

# PRIVACY IMPACT ASSESSMENT

(Rev. 2/2020)

(All Previous Editions Obsolete)

Please submit your responses to your Liaison Privacy Official. *All entries must be Times New Roman, 12pt, and start on the next line.* If you need further assistance, contact your LPO. A listing of the LPOs can be found here:

[https://usepa.sharepoint.com/:w:/r/sites/oei\\_Community/OISP/Privacy/LPODoc/LPO%20Roster.docx](https://usepa.sharepoint.com/:w:/r/sites/oei_Community/OISP/Privacy/LPODoc/LPO%20Roster.docx)

<b>System Name: Coal Combustion Residuals, Water Quality, and Fetal Health</b>	<b>System Owner: Wes Austin</b>
<b>Preparer: Wes Austin</b>	<b>Office: AO/OP/NCEE</b>
<b>Date: 2/5/2021</b>	<b>Phone: 202-566-2397</b>
<b>Reason for Submittal: New PIA <input checked="" type="checkbox"/> Revised PIA <input type="checkbox"/> Annual Review <input type="checkbox"/> Rescindment <input type="checkbox"/></b>	
<b>This system is in the following life cycle stage(s):</b>	
Definition <input type="checkbox"/> Development/Acquisition <input checked="" type="checkbox"/> Implementation <input type="checkbox"/>	
Operation & Maintenance <input type="checkbox"/> Rescindment/Decommissioned <input type="checkbox"/>	
<b>Note: New and Existing Systems require a PIA annually, when there is a significant modification to the system or where privacy risk has increased to the system. For examples of significant modifications, see <u><a href="#">OMB Circular A-130, Appendix 1, Section (c) (1) (a-f)</a></u>.</b>	
<b>The PIA must describe the risk associated with that action. For assistance in applying privacy risk see <u><a href="#">OMB Circular No. A-123, Section VII (A) (pgs. 44-45)</a></u>.</b>	

## **Provide a general description/overview and purpose of the system:**

The purpose of the system is to facilitate publication of dissertation research. The publication process will require access to the core data used in the dissertation research analysis. To provide some background, my dissertation assessed the relationship between coal ash surface water discharges and three relevant outcomes: surface water quality, municipal system water quality, and fetal health indicators from a birth certificate database in North Carolina. While almost all of the data used for this project is publicly available, the birth certificates required special approval to access and use for research purposes. Specifically, I acquired Institutional Review Board approval from Georgia State University, HIPAA approval, and approval from the State Registrar of North Carolina to store birth certificate information including SPII on a secure server at Georgia State University. A working paper of the study is now complete and publicly available. Proceeding with publication will require additional sensitivity tests and

analyses that I cannot continue to perform from my former secure workstation at Georgia State University. I have arranged a secure data transference procedure with my former employers to move this data to my current EPA work computer through an encrypted volume, secure transport server, virtual private network connection, and encrypted transfer using VeraCrypt software. Pending official Office of General Counsel, IRB, PIA, and Human Research Ethics Council approval, the State Center for Health Statistics in North Carolina will review and approve a revised data use agreement allowing the transfer of data to my new workstation at the EPA. The data to be stored on my secure government-provided work laptop include birth certificate information from North Carolina, which includes SPII in the form of mothers' names and addresses at time of birth. To this will be merged publicly available information including Toxic Release Inventory data, municipal water system service zones, a property parcel database, and air quality metrics.

## **Section 1.0 Authorities and Other Requirements**

### **1.1 What specific legal authorities and/or Executive Order(s) permit and define the collection of information by the system in question?**

The study aims to provide insights about the importance of coal combustion residuals with respect to EPA's authority in regulating waters of the United States through the Clean Water Act, drinking water of the United States through the Safe Drinking Water Act, and waste materials of the United States through the Resource Conservation and Recovery Act, specifically Subtitle D of this Act covering management of coal ash wastes.

Executive Order 13045, "Protection of Children from Environmental Health Risks and Safety Risks" (1997), states that each Federal agency shall ensure that its policies, programs, activities, and standards address disproportionate risks to children that result from environmental health or safety risks. The Environmental Protection Agency directly regulates the disposal of coal combustion residuals and release of waste effluent from the same through subtitle D of the Resource Conservation and Recovery Act as well as through the Effluent Limitation Guidelines of the Clean Water Act. Executive Order 13045 therefore directly charges the EPA with ensuring these policies do not pose disproportionate risk to newborns. The data used in this study allows the researchers to study how water quality fluctuations from coal combustion residual pollution, which are likely less harmful to adults, may adversely impact the fetal environment and contribute to such child health disparities.

