



U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT

- Position Title:** Data Science and Curation Intern
- Dept/Office:** Bureau for Management, Office of the Chief Information Officer
- Team:** USAID Data Services Team (with multiple placement options)
- Location:** Interns will be physically located in USAID facilities in the downtown Washington, D.C. area and / or Arlington, VA
- Duration:** May 2018 - September 2018 (flexible / extendible)
- Compensation:** This is an unpaid internship. Interns will be responsible for covering housing and living expenses in the immediate Washington, D.C. area. Interns can expect to have access to multiple professional development and learning opportunities within USAID and in the D.C. area.

Background:

The United States Agency for International Development (USAID) is the world's premier international development agency and a catalytic actor driving development results. The Agency is committed to using data and evidence to accelerate innovative approaches in development and to transform development outcomes, and understands that opening data and sharing it widely can lead to new breakthroughs in support of the Agency's mission.

In October 2014, [USAID announced](#) the release of its first ever Development Data Policy (ADS Chapter 579). This policy established the [Development Data Library \(DDL\)](#) as the Agency's central digital repository for USAID-funded, machine readable data. It also directed Agency staff and implementing partners to register or submit USAID-funded data to the repository. The DDL serves as USAID's source for advancing the availability and access to rich, heterogeneous microdata. Most importantly, it facilitates reuse of data for the next generation of development work.

In 2018, USAID plans to launch a new, improved DDL platform that will catalyze the availability, access, and use of USAID data. The new DDL will offer a redesigned data submission process with improved navigation, enhanced organization for data holdings, and expanded features to accelerate data use, such as in-platform visualizations.



Position Description:

USAID Data Science and Curation Interns will be a part of USAID's data community working to advance the availability, accessibility, and re-use of development data. Interns will provide support related to managing data at multiple points throughout the information lifecycle. This includes supporting business processes related to identifying, reformatting, redacting, submitting, tagging, approving, publishing, and archiving datasets generated with USAID funding. The intern will provide analytical products and services using foreign assistance data and information resources, which may require running queries on databases, enterprise systems, and other primary data sources.

Each intern will, at minimum, support the launch of USAID's new Development Data Library (DDL), to include:

- a) Learning and apply best practices for preserving, manipulating, and transforming data to migrate complex datasets into the new DDL;
- b) Enhancing dataset metadata and documentation to improve the usability of datasets through the DDL;
- c) Learning and applying best practices for data publication and sharing to help advance the responsible release of datasets through the DDL;
- d) Developing training materials for USAID staff, partners, and public users who are new to the platform.

In addition, the intern will provide ancillary support to other USAID operating units based on the unique needs of that unit. Placement opportunities are as follows (generally limited to one intern per unit):

- Bureau for Management, Office of the Chief Information Officer (M/CIO)
- Bureau for Global Health (GH)
- Bureau for Economic Growth, Education, and Environment (E3)
- Bureau for Food Security (BFS)
- Global Development Lab, Center for Development Research (LAB/CDR)



Placement: Office of the Chief Information Officer (M/CIO)

- As part of the core team leading the launch of USAID's new Development Data Library (DDL), assist with all aspects of the DDL launch, to include data migration, feature testing, socialization, training / orientation, and communications.
- Support Agency-wide efforts to develop a data literacy training curriculum
- Contribute to strategic planning sessions focused on modernizing a multitude of USAID data-related websites
- Identify and perform analyses and visualizations using public datasets that demonstrate the benefits of data sharing and publication to USAID staff and partners;
- Assist with drafting official USAID guidance related to data standardization; creation of data management plans; obtaining informed consent during data collection; and protocols for conducting risk analysis;
- Support the launch and implementation of USAID's new Analytics, Data, Visualization, and Information Services (ADVISE) contract.

