

Appendix: Data Sharing Plan

This project will generate computer code, data products, and research findings. All computer code will be shared on the ICPSR git repository (<https://github.com/ICPSR>). Links to the repository, as well as to all public use data products will be shared through LinkageLibrary (<https://linkagelibrary.icpsr.umich.edu/>), an NSF-funded repository at ICPSR to support the record linkage community. Research findings will be documented in publicly available working papers available via the University of Michigan's Deep Blue repository (<https://deepblue.lib.umich.edu/>), via Census Bureau Working Paper Series when applicable, and ultimately via published research papers.

ICPSR has a 60-year track record of reliably storing and distributing research data and documentation according to best practices. We have obtained numerous certifications and awards for our policies and practices in these areas. ICPSR was a test case for Trusted Repositories Audit & Certification (TRAC, now ISO 16363) and was one of the first six data repositories in the world to earn the Data Seal of Approval, in 2011. ICPSR has also achieved FISMA-Moderate "Authority to Operate" status, CoreTrustSeal accreditation and World Data System accreditation. The Consortium of European Social Science Data Archives (CESSDA) has chosen ICPSR as an "industry best" or "gold standard" organization by which to benchmark the levels and qualities of service that their national data archives provide to the research community.

In addition to public products, our project will produce resources protected by Title 13. Code and output written in a Title 13 environment will be carefully documented and archived in the FSRDCs and, when possible, cleared through the Disclosure Review Board and housed permanently at ICPSR. Restricted Panel Study of Income Dynamics (PSID) data can be requested by researchers via the Michigan Center for the Demography of Aging's Virtual Data Enclave at <https://simba.isr.umich.edu/restricted/RestrictedUse.aspx>. Criminal Justice Administrative Records System (CJARS) data is available at <https://cjars.isr.umich.edu/> or through the FSRDCs at <https://www.icpsr.umich.edu/web/pages/appfed/index.html>

For Priority Area 1 (development of new methods), we will disseminate preliminary versions of code, data products, and research findings as early as possible, to encourage feedback and iteration during the project period. The pilot products developed in this priority area will have high-quality documentation suitable for other researchers to understand and replicate team work. Any data files we generate will be documented following the Data Documentation Initiative (DDI) framework. DDI is an international standard for describing surveys, questionnaires, and statistical data files. ICPSR's lead archivist is the Director of the DDI Alliance.

Table 1 shows our anticipated product release plan.

Table 1. High-level tasks and deliverables for Priority Area 1

	Year 1	Year 2	Year 3	Year 4	Year 5
Priority Area 1: Develop new methods for linkage					
Develop ML-based approach	█	█			
Develop quality measures	█	█			
Release preliminary code		█			
Adjust new methods based on quality findings		█	█		
Document improvement over current methods		█	█		
Document how to use quality measures			█		
Release final pilot code and documentation			█		