Commercial ADRF Product Information

ADMINISTRATIVE DATA RESEARCH FACILITY (ADRF)

The Administrative Data Research Facility ("**ADRF**") is a cloud-based research analytics and collaboration system, which is deployed as a Software-as-a-Service ("**SaaS**"). The ADRF is hosted on the Federal Risk and Authorization Management Program ("**FedRAMP**")-authorized Amazon Web Services Government Community Cloud Infrastructure as a Service platform ("**AWS GovCloud**"). The ADRF is a FedRAMP authorized cloud service provider ("**FEDRamp authorized"**) and its current FedRAMP authorization status can be viewed here: https://marketplace.fedramp.gov/#!/products?sort=productName.

The ADRF provides a secure environment for research analysts to discover, access, and use securely-stored data, and allows data stewards to monitor and strictly manage the use of such data. Carefully balancing security and usability requirements, the ADRF provides a comprehensive facility for supporting effective data analytics across broad and previously unconnected datasets.

Each customer's ADRF instance ("**Enclave**") is designed, implemented and operated in accordance with the ADRF's FedRAMP authorization requirements.

ADRF ARCHITECTURE OVERVIEW

The ADRF is a SaaS platform hosted on the AWS GovCloud platform, which is authorized under the package ID F1603047866. AWS GovCloud is designed as a security policy and compliance framework for government data and therefore offers significant protections for commercial users. As a SaaS offering, ADRF is standalone and does not part of a larger enterprise architecture.

- ADRF Customer User Access - ADRF customer users access the environment using SSL through a web browser referencing the ADRF public URL. When connecting through the web (browser) client, the user does not need to install anything on their computer. All the authentication and permission mechanisms comply with the FedRAMP parameters. Users are authenticated using the following steps:

- The user will provide username, password and multi-factor credential;

- After a successful authentication following ADRF policies, the virtual data enclave will create a session for the user on one of the workspace nodes

- ADRF will grant access to the authenticated user and route the session through the firewall back to the user; and

- After the users connect to ADRF, they will have access to the customer data for which they are permissioned and the purchased analytical tools available within the ADRF.

SECURITY GUIDELINