements (e.g., name, email, address, phone number, date of birth, Social Security, etc.) that the system collects, uses, disseminates, or maintains.

The system collects and uses the following identifiers:

(1) Name

(2) Social Security number

(3) Date of birth

(4) Address

(5) Telephone number

(6) Gender

(7) Ethnicity

(8) Race

(9) Adult education student identification number

(10) Community college student identification number

(11) Vocational Rehabilitation client number

(12) WIOA program participant identification number

(13) Unemployment Insurance employer number of wage earner's employer

3.2. Does the system collect only the minimum amount required to achieve the purpose stated in Question 1.2?

Yes, see response in 1.2 for further details.

3.3. What are the sources of PII collected (e.g., individual, school, another agency, commercial sources, etc.)?

(1) Education records from the Maricopa County Community College District.

(2) Education records from the Pima County Community College District.

(3) Participant records from the Adult Education program in the Arizona Department of Education.

(4) Participant records from the Workforce Innovation & Opportunity Act (WIOA) programs in the Arizona Department of Economic Security.

(5) Wage records from the Unemployment Insurance program in the Arizona Department of Economic Security.

3.4. How is the PII collected from the stated sources listed in Question 3.3 (e.g., paper form, web page, database, etc.)?

The IDS collects PII in electronic files and in a database.

3.5. How is the PII validated or confirmed to ensure the integrity of the information collected? Is there a frequency at which there are continuous checks to ensure the PII remains valid and accurate?

The IDS uses existing data that is collected by state agencies and educational institutions as part of their routine administrative operations. Although the system measures data quality and in certain cases reports back to the originating agency, it does not attempt to improve the integrity of the information or the PII.

3.6. Describe how the PII is used to achieve the purpose stated in Question 1.2 above.

In the IDS, the direct identifiers in the data are only used for record linkage, which is the process of identifying and matching records that correspond to the same individual from disparate administrative data sources. The system of linked records created from record linkage operations is used for statistical purposes such as program evaluation, policy research and cost-benefit analyses.

3.7. Is the system using PII for testing/researching new applications or information systems prior to deployment for training employees?

No.

3.7.1. If the above answer is YES, what controls are in place to minimize the risk and protect the data?

N/A

3.8. Does the system collect Social Security Numbers? Note that if the system maintains Social Security Numbers but does not explicitly collect them, answer 3.8.1 to address the purpose for maintaining them.

Yes.

3.8.1. If the above answer is YES, explain the purpose for its collection, and how the Social Security Number will be used.

The IDS uses existing data that is collected by state agencies and educational institutions as part of their routine administrative operations. Social Security numbers in the data received from participating agencies and institutions are only used for record linkage, which is the process of identifying and matching records that correspond to the same individual from disparate administrative data sources.

3.8.2. Specify any alternatives considered in the collection of Social Security Numbers and why the alternatives were not selected.

As the Social Security number is a numeric identifier that is unique to a person, it is the most important matching field in record linkage operations. However, Social Security numbers are not always available or accurate, so other personal identifiers like names and dates of birth are used in addition to Social Security numbers for record linkage.

4. NOTICE

4.1. How does the system provide individuals with notice about the collection of PII prior to its collection (e.g., direct notice, such as a Privacy Act Statement (if applicable) or public notice, such as a SORN, PIA,)? If notice is not provided, explain why not.

The IDS does not collect PII directly from individuals, so it does not give prior notice to individuals in the data. This PIA is published on the website https://ids.az.gov.

4.2. Provide the text of the notice or the link to the webpage where the notice is posted if notice is provided other than by SORN or PIA.

This PIA is published on the website https://ids.az.gov. OEO will provide no other notice of PII collection and use.

4.3. What opportunities are available for individuals to consent to uses (including new uses of previously collected PII), decline to provide PII, or opt out of the project?

Educational authorities provide education records for use in the IDS under exceptions to individual consent granted to them by FERPA law. There is no federal or state requirement that the state UC Agency, state VR agency or the state workforce agency, participating in the IDS, obtain informed written consent from individuals prior to the use of their personal information for IDS purposes.

4.4. Is the notice referenced in Question 4.1 reviewed and revised when there are changes in the practice, policy, or activities that affect the PII and privacy to ensure that individuals are aware of and can consent to, where feasible, these changes?

This PIA will be updated whenever new data from state agencies and educational institutions is added to the IDS.

5. INFORMATION SHARING AND DISCLOSURES

5.1. Will PII be shared internally with other OEO departments? If the answer is NO, please skip to Question 5.4.

No. Only OEO staff dedicated to IDS operations will have access to PII.

5.2. What PII will be shared and with whom?

N/A

5.3. What is the purpose for sharing the specified PII with the specified internal organizations?

N/A

5.4. Will the PII contained in the system be shared with external entities (e.g. another agency, school district, the public, etc.)? If the answer is NO, please skip to Question 6.1.

Yes. A Project Review Committee, composed of subject matter experts from each of the government programs and education institutions contributing data to the IDS, oversees data access and use.