Executive Order 12898, "Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations" (1994), calls on each Federal agency to make achieving EJ part of its mission "by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations and low-income populations." My study demonstrates that Black and Hispanic mothers are more likely to be exposed to this form of pollution than white mothers, with resulting disproportionate adverse health effects. The Environmental Protection Agency regulates the disposal of coal combustion residuals and release of waste effluent from the same through subtitle D of the Resource Conservation

and Recovery Act as well as through the Effluent Limitation Guidelines of the Clean Water Act. Executive Order 12898 charges the EPA with assessing the disproportionate impact of these rules on populations of concern; the data used in this study allows such assessment.

By using data to assess the environmental and human health impact of EPA policies, this project falls under the scope of the Foundations for Evidence-Based Policymaking Act of 2018 and OMB Circular A-11. OMB Circular A-11, "Preparation, Submission, and Execution of the Budget," states that, "Agencies should strengthen the use of evidence (defined in section 200.22) and data to drive better decision making and more effectively deliver on mission." Section 200.22 explains, "As defined by 5 U.S.C. 311(3), 'evaluation' means an assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency. Evaluations can provide critical information to inform decisions about current and future programming, policies, and organizational operations." Finally, Circular A-11 notes that, "The Foundations for Evidence-Based Policymaking Act of 2018 (or "Evidence Act") advances data and evidence-building functions in the Federal government by statutorily mandating Federal evidence-building activities, open government data, and confidential information protection and statistical efficiency."

**1.2 Has a system security plan been completed for the information system(s) supporting the system? Does the system have or will the system be issued an Authorization-to-Operate? When does the ATO expire?**

This is not an actual system or application but a process that involves me receiving data stored on a secure server in GSU's data center or at the North Carolina birth statistics office. The birth certificate records exist in three State datasets that will be stored on my secure government-issued laptop with access only granted to myself. The ATO and associated security documentation is a part of the agency's Enterprise Services System (ESS), General Support System which falls under the control of the Office of Mission Support (OMS). This ATO is dated to This ATO is dated to expire on May 21, 2021."

**1.3 If the information is covered by the Paperwork Reduction Act (PRA), provide the OMB Control number and the agency number for the collection. If there are multiple forms, include a list in an appendix.**

An Information Collection Request (ICR) is not required.

**1.4 Will the data be maintained or stored in a Cloud? If so, is the Cloud Service Provider (CSP) FedRamp approved? What type of service (PaaS, IaaS, SaaS, etc.) will the CSP provide?**

Yes, the data will be stored on the MSO365 One Drive - SaaS.

## **Section 2.0 Characterization of the Information**

*The following questions are intended to define the scope of the information requested and/or collected, as well as reasons for its collection.*

**2.1 Identify the information the system collects, uses, disseminates, or maintains (e.g., data elements, including name, address, DOB, SSN).**

The data consist of existing administrative records compiled by the State Center for Health Statistics of North Carolina in the form of birth certificate forms. The data includes confidential information, or SPII, in the form of mother's names, mother's date of birth, children's names, and local home address at time of birth. The data does not include social security numbers. These birth certificate forms also include a variety of conditions of the newborn such as birthweight, estimated gestation length, delivery method, date and time of birth, and presence of a congenital anomaly. Mother's information includes lifestyle attributes such as smoking behavior, age, weight, number of prenatal visits, educational history, and race/ethnicity. Father information includes race/ethnicity, educational status, and age. The birth statistics is merged to several publicly available pieces of information discussed in the reply to question 2.3.

**2.2 What are the sources of the information and how is the information collected for the system?**

The source of the information is the State Center for Health Statistics of North Carolina. The Center collects this information as part of its mandate to record natality and death records for the state of North Carolina.

**2.3 Does the system use information from commercial sources or publicly available data? If so, explain why and how this information is used.**

Yes. The birth records are merged to a publicly available property parcel geodatabase for the state of North Carolina. I then overlay a publicly available municipal water service zone geodatabase on the parcel database, allowing me to characterize births in the state as having occurred while a given mother (i.e., address) was serviced by a specific water company. Finally, these records are linked to a geodatabase of municipal water systems affected by coal combustion residuals, which is provided freely by the Southern Environmental Law Center (SELC). For each water system affected by coal ash according to the SELC, I characterize upstream coal ash pollution releases using the Toxic Releases Inventory. Therefore, for each birth record in the state of North Carolina, I may make comparisons not simply across siblings of the same mother, but also across years in which more or less pollution is released upstream from the mother's water system intake. Finally, I link these births to satellite-based air quality estimates created by the Atmospheric Composition Analysis Group at Dalhousie University. With the exception of birth certificates, all records are freely available to the public.

