

**Privacy Impact Assessment (PIA)**

for the

**IES Data Center (IESDC)**

**Jan. 9, 19**

**For PIA Certification Updates Only:** This PIA was reviewed on June 9, 2019 by Ed Vaden certifying the information contained here is valid and up to date.

**Contact Point**

**Contact Person/Title: Ed Vaden, IES ISSO**

**Contact Email: ed.vaden@ed.gov**

**System Owner**

**Name/Title: Brian Taylor, IES Director of Technology**

**Principal Office: Institute of Education Sciences (IES)**

**Please submit completed Privacy Impact Assessments to the Privacy Office at** **privacysafeguards@ed.gov**

*Please complete this* ***Privacy Impact Assessment (PIA)*** *on how personally identifiable information (PII) is collected, stored, protected, shared, and managed electronically by your system. You may wish to consult with your ISSO in completing this document.*

***If a question does not apply to your system, please answer with N/A.***

# Introduction

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

The IES Data Center (IESDC) hosts IES survey collection and data dissemination websites. Applications hosted in this data center support the four Centers in IES:

* National Center for Education Evaluation and Regional Assistance – NCEE
* National Center for Education Research – NCER
* National Center for Education Statistics – NCES
* National Center for Special Education Research –NCSER

The IESDC also hosts survey collections and data dissemination websites for the Office of Postsecondary Education (OPE). These OPE websites are related to Integrated Postsecondary Education Data System (IPEDS) websites. IPEDS is part of NCES.

The data hosted in IESDC are covered by existing PIAs and OMB Clearance Packages.

* 1. Describe the purpose for which the personally identifiable information (PII)[[1]](#footnote-1) is collected, used, maintained or shared.

The Institute of Education Science is expected to provide national leadership in expanding fundamental knowledge and understanding of education from early childhood through postsecondary study, in order to provide parents, educators, students, researchers, policymakers, and the general public with reliable information about— the condition and progress of education in the United States, including early childhood education and special education; educational practices that support learning and improve academic achievement and access to educational opportunities for all students; and the effectiveness of Federal and other education programs.

Some data collections managed through IESDC include collection of PII for purposes of following up with respondents during later stages of multi-timepoint data collections and to support matching to administrative record data from other sources. This latter aspect of the work helps reduce respondent burden during the survey and typically provides more accurate information that a respondent might be able to provide during a relatively short survey.

* 1. Is this a new system, or one that is currently in operation?
	Currently Operating System
	2. Is this PIA new, or is it updating a previous version?
	Updated PIA
	3. Is the system operated by the agency or by a contractor?

Contractor

* + 1. If the system is operated by a contractor, does the contract or other acquisition-related documents include privacy requirements?

Yes

# Legal Authorities and Other Requirements*If you are unsure of your legal authority, please contact your program attorney.*

* 1. Whatspecific legal authorities and/or agreements permit and regulate the collection and use of data by the system? Please include name and citation of the authority.

IES is authorized in law under Title 20 – EDUCATION, CHAPTER 76 - Education Research, Statistics, Evaluation, Information, and Dissemination, Subchapter I – Education Sciences Reform Act of 2002. Data collected through IESDC and used by OPE are authorized under this same legislation.

**SORN**

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

	Yes

		1. If the above answer is **YES,** this system will need to be covered by Privacy Act System of Records Notice(s) (SORN(s)).[[2]](#footnote-2) Please provide the SORN name, number, Federal Register citation and link, or indicate that a SORN is in progress.



IES SORNs cover records maintained in the IES Data Center. For a complete list of applicable SORNs please see the following webpage: https://www2.ed.gov/notices/ed-pia.html#ies

* + 1. If the above answer is **NO**, explain why a 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 Department, etc.
		

Click here to enter text.

**Records Management**

**If you do not know your records schedule, please consult with your records liaison or send an email to** **RMHelp@ed.gov**

* 1. What is the records retention schedule 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.

NARA records retention schedules associated with IES are currently being updated. Data produced by IES for statistical analyses and reporting are generally considered permanent records. Information needed to collect or process these data that do not become part of the official data are temporary and may be retained for a period of 3-5 years or until the completion of the study.