Placement: Bureau for Global Health (GH)

- Assist with the further development of open data training modules, for our bureau, field offices, and implementing partners.
- Assist with providing Bureau wide AOR/COR open data training as well as informational presentations.
- Assisting with converting training materials into online content
- Work with individual offices to develop internal guidance on the clearance of data packages
- Work directly with offices to assist with the preparation of data packages, for DDL submission
- Work with our various offices, and their Data Stewards, to develop project plans for improved data release.
- Work with implementing partners, through our various offices, and their Data Stewards, as well as relevant AOR/CORs, to develop Data Management Plans.

Placement: Bureau for Economic Growth, Education, and Environment (E3)

- Coordinate across E3/ED technical teams to review and prepare the 2018 Education Activity roster
- Support preparation for the 2018 Annual Data Call, including updating guidance and preparing data capture sheets as necessary
- Work with E3/ED and OUs to assist with the preparation of data packages for DDL submission
- Support communication with OUs to obtain clearance and prepare data packages for DDL submission



Placement: Bureau for Food Security (BFS)

- Assist AOR/CORs and partners with data quality review of Feed the Future datasets submitted to the Agency.
- Review Data Management Plan (DMP) inventory status and assist BFS Offices and partners with creating new and updating existing DMPs.
- Identify incomplete data package submittals and work with AOR/CORs and partners to update.
- Assist with refining training materials for partners, Washington bureau, and FTF missions and identify best practices for data curation and management.
- Assist with bureau-wide Open Data training to ensure compliance with policy.
- Assist with identifying and/or building data use cases to support USAID Open Data initiatives as they relate to, or support achieving Global Food Security Strategy goals and objectives. Provide support for a datathon or other events to raise awareness of and demonstrate the value of open development data.
- Identify existing food security data, and/or make recommendations for the curation of new data that could be used to support the development of a strategy to integrate machine learning tools into the Monitoring, Evaluation, and Learning (MEL) processes.

Placement: Global Development Lab, Center for Development Research

- Stocktaking of public access/ open data trainings currently existing across USAID and compiling relevant materials and content
- Create a list of suggested best practices from these materials for the whole agency to consider/adopt
- Create an SOP for rolling out public access/open data for PEER, HESN, and new mechanisms

Education:

- It is generally preferred that candidates be in the process of completing a graduate degree in economics, various sciences (including computer or data science), mathematics, or relevant quantitative subject area that requires an understanding of technical methods in data management to solve complex problems.

Skills:

- Experience working with development data or related disciplinary data (such as global/public health data, agriculture and food security data, social/behavioral or political science data, education data, environmental data, geospatial data)



- Ability to work with different types of data from various sources such as static (snapshots of a project at a single point in time) and dynamic (ever-changing data generated continuously)
- Ability to understand and work closely with data repository workflows and business processes
- Ability to work with minimal supervision; entrepreneurial and self-motivated
- Ability to work collaboratively with a variety of individuals at different levels within the Agency and outside the Agency on complex or politically sensitive issues

Citizenship / Clearance: Interns must be U.S. Citizens and will be required to pass a background check and interim clearance process to obtain facility and network access.

To Apply:

Please contact your **first choice** placement Point of Contact (POC) by emailing one of the addresses below. Your first choice placement should most match your skills and expectations for the internship. **Applications received prior to May 11** will receive priority consideration. When doing so, please also include:

- A copy of your resume
- A brief statement on why you are pursuing the internship
- A ranking of your 2nd, 3rd, 4th, and 5th placement preferences

Placement POCs:

- Bureau for Management, Office of the Chief Information Officer (M/CIO)
 - Scott Thompson <dataservices@usaid.gov>
- Bureau for Global Health (GH)
 - Jason Wucinski <jwucinski@usaid.gov>
- Bureau for Economic Growth, Education, and Environment (E3)
 - Ben Sylla <bsylla@usaid.gov>
- Bureau for Food Security (BFS)
 - Zachary Baquet <zbaquet@usaid.gov>
 - Anna Brenes <abrenes@usaid.gov>
- Global Development Lab, Center for Development Research (LAB/CDR)
 - Annica Wayman <awayman@usaid.gov>