**2.4 Discuss how accuracy of the data is ensured.**

The institutions that collect the data (State Center for Health Statistics of North Carolina) are responsible for ensuring accuracy of the data.

## **2.5 Privacy Impact Analysis: Related to Characterization of the Information**

*Discuss the privacy risks identified for the specific data elements and for each risk explain how it was mitigated. Specific risks may be inherent in the sources or methods of collection, or the quality or quantity of information included.*

### **Privacy Risk:**

There is a small risk that the data on the birth certificates, such as conditions of the newborn or mother's address at time of birth, could be released to non-authorized users.

### **Mitigation:**

Only authorized users will have access to the data, and access controls described in 3.1 will be used.

## **Section 3.0 Access and Data Retention by the System**

*The following questions are intended to outline the access controls for the system and how long the system retains the information after the initial collection.*

### **3.1 Do the systems have access control levels within the system to prevent authorized users from accessing information they don't have a need to know? If so, what control levels have been put in place? If no controls are in place why have they been omitted?**

Yes. Data will be stored electronically, and only authorized users will have access. There will be one authorized user, and I will have access to all data elements. I am an EPA employee, not a contractor, and I am included on the Institutional Review Board (IRB) application through Georgia State University. Access will be restricted to authorized user by requiring computer log-on with unique user account, PIN number, and FIP card. Data will be uploaded to individual user's local hard drives and the MS0365 One Drive, thus preventing access by anyone other than the authorized user. The controls meet the moderate security controls required by NIST publication 800-53 revision 4, including 2-factor authentication (PIN number and FIP card).

### **3.2 In what policy/procedure are the access controls identified in 3.1, documented?**

The access controls identified in 3.1 will be documented in my modified IRB application under question 7.15 and further in section 10.3.

### **3.3 Are there other components with assigned roles and responsibilities within the system?**

No.

### **3.4 Who (internal and external parties) will have access to the**

**data/information in the system? If contractors, are the appropriate Federal Acquisition Regulation (FAR) clauses included in the contract?**

I am the only party who will have access to the data. No contractor will have access to the data.

**3.5 Explain how long and for what reasons the information is retained. Does the system have an EPA Records Control Schedule? If so, provide the schedule number.**

I estimate that the data will need to be retained for approximately one and a half years to allow for continuity of research plan, data analysis, revision of research papers, and publication of research papers. This project will use the existing EPA Records Schedule 1035 and NARA Disposal Authority DAA-0412-2013-0021-0005 (<http://intranet.epa.gov/records/schedule/final/1035.html>)

**3.6 Privacy Impact Analysis: Related to Retention**

*Discuss the risks associated with the length of time data is retained. How were those risks mitigated? The schedule should align the stated purpose and mission of the system.*

**Privacy Risk:**

The small risk that birth-specific data will be released to non-authorized users will remain during the length of time the records are retained. The schedule is determined by the research needs and the length of time to publish aggregate results in peer-reviewed journals.

**Mitigation:**

Only authorized users will have access to the data the entire time it is stored in EPA systems. Records will be permanently deleted consistent with NARA Disposal Authority DAA-0412-2013-0021-0005 when no longer needed.

**Section 4.0 Information Sharing**

*The following questions are intended to describe the scope of the system information sharing external to the Agency. External sharing encompasses sharing with other federal, state and local government, and third-party private sector entities.*

**4.1 Is information shared outside of EPA as part of the normal agency operations? If so, identify the organization(s), how the information is accessed and how it is to be used, and any agreements that apply.**

No, EPA will not share individual-level birth certificate data received from the State Center of Health Statistics with any systems, agencies, governments, or other external parties. The authorized user is the only EPA researcher with access to the data. Research results,

including aggregate statistics (but no individual-level data), will be disseminated in research reports or peer-reviewed journals.

#### **4.2 Describe how the external sharing is compatible with the original purposes of the collection.**

Disseminating research results through working papers and peer-reviewed publications is consistent with the purpose of the project, which is to gain insights about the potential for coal combustion residual water pollution to impact fetal health and to share these insights with public water systems, the state of North Carolina, the research community, and EPA rule makers to aid in implementation of the Safe Drinking Water Act, the Clean Water Act, and the Resource Conservation and Recovery Act Subtitle D.

#### **4.3 How does the system review and approve information sharing agreements, MOUs, new uses of the information, new access to the system by organizations within EPA and outside?**

Not applicable because EPA will not share individual data with any systems, Departments/agencies, governments, or other external parties.

#### **4.4 Does the agreement place limitations on re-dissemination?**

No.

#### **4.5 Privacy Impact Analysis: Related to Information Sharing**

*Discuss the privacy risks associated with the sharing of information outside of the agency. How were those risks mitigated?*

##### **Privacy Risk:**

None because individual data are not shared outside the agency.