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

	Yes

# Characterization and Use of Information

**Collection**

* 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 various data collections include information on educational outcomes, student assessments, education program offerings and program participation, staffing, qualifications of staff, individual characteristics (e.g., age, race/ethnicity, gender), salaries, and SSN’s in longitudinal studies for tracking and record linkages to administrative data.

Most studies at some point have contact information data for respondents. Generally, the contact information is not retained past the point when the data are processed for final database development. However, some studies include follow-up work with respondents and these studies retain contact information for extended periods. Some collections also have to include the official contact information for administrators (e.g., public school district superintendents) to be made public and these data are also retained past the point where a final database has been produced.

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

Yes

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

Most often sources of information are collected directly from individual students, parents, teachers, school administrators, childcare providers, or educational entities but may be collected by other entities specified in the corresponding SORN(s).

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

The mode of data collection depends on the specific study and the ability of respondents to respond using different communication tools. The more common modes of data collection include paper forms, web-based data collection tools, telephone interviews, and in-person interviews.

* 1. How is the PII validated or confirmed to ensure the integrity of the information collected?[[3]](#footnote-3) Is there a frequency at which there are continuous checks to ensure the PII remains valid and accurate?

The records which are collected and maintained in the Data Center are not for the purpose of providing a benefit to the individual whose records are collected or for enforcement purposes. Therefore, individual-level misidentification in records maintained in the IESDC is not a significant concern in terms of the PIA. However, the collection of these records is in service of providing official statistics, data and other information to support the work and research of policy makers, education practitioners, researchers, the general public and a wide range of additional stakeholders. To this end, IES applies current state of the art techniques to ensure the quality of the data.

**Use**

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

IESDC hosts publicly available data on publicly accessible websites. It also hosts IES survey collection websites. Restricted access is provided by a data licensing program, not through the IES Data Center. Restricted access data are shared for the educational research and analysis that is consistent with the intended use of the specific data file. The IES data licensing program is extended only to qualified researchers at institutions in the United States. The license is between IES and the qualified researcher and their institution, IES does not license individuals not affiliated with an established research firm or educational institution.

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

	No

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

Click here to enter text.

**Social Security Numbers**

*It is the Department’s Policy that, in order to collect Social Security Numbers, the System Owner must state the collection is: 1) authorized by law, 2) necessary for an agency purpose, and 3) there is no reasonable alternative.*

* 1. 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

* + 1. If the above answer is **YES**, explain the purpose for its collection, and how the SSN will be used.
		

Not all collections include the collection of Social Security numbers (SSNs). In the event that they are, SSN's are collected with OMB concurrence for the express purpose of record linkage with Department records systems and in some cases, for sample surveys, state or local student records. In the case of longitudinal studies, the SSN's may also be used for tracking study respondents as they change schools and/or geographic locations over the course of the study (e.g., to be able to follow the study respondents as they complete school and transition to the work force or postsecondary education or to follow study respondents as they transition from the completion of baccalaureate program into the work force or additional higher education).

* + 1. Specify any alternatives considered in the collection of SNNs and why the alternatives were not selected.
		

IES keeps abreast of studies that focus on the use of names for such linkages and follow-up studies, but the return rate on matches based on other forms of identification remain substantially lower than that obtained through the use of SSN's.

# Notice

* 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.

Every IES data collection has a System of Records Notice. In the case of surveys conducted in educational institutions advance letters are sent to the administrator for the institution describing the study, explaining the voluntary nature of the study and describing the pledge of data confidentiality. In addition, every IES data collection that includes individually identifiable information includes a description of the voluntary nature of the data collection and a pledge of confidentiality, per OMB standards and NCES Statistical Standard 4.2. The text of the pledge includes the following: Your answers may be used only for statistical or research purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law.

Furthermore, the routine statistical purposes for which the data may be used must be explained. If an individual educational institution requires informed consent from parents or adult students, that is included in the data collection procedures, otherwise each respondent is informed of the voluntary nature of their participation as it applies to both the entire data collection and to individual questions within the data collection. In the event of legally mandated participation in a data collection, the data provider (usually a representative of an institution) is provided a description of the legal requirement and all data protections that are afforded are provided, if the institutional data include data that are potentially disclosive of individual characteristics IES uses professional best practices to protect the identity of individuals in an institution when publishing data from the collection.