5.5. What PII will be shared and with whom? List programmatic disclosures only.

We define personally identifiable information (PII) as information that can be used to distinguish or trace an individual's identity. It includes direct and indirect identifiers. A direct identifier such as a name or a Social Security number directly identifies a specific individual. On the other hand, an indirect identifier such as a date of birth does not identify any specific individual by itself but can be aggregated and "linked" with other information to identify an individual.

Only OEO staff dedicated to IDS operations will have access to direct identifiers. Direct identifiers are used in the IDS exclusively for record linkage and will not be shared or disclosed further to another party.

The system of linked records, created by record linkage, is used for statistical purposes in authorized projects. It does not include direct identifiers. State laws and federal laws and regulations determine who gets access to the system of linked records. Typically, projects are carried out by agencies and educational institutions that are participating by contributing data to the IDS.

Certain indirect identifiers, such as gender, ethnicity, race, date of birth and zip code are important for statistical analysis and removing them may damage the utility of a dataset. When compiling data from the system of linked records for use in approved projects, indirect identifiers are de-identified where possible to minimize the risk of re-identification. Techniques for de-identification of indirect identifiers include suppression, such as the removal of the date and month components in date of birth, and generalization, such as the replacement of the zip code 12345 with an interval 12000 to 12999.

5.6. What is the purpose for sharing the PII with the specified external entities?

The system of linked records in the IDS is used for authorized projects that have a statistical purpose. The research and analysis work done in these projects generate or study aggregate indicators, performance metrics, and population descriptions, trends and correlations, to name a few.

5.7. Is the sharing with the external entities authorized?

Yes. A project review committee, composed of subject matter experts from each of the government programs and education institutions contributing data to the IDS, oversees data access and use.

The following procedure governs the access and use of data:

(a) Individuals and organizations who wish to use the IDS will first contact OEO with a project proposal.

(b) OEO staff will make an initial review of the proposal, identifying any obvious data availability or legal challenges. If it is clear that requested data may not or cannot be provided in accordance with the proposal as submitted, OEO staff will contact the applicant with suggestions for revision (if possible) or a refusal to fulfill the request with an explanation regarding why the request cannot be fulfilled.

(c) When OEO staff has a project proposal that is clear and seemingly permissible, OEO staff will forward the proposal to members of the Project Review Committee whose data are requested in the proposal.

(d) The Project Review Committee members, upon review of the proposal, may request the applicant provide needed clarification through OEO staff. Once a proposal is clear, the Project Review Committee members will shepherd the request through the relevant program offices at their respective agencies.

(e) Agency program offices will process the request and let OEO staff know if they approve or reject the data request. If the program office/offices reject the request, OEO staff will notify the applicant. The applicant then may revise and resubmit the proposal to OEO staff, restarting the process; choose not to revise the proposal but appeal to the full Project Review Committee; or the applicant may opt to not proceed with the project. Even if an applicant appeals or resubmits a revised proposal, it must be approved by all relevant program offices before a data request can be fulfilled.

(f) If all the relevant program offices approve a request, OEO will notify the applicant and prepare data sharing agreements for the applicant to sign and submit to OEO.

(g) Once the agreements have been signed, OEO staff will create a secure research platform and provide access to the requested data to the individuals identified in the proposal for the period authorized by the agreements.

(h) Once data analysis is completed, the project team must submit any material intended for publication or release or sharing to the Project Review Committee for statistical disclosure review. The Project Review Committee will review research and results, working with the project team as necessary, before authorizing dissemination of results.

5.8. Is the system able to provide and retain an account of any disclosures made and make it available upon request?

Yes. To fulfill a request from individuals requesting details about the disclosure of their records to third parties conducting statistical analysis, OEO will generate the following information for each instance of disclosure: the purpose of disclosure; the date of disclosure; the identity of the party to which data was disclosed; and a description of the data disclosed.

5.9. How is the PII shared with the external entity (e.g. email, computer match, encrypted line, etc.)?

OEO restricts access to the de-identified system of linked records by allowing projects to only be carried out in a secure research workspace in the IDS. The secure research workspace is a protected enclave containing analytical tools, data storage and general computing resources. It prevents individual-level data from leaving the IDS.

5.10. Is the sharing pursuant to a Computer Matching Agreement (CMA), Memorandum of Understanding (MOU), or other type of approved sharing agreement with another agency?

Yes. State agencies and educational institutions participating in the IDS have signed data sharing agreements with OEO that conform to federal and state regulations governing such agreements. Individuals/organizations who wish to carry out projects in the IDS also enter into agreements with the state agencies and educational institutions whose data are requested in their project proposal.

5.11. Does the project place limitations on redisclosure?

Administrative safeguards that help prevent improper disclosure include background checks on all system operators and data users, and data sharing contracts that prohibit data users from reidentifying data subjects and redisclosing confidential information.

6. REDRESS

6.1. What are the procedures that allow individuals to access their own information?

OEO is the point of contact for individuals seeking to know the content of their records in the IDS. State agencies and educational institutions contributing data to the IDS have agreed to develop a procedure by May 31, 2022 for fulfilling such requests. The procedure will address the following:

(a) How to handle requests from individuals made in person, by mail, by email and by phone.