##### **Mitigation:**

None. Only aggregate statistics will be disseminated.

## **Section 5.0 Auditing and Accountability**

*The following questions are intended to describe technical and policy-based safeguards and security measures.*

#### **5.1 How does the system ensure that the information is used as stated in Section 6.1?**

Review by the Institutional Review Board (IRB) for the initial project will be amended to reflect the new storage of information on secure government-issued computer with PIN card and password authentication will ensure that the information is used as stated in Section 6.1. In addition, the IRB will review and renew our application annually.

**5.2 Describe what privacy training is provided to users either generally or specifically relevant to the system/collection.**

All authorized users will take annual EPA Information Security and Privacy Awareness Training identifying the user's role and responsibilities for protecting information resources, as well as consequences for not adhering to the policy.

**5.3 Privacy Impact Analysis: Related to Auditing and Accountability**

**Privacy Risk:**

Risk would occur if authorized user is not trained properly on privacy and accountability.

**Mitigation:**

Annual Information Security and Privacy training is required to be completed by all EPA staff. I have completed this training.

**Section 6.0 Uses of the Information**

*The following questions require a clear description of the system's use of information.*

**6.1 Describe how and why the system uses the information.**

The authorized user is a research economist at the EPA's National Center for Environmental Economics and will use the data to complete and publish a study on the health impacts of coal combustion residual pollution. The study will require econometric analyses and sensitivity tests to demonstrate the robustness of the findings to journal reviewers. The study aims to provide insights about the importance of coal combustion residuals with respect to the EPA's mandate in regulating waters of the United States through the Clean Water Act, drinking water of the United States through the Safe Drinking Water Act, and waste materials of the United States through the Resource Conservation and Recovery Act.

**6.2 How is the system designed to retrieve information by the user? Will it be retrieved by personal identifier? Yes\_\_\_ No\_X\_. If yes, what identifier(s) will be used. (A personal identifier is a name, social security number or other identifying symbol assigned to an individual, i.e. any identifier unique to an individual. Or any identifier that can be linked or is linkable to an individual.)**

The dataset is designed to conduct large-scale statistical analysis of the extent to which coal combustion residuals impact fetal health; users therefore only retrieve aggregate statistics of population-level relationships of interest. The system is not designed to retrieve individual records. Users will only retrieve summary statistics (e.g., means) over the entire population or sub-populations that are and are not exposed to coal ash.

### **6.3 What type of evaluation has been conducted on the probable or potential effect of the privacy of individuals whose information is maintained in the system of records?**

Data will be stored electronically, and only one authorized user will have access. The authorized user is an EPA employee (not contractors). Access will be restricted to authorized user by requiring computer log-on with unique user account, PIN number, and FIP card. Data will be uploaded only to individual user's local hard drive and the MS0365 One Drive, thus preventing access by anyone other than the authorized user. The controls meet the moderate security controls required by NIST publication 800-53 revision 4, including 2-factor authentication (PIN number and FIP card).

### **6.4 Privacy Impact Analysis: Related to the Uses of Information**

*Describe any types of controls that may be in place to ensure that information is handled in accordance with the uses described above.*

#### **Privacy Risk:**

There is a small risk that birth certificate information could be used by non-authorized users.

#### **Mitigation:**

Only authorized user will have access to the data, and access controls described in 3.1 will be used.

**\*If no SORN is required, STOP HERE.**

*The NPP will determine if a SORN is required. If so, additional sections will be required.*

## **Section 7.0 Notice**

*The following questions seek information about the system's notice to the individual about the information collected, the right to consent to uses of information, and the right to decline to provide information.*

**7.1 How does the system provide individuals notice prior to the collection of information? If notice is not provided, explain why not.**

**7.2 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the collection or sharing of their information?**

**7.3 Privacy Impact Analysis: Related to Notice**

*Discuss how the notice provided corresponds to the purpose of the project and the stated uses. Discuss how the notice given for the initial collection is consistent with the stated use(s) of the information. Describe how the project has mitigated the risks associated with potentially insufficient notice and opportunity to decline or consent.*

**Privacy Risk:**

**Mitigation:**

## **Section 8.0 Redress**

*The following questions seek information about processes in place for individuals to seek redress which may include access to records about themselves, ensuring the accuracy of the information collected about them, and/or filing complaints.*

**8.1 What are the procedures that allow individuals to access their information?**

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

**8.3 Privacy Impact Analysis: Related to Redress**

*Discuss what, if any, redress program the project provides beyond the access and correction afforded under the Privacy Act and FOIA.*

**Privacy Risk:**

**Mitigation:**