* 1. 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.
	

The notice is tailored to specific respondent types and will vary by study. For an example of the notices, please see the following teacher questionnaire for the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 at https://nces.ed.gov/ecls/pdf/secondgrade/Spring\_2013\_Teacher\_Ques\_Teacher\_Level.pdf.

* 1. 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?

Every IES data collection that includes individually identifiable information includes a description of the voluntary nature of the data collection and a pledge of confidentiality per OMB standards and NCES Statistical Standard 4.2. Each respondent is informed of the voluntary nature of their participation as it applies to both the entire data collection and to individual questions within the data collection. Statements like the following are shared with respondents before they provide information, "If you have comments or concerns regarding the status of your individual response to this survey, write directly to: … Participation is voluntary. You may skip questions you do not wish to answer;... All responses that relate to or describe identifiable characteristics of individuals may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law. Data will be combined to produce statistical reports. No individual data that links your name, address, telephone number, or identification number with your responses will be included..." For a complete example of how this is conveyed to respondents, please see language on the first page of the example cited in response to question 4.2 above.

In the event of legally mandated participation, the data providers (usually representatives of institutions) are provided a description of the legal requirement and any data protections afforded or provided. At the present time all collections with mandated participation collect institution-level rather than individual-level data.

* 1. 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?

Yes

# Information Sharing and Disclosures

**Internal**

* 1. Will PII be shared internally with other ED principal offices? If the answer is **NO**, please skip to Question 5.4.

	No
	2. What PII will be shared and with whom?
	

Click here to enter text.

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

Click here to enter text.

**External**

* 1. 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
	2. What PII will be shared and with whom? List programmatic disclosures only.[[4]](#footnote-4)
	**Note: If you are sharing Social Security Numbers externally, please specify to whom and for what purpose**.
	

Public access data include disclosure avoidance protections and may be shared with the general public. Restricted access data are shared with staff in ED and other federal agencies through a Memorandum of Understanding (MOU) that mirrors the requirements described above for the data licensing program (see discussion of IES restricted data licenses in 3.6). This program is not part of the IES Data Center. Note that Social Security Numbers are never shared. A recent example of such an MOU is one with the U.S. Census Bureau to link an IES study of college graduates to employment data to study relationships between college degree attainment and career trajectories.

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

The information is shared for statistical or research purposes in alignment with question 1.2.

* 1. Is the sharing with the external entities authorized?
	

Yes

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

Yes

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

All public access data have been approved for release by IES senior management based upon analyses overseen and directed by the IES Disclosure Review Board. Each memorandum of Understanding (MOU) with another federal program and each license with a qualified external researcher includes a description of the planned research and a statement explaining why publicly available data are not sufficient for the proposed analysis. The MOU/License process includes the use of a legally binding agreement with the other federal office, agency or institution; affidavits of nondisclosure from each authorized data user on the MOU/License; verification that the computer system on which the data will be used has full certification and accreditation in the case of a federal agency, alternatively verification that the data will only be used on a standalone computer in the case of a license; participation in unannounced security/compliance inspections; compliance with IES reporting standards for the specific data file; and the submission of all work products to IES for disclosure review prior to release to anyone not permitted access to the data through the researcher's MOU/License agreement.

Data that do not require MOU or License agreements have had PII removed and have been subject to additional data perturbations to prevent indirect identification. Researchers, policymakers, and other stakeholders use these data for their own research and information purposes. Similarly, stakeholders can access information through published reports in the IESDC. These reports do not contain PII.

* 1. 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

* 1. Does the project place limitation on re-disclosure?
	
	Yes

# Redress

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

Individuals must contact the system manager listed in the respective SORN and provide the necessary particulars listed in order to request access to their own information. the system manager and necessary particulars requested vary by study and SORN. In general, individuals are requested to provide name, contact information, and any other identifying information requested to distinguish between individuals with the same name.

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

Individuals must contact the system manager listed in the respective SORN and provide any identifying information requested to correct inaccurate or erroneous information.

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

Information regarding the procedures for correcting information can be found in each study's respective SORN.

