Privacy Impact Assessment
for the

Evidence Platform

MCC/PIA-001

July 14, 2021

Contact Point of Contact

Miguel G. Adams
Chief Information Security Officer and
Deputy Privacy Officer
Department of Administration and Finance
Millennium Challenge Corporation
(202) 521-3574

Reviewing Official

James C. Porter
Chief Information Officer and
Chief Privacy Officer
Department of Administration and Finance
Millennium Challenge Corporation
(202) 521-3716

Abstract

The Evidence Platform is a public-facing platform that will be used for disseminating the data and documentation produced by data activities funded by the Millennium Challenge Corporation (MCC). There are three events where the collection of personally identifiable information will be required for the public to obtain the benefit of accessing MCC-funded data and documentation.

Overview

MCC is committed to delivering sustainable economic growth and poverty reduction throughout the lifecycle of its investments. MCC's evidence-based approach is rooted in that mission, requiring investment in data activities that predict and measure the results of MCC funded activities.

MCC's Guidelines for Transparent, Reproducible, and Ethical Data and Documentation (TREDD - https://www.mcc.gov/resources/doc/guidance-mcc-guidelines-tredd) govern the data and documentation sharing, which is predicated on the principles of accountability, transparency, and learning:

- **Accountability** refers to the obligation to report on and accept responsibility for all funded activities.
- **Transparency** refers to disclosing data activity analysis and findings (both data and documentation) in a public and transparent manner and share the information.
- **Learning** refers to improving the understanding of the measured results of funded activities, particularly in terms of poverty reduction and growth to improve current and future activities.

The Evidence Platform will be a **public-facing platform for disseminating the data and documentation produced by MCC-funded data activities**. The platform will directly replace the existing MCC Evaluation Catalog (https://data.mcc.gov/evaluations/index.php/catalog) and broaden the platform's scope and functionality to share the data and documentation of any MCC-funded data activity.

Through a full and open competition, MCC contracted with the University of Michigan's Interuniversity Consortium for Political and Social Research (ICPSR - https://www.icpsr.umich.edu/web/pages/) to leverage their existing data stewardship and data platform systems to meet MCC's objectives for the public facing platform. The system will:

- **Store and share documentation and de-identified data for public-use** – To directly replace the MCC Evaluation Catalog, the platform will be the mechanism by which MCC shares data

and documentation that is de-identified by MCC contractors, reviewed for clearance by the MCC Disclosure Review Board (DRB), and posted to the new platform.

- **Store and share restricted-access data** – Given promises of confidentiality and MCC's commitments to protecting the privacy of data activity participants, some data cannot be de-identified in a way that reduces re-identification risk and retains the usability of the data for accountability and learning objectives. In cases where limited data sharing is facilitated by the informed consent process, this data can be prepared for sharing through a restricted-access mechanism which carefully protects access to data for specific statistical analysis. The platform will be the mechanism by which MCC shares restricted-access data through an ICPSR-managed Virtual Data Enclave (VDE) following preparation by MCC contractors, review by the MCC DRB, and deposit of the restricted-access data with ICPSR.
- **Collect data on users** – The system collects tracking data and under certain circumstances, the system will collect personally identifiable information (PII) on users through three channels:
    - **Website Visitors**. The system automatically collects and stores tracking data for website visitors for statistical purposes and to help make the site more useful to visitors: (i) IP address of the computer used to access the site; (ii) geographic coordinates based upon the computer's IP address; (iii) date and time of the visit; (iv) pages the user visited; (v) address of the website the user came from when visiting the ICPSR site; (vi) operating system of the user's computer; (vii) version and type of browser used when visiting the site.
    - **Public-use Data Users**. The system collects PII and other data on users who download public-use data to allow for better reporting about who is using the data and to communicate with data users in the event more information about an object they have downloaded can be provided or to request data destruction. The following information is collected through the public use data users ICPSR MyData account: (i) email address; (ii) password (created by user); (iii) first and last names; (iv) type of organizational affiliation; and (v) department/field of specialization (Economics, Political Science, Psychology, etc). Postal address is collected if provided, but that information is not required to enroll in MyData.
    - **Restricted-use Data Users**. In addition to the ICPSR MyData account PII described above, data users requesting access to restricted-use data must submit the following information: (i) Restricted Data Use Agreement (RDUA) signed by both the data user and a representative of their institution; (ii) Documentation of Institutional Review Board (IRB) approval or exemption; (iii) research proposal; (iv) name and contact information of all researchers at the institution who will have access to the data; (v) list of data sets requested and why needed; and (vi) CV/Resume/Biosketch for each user that will access the VDE. When the data user's application is ready to approve,

ICPSR collects the applicant's responses to four knowledge check questions in order to confirm that applicant has watched the Virtual Data Enclave training video. After an application is approved, the applicant provides the following information for each member of their team in order to create a UM Sponsored Account, which is used to create the VDE user accounts: (i) full name; (ii) email address; (iii) home address; (iv) gender; and (v) birthday.