(b) How to identify and authenticate the person requesting information with the required degree of confidence that he/she is who that person claims to be.

(c) When to redirect inquiries to the individual state agencies and educational institutions.

(d) How promptly the request should be fulfilled.

(e) How state agencies and educational institutions will work with OEO to fulfill the request.

(f) What to include in the data disclaimer that advises the applicant that data undergoes corrections and revisions over time.

6.2. What procedures are in place to allow the subject individual to correct inaccurate or erroneous information?

For IDS operations, OEO uses existing data that is collected by state agencies and educational institutions as part of their routine administrative operations. Individuals who seek to correct inaccurate or erroneous information in IDS records will be directed to the pertinent program office in the state agency or educational institution.

6.3. How does the project notify individuals about the procedures for correcting their information?

When fulfilling a “right to inspect” request from an individual, OEO will also provide contact information for the state agency or educational institution maintaining the original system of records.

7. SAFEGUARDS

7.1. Does the system owner work with their CSO/ISSO to build privacy & security into the system and build privacy extensions to the extent feasible?

The system owner works with the State Information Security and Privacy Office to implement the security controls prescribed by National Institute of Science and Technology (NIST) and enhanced by the Arizona Risk and Authorization Management Program (AZRamp). The selected security controls are commensurate with the protection needed for confidential information stored and processed by the IDS.

The system owner worked with the U.S. Department of Education’s Privacy Technical Assistance Center (PTAC) on FERPA compliance issues and best practices for protecting student privacy.

7.2. Is an Authority to Operate (ATO) required?

OEO uses an ATO issued by its CIO/CISO for authorization of IDS operations and reauthorization after significant changes to the IDS. The ATO is based on the CIO/CISO review of a recent System Security Plan, penetration test and security assessment, and any relevant Plan of Action and Milestones.

7.3. Under NIST FIPS Pub. 199, what is the security categorization of the system: Low, Moderate, or High?

Moderate.

7.4. What administrative, technical, and physical safeguards are in place to protect the information?

Section 5.7 describes the authorization procedure that governs data access and use.

Sections 5.1 and 5.9 describe technical safeguards such as de-identification of direct and indirect identifiers, and restricted access using protected enclaves.

Administrative safeguards that help prevent improper disclosure include background checks on all system operators and data users, and data sharing contracts that prohibit data users from reidentifying data subjects and redisclosing confidential information.

7.5. Is the information in the system appropriately secured in accordance with the IT security requirements and procedures as required by Federal law and policy?

The information in the IDS is secured in accordance with statewide security policies and standards set by the Arizona Department of Administration (ADOA). A.R.S. § 18-104 requires ADOA to adopt statewide security standards for information technology.

7.6. Has a risk assessment been conducted where appropriate security controls to protect against that risk have been identified and implemented?

A risk assessment is underway.

7.7. Please describe any monitoring, testing, or evaluation conducted on a regular basis to ensure the security controls continue to work properly at safeguarding the PII.

OEO has a comprehensive continuous monitoring strategy for information security that covers technology, processes, procedures, operating environments and people. The continuous monitoring program encompasses security control assessments, security status monitoring and reporting for risk-based decision making. The continuous monitoring program collects data and reports findings that are addressed in a timely manner to safeguard the system and its PII.

8. AUDITING AND ACCOUNTABILITY

8.1. How does the system owner assess and ensure that the PII is used in accordance with stated practices in this PIA?

The system owner audits access to information assets that include PII and limits access to the audit logs to only a few administrators.

8.2. Does the system owner continuously monitor and audit the privacy controls to ensure effective implementation?

The system owner regularly monitors PII processing and transparency controls. The system owner does the following, among other things:

(1) Reviews this privacy impact assessment and updates it when there are changes in the data and in the laws and regulations governing data use.

(2) Identifies the minimum PII elements that are relevant and necessary to accomplish the IDS purposes and regularly reviews the PII holdings.

(3) Reviews and tests the security and privacy incident response plan annually.

(4) Oversees information sharing with third parties carrying out authorized projects in the IDS by working closely with the Project Review Committee on data access and disclosure on an ongoing basis.

8.3. What are the privacy risks associated with this system and how are those risks mitigated?

Personal identifiers such as Social Security numbers and names are not necessary for carrying out statistical activities, so the data used for projects are de-identified by removing these identifiers. As records without the personal identifiers still carry the risk of reidentification by linkage with similar records in some external dataset that contains both the linking information and the identity of the data subject, individual-level de-identified data is never released. Projects are carried out by analysts using a secure research workspace in the IDS — a protected enclave that prevents analysts from moving data in or out of the IDS.

Publication and dissemination of results from projects carried out in the IDS are usually in the form of statistical tables with aggregate information. Statistical tables also carry the risk of disclosing information about identifiable persons when aggregate data does not have a sufficient number of data subjects to disguise the attributes of a single data subject. The required protection is achieved by the application of statistical disclosure limitation procedures whose purpose is to ensure that the risk of disclosing confidential information about identifiable persons will be very small.

VS