# Safeguards*If you are unsure which safeguards will apply, please consult with your* [*ISSO*](https://share.ed.gov/teams/OCIO/IA/SEA/SEA_Collab/SitePages/ED%20ISSOs.aspx)*.*

* 1. Does the principal office work with their CSO/ISSO to build privacy & security into the system and build privacy extensions to the extent feasible?

Yes

* 1. Is an Authority to Operate (ATO) required?

	Yes
	2. Under [NIST FIPS Pub. 199](https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf), what is the security categorization of the system: **Low, Moderate, or High?**
	Moderate
	3. What administrative, technical, and physical safeguards are in place to protect the information?

The IES Data Center (IESDC) houses the survey collection and data dissemination websites for the Institute of Education Sciences (IES). IESDC is hosted in a secure environment in the Amazon Web Services (AWS) GovCloud (US) region. AWS GovCloud (US) is an isolated AWS Region designed to allow US government agencies and customers to move sensitive workloads into the cloud by addressing their specific regulatory and compliance requirements. AWS Govcloud (US) has Federal Risk and Authorization Management Program (FedRAMP) certification. IESDC is connected to the Internet via a certified Managed Trusted Internet Protocol Services (MTIPS) connection which includes a security operation center (SOC) as well as EINSTEIN monitoring. IESDC is rated Moderate by the FIPS 199 Security Categorization. IESDC received an Authority to Operate (ATO) on 11/21/17 and participates in yearly FISMA self-assessments. IESDC uses Access Conrol Lists (ACLs), firewalls, NTFS permissions, Microsoft BitLocker, McAfee ePO and Virus Scan, and BigIP filtering to protect the information.

The controls listed in 7.2 also help safeguard the information.

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

Yes

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

Yes

* 1. 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.

IESDC received an Authority to Operate (ATO) on 11/21/17 from a FISMA Security Assessment. These assessments take place every 3 years. IESDC participates in yearly FISMA self-assessments on the years the full assessments do not take place. IESDC performs monthly internal and external vulnerability scans on all systems. Additional vulnerability scans take place as necessary to make sure system changes do not cause security issues. IESDC runs daily quick virus scans and weekly full virus scans on all systems. System changes are tracked through the IES Members Site's change control system.

# Auditing and Accountability

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

All public access data have been approved for release by IES senior management, based upon the analysis and recommendations from the IES Disclosure Review Board. In addition, in those instances in which data are made available for analysis through the use of an on-line analysis tool, internal controls are programmed into the analysis tools to avoid potential disclosure of any identifying information (e.g.. the suppression of small cells, limitations on the specific types of analyses available, constraints on the publication of unrounded and unweighted sample size, and electronic protections of the underlying data). Users of public-facing data are notified that use of such data for purposes other than statistical or research purposes (e.g., trying to identify individuals in a study) is subject to severe criminal and civil penalties. More sensitive data that are shared are governed by IES Memoranda of Understanding (MoU) or restricted-use data licenses. Each MoU or license with an approved qualified agency or researcher includes a description of the planned research and a statement explaining why publicly available data are not sufficient for the proposed analysis. IES reviews all analyses generated through such MoUs or licenses for disclosure risk before information is shared beyond those named in the MoU or license agreement.

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

Yes

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

Privacy risks would be unauthorized individuals gaining access to PII of respondents in IES data collections. Information provided in section 7 discuss how the information systems are used to protect the data. Discussion in question 8.1 discusses how data perturbations, specific adjustments to on-line analysis tools, and legal protections further prevent disclosure risks.

1. The term “personally identifiable information” refers to information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc. [OMB Circular A-130, page 33](https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf) [↑](#footnote-ref-1)
2. A System of Records Notice (SORN) is a formal notice to the public that identifies the purpose for which PII is collected, from whom and what type of PII is collected, how the PII is shared externally (routine uses), and how to access and correct any PII maintained by ED. <https://connected.ed.gov/om/Documents/SORN-Process.pdf> [↑](#footnote-ref-2)
3. Examples include restricted form filling, account verification, editing and validating information as it’s collected, and communication with the individual whose information it is. [↑](#footnote-ref-3)
4. If this information is covered by Privacy Act System of Records Notice (SORN) please list only relevant programmatic disclosures listed under the Routine Uses section. [↑](#footnote-ref-4)