In summary, PII collected includes the following: (i) email address; (ii) first and last name; (iii) type of organizational affiliation; (iv) department/field of specialization; (v) postal address; (vi) name and contact information of all researchers at the institution who will have access to the data; (vii) gender; and (viii) birthday.

Section 1.0 Authorities and Other Requirements

1.1 What specific legal authorities and/or agreements permit and define the collection of information by the project in question?

The Paperwork Reduction Act of 1995 (44 U.S.C. 3501 et. Seq.) authorizes the Millennium Challenge Corporation's (MCC) collection of public information and governs who has access to that information.

1.2 What Privacy Act System of Records Notice(s) (SORN(s)) apply to the information?

MCC is in the process of creating a SORN specific to the Evidence Platform and is in the process of filing it with the Office of Management and Budget (OMB) – the Federal Register.

1.3 Has a system security plan been completed for the information system(s) supporting the project?

A system security plan is currently being drafted for the Evidence Platform.

1.4 Does a records retention schedule approved by the National Archives and Records Administration (NARA) exist?

Yes. MCC maintains an approved records retention schedule for information collected on Restricted-Use Data Users. No records retention schedule is required for website visitor data or public-use data users.

1.5 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.

The information collected on Restricted-use Data Users is covered by the Paperwork Reduction Act. MCC is the process of filing the Information Request with OMB.

Section 2.0 Characterization of the Information

2.1 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.

The information collected on Restricted-use Data Users is covered by the Paperwork Reduction Act. MCC is the process of filing the Information Request with OMB.

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

The source of information collected on data users is collected through three channels:

- **Website Visitors**. The system automatically collects and stores tracking data for website visitors for statistical purposes and to help make the site more useful to visitors: (i) IP address of the computer used to access the site; (ii) geographic coordinates based upon the computer's IP address; (iii) date and time of the visit; (iv) pages the user visited; (v) address of the website the user came from when visiting the ICPSR site; (vi) operating system of the user's computer; (vii) version and type of browser used when visiting the site.
- **Public-use Data Users**. The system collects PII and other data on users who download public-use data to allow for better reporting about who is using the data and to communicate with data users in the event more information about an object they have downloaded can be provided or to request data destruction. The following information is collected through the public use data users ICPSR MyData account: (i) email address; (ii) password (created by user); (iii) first and last names; (iv) type of organizational affiliation; and (v) department/field of specialization (Economics, Political Science, Psychology, etc). Postal address is collected if provided, but that information is not required to enroll in MyData.
- **Restricted-use Data Users**. In addition to the ICPSR MyData account PII described above, data users requesting access to restricted-use data must submit the following information: (i) Restricted Data Use Agreement (RDUA) signed by both the data user and a representative of their institution; (ii) Documentation of Institutional Review Board (IRB) approval or exemption; (iii) research proposal; (iv) name and contact information of all researchers at the institution who will have access to the data; (v) list of data sets requested

and why needed; and (vi) CV/Resume/Biosketch for each user that will access the VDE. When the data user's application is ready to approve, ICPSR collects the applicant's responses to four knowledge check questions in order to confirm that applicant has watched the Virtual Data Enclave training video. After an application is approved, the applicant provides the following information for each member of their team in order to create a UM Sponsored Account, which is used to create the VDE user accounts: (i) full name; (ii) email address; (iii) home address; (iv) gender; and (v) birthday.

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

The Evidence Platform does not itself use information from commercial source or publicly available data; however, the Evidence Platform does use Google Analytics concerning the characteristics and activities of users for reporting purposes. Google Analytics collects anonymized data on site visitors, aggregates it, and offers reports on where the traffic is coming from, what pages they browsed, for how long, etc. Google Analytics uses a first-party browser cookie containing a randomly-generated persistent ClientID.

2.4 Discuss how accuracy of the data is ensured.

For website visitors, the Evidence Platform relies on the browser identification string as noted in Section 2.2.

For public-use data users, the Evidence Platform uses email verification for newly created MyData accounts. When a user creates a new MyData account, the Evidence Platform sends a verification message to the email address that was used. To complete the account creation process, one must then click on a link (URL) contained in the verification message. This verification step is similar to the process used at service providers, social networks sites, online merchants, etc. Public-use data users must be logged into their MyData account in order to download data.

For restricted-use data users, ICPSR project staff confirm the researchers' information online (institution email, institutional profile, research journal publications, grants, etc.) to verify that they are affiliated with the institution in the RDUA. They also confirm the institutional representative's information in a similar manner to ensure they do indeed have legal signatory authority.

2.5 Privacy Impact Analysis: Related to Characterization of the Information

Privacy Risk:
As noted in section 2.2, MyData account enrollment collects email address password, full name, organizational affiliation, and (v) department/field of specialization. Postal address is collected if provided, but that information is not required to enroll in MyData.

In the event of a system breach, exposure of these data elements to unauthorized entities would pose a risk to system users:

- compromised email address and password would pose a risk of having additional information accessed on other sites for individuals who use those same credentials to access other systems or services, and
- compromised postal address could contribute to the risk of identity theft.

Mitigation:
In the event of a breach, ICPSR will follow University of Michigan Information Security Incident Reporting (SPG 601.25).

Privacy Risk:
The information collected from website visitors (IP address, pages visited, date and time of the visit) is collected automatically and does not require individual participation. If the data collection mechanism is unreliable or inaccurate, then the statistics derived from this information may not truly reflect user behavior.

Mitigation:
Ensure that the collection mechanism uses proven and tested technology and that it has the same downtime and uptime schedule as the website.

Section 3.0 Uses of the Information

3.1 Describe how and why the project uses the information.

Information on Website Visitors is used by ICPSR for statistical purposes and to help make the site more useful to visitors. Information collected on Public-use Data Users allows ICPSR to report on who is using the data and to communicate with the users if more information about an object that they have downloaded can be provided or to request data destruction. Information about Restricted-use Data Users allows ICPSR to determine if the user meets MCC's criteria for access to Restricted-use data in the Virtual Data Enclaves. ICPSR may report to MCC aggregate information on users – such total numbers and distributions by geography and sector. MCC may also request ICPSR to send requests to users, such as requests to share any knowledge products produced because of the data users' access to the public and/or restricted-access data. Information required for

the data user's UM/VDE accounts is used to secure their login information and/or assist in the event the user requires a password reset.

3.2 Does the project use technology to conduct electronic searches, queries, or analyses in an electronic database to discover or locate a predictive pattern or an anomaly? If so, state how MCC plans to use such results.

No.

3.3 Are there other Components with assigned roles and responsibilities within the system?

MCC staff will create MyData Account profiles and will be assigned roles and responsibilities in the system to upload and publish cleared data and documentation to the Evidence Platform.

3.4 Privacy Impact Analysis: Related to the Uses of Information.

Privacy Risk:
If ICPSR were to transform, modify, edit, or alter user data, then MCC would not have an accurate view of the users who visit the Evidence Platform and/or have a clear understanding of the users who have been granted access to the Restricted Use data.

Mitigation:

As defined in the Restricted-Use Data Deposit and Dissemination Agreement (RUDDDA) and Public-Use Data Deposit and Dissemination Agreement (PUDDDA), ICPSR may not alter any user information except for the limited purpose of standard data processing. ICPSR may not transform, modify, edit, or alter user information without the express written authorization of MCC.

Section 4.0 Notice

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

ICPSR's system provides notification through the following channels:

- **Website Visitors**. ICPSR's Privacy Policy - https://www.icpsr.umich.edu/web/pages/about/privacy.html - defines what data is collected on users to the website.
- **Public-use Data Users**. To access public data files, data users must establish a MyData account where they voluntarily provide the requested data in order to be granted access to the data files.
- **Restricted-use Data Users**. To access restricted data files, data users must establish a MyData account where they voluntarily provide the requested data to be granted access to the data files. They must also voluntarily provide the requested documentation for the restricted-access data and create a separate UM sponsored account as described above.

4.2 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the project?

ICPSR's system provides notification through the following channels:

- **Website Visitors**. ICPSR's Privacy Policy - https://www.icpsr.umich.edu/web/pages/about/privacy.html - defines what data is collected on users to the website. Data users may volunteer to continue using the site or opt out.
- **Public-use Data Users**. To access public data files, data users must establish a MyData account where they voluntarily provide the requested data to be granted access to the data files.
- **Restricted-use Data Users**. To access restricted data files, data users must establish a MyData account where they voluntarily provide the requested data to be granted access to the data files. They must also voluntarily provide the requested documentation for the restricted-access data and create a separate UM sponsored account as described above.

4.3 Privacy Impact Analysis: Related to Notice Privacy

Privacy Risk:
ICPSR could potentially use information collected for purposes not listed within the Privacy Policy.

Mitigation:
The Restricted-Use Data Deposit and Dissemination Agreement (RUDDDA) and the Public-Use Data Deposit and Dissemination Agreement (PUDDDA) states that ICPSR cannot use information collected from users for purposes not listed within the Privacy Policy.

Section 5.0 Data Retention

5.1 Explain how long and for what reason the information is retained.

MyData records are retained indefinitely for auditing purposes. If a user deletes their MyData account, the email address is deleted from the record, but the remaining elements are retained. Any federated login information is deleted as well. The risks of a breach for this PII is described in Section 2.5.

ICPSR will retain restricted data user information for a period of two years from the date that it is collected. This information will be used to determine and record user eligibility for access to Restricted-Use data, as well as regular reporting on website usage statistics. UM sponsored accounts have a default 1 year before renewal is required or expiration occurs. ICPSR renews these accounts as long as the data user has a continued active RDUA.

5.2 Privacy Impact Analysis: Related to Retention

Privacy Risk: ICPSR's internal system for data retention may fail.

Mitigation: For public-use data users, MyData database records are backed up nightly to tape as part of ICPSR routine backup processes. If there were a server failure, the maximum amount of information lost would constitute new accounts and downloads within the previous 24 hours. For restricted data users, ICPSR will submit user data annually to MCC following agreed record retention schedule.

Section 6.0 Information Sharing

6.1 Is information shared outside of MCC as part of the normal agency operations? If so, identify the organization(s) and how the information is accessed and how it is to be used.

The user information collected is not shared outside of MCC as a part of the normal agency operations.

6.3 Describe how the external sharing noted in 6.1 is compatible with the SORN noted in 1.2.

Not Applicable

6.3 Does the project place limitations on re-dissemination?

Not Applicable

6.4 Describe how the project maintains a record of any disclosures outside of the MCC.

Not Applicable

Section 7.0 Redress

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

Individuals can access their MyData accounts at any time through the MyData dashboard as long as their account remains active.

If they delete their account, they can no longer access that information even though some of that information is retained as per section 5.1.

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

MyData users can edit this information themselves directly through their account dashboard or contact the ICPSR user support help desk for assistance in updating erroneous information. In addition, if ICPSR finds a user with incorrect contact details (e.g., e-mail address) and the user has logged into their account in the past 12 months, ICPSR will contact the user to obtain permission to update the user's account. Users who do not respond within four weeks will have their MyData accounts manually edited and updated by ICPSR. If the user has not logged into their account in the past 12 months, ICPSR will manually edit the user's MyData account to reflect the new, corrected information.

Restricted data users are required, as per the Restricted Data Use Agreement, to provide annual reports to ICPSR staff upon request. Restricted data users are required to notify ICPSR of updates to their project description, IRB documentation, and staffing information, including notifying ICPSR of a change in institutional affiliation at least six weeks prior to the last day of employment and notifying ICPSR of the addition or removal of research staff as soon as reasonably possible. For UM/VDE accounts, the PI and/or project admin are able to modify the information provided via their project page. Lastly, restricted data users can also change any erroneous information at any time by contacting the ICPSR user support help desk for assistance.

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

MyData users are notified via links to the ICPSR user support help desk across all ICPSR sites and via their MyData account dashboard. Restricted data users are notified via the Restricted Data Use Agreement.

Section 8.0 Auditing and Accountability

8.1 How does the project ensure that the information is used in accordance with stated practices in this PIA?

The Evidence Platform will be audited in accordance with the MCC Information System Security Policy. The Chief Information Security Officer (CISO) reviews and updates the policy annually to ensure it remains compliant with federal law, external mandates and MCC business decisions.

8.2 Describe what privacy training is provided to users either generally or specifically relevant to the project.

All MCC contractors must complete the Information Security and Privacy training when partnering with the agency. All contractors must complete role-specific training if they will have access to PII. Contractors must adhere to yearly training certification requirements.

8.3 What procedures are in place to determine which users may access the information and how does the project determine who has access?

Contractors responsible for maintaining the Evidence Platform have specified identification and password-protected access to the system. Access to the Evidence Platform is role-based and on a need-to-know and least-privileged basis.

8.4 How does the project review and approve information sharing agreements, MOUs, new uses of the information, new access to the system by organizations within MCC and outside?

The project establishes a data sharing agreement between ICPSR and restricted-use data users through the Restricted Data Use Agreement.

Responsible Officials

MIGUEL G. ADAMS (affiliate)

Digitally signed by MIGUEL G. ADAMS (affiliate)
Date: 2021.07.14 13:04:56 -04'00'

Miguel G. Adams
Chief Information Security Officer and
Deputy Privacy Officer
Department of Administration and Finance
Millennium Challenge Corporation

Approving Signature

JAMES C PORTER (affiliate)

Digitally signed by JAMES C PORTER (affiliate)
Date: 2021.07.14 15:23:05 -04'00'

SIGNATURE
James C. Porter
Chief Information Officer and
Chief Privacy Officer
Department of Administration and Finance
Millennium Challenge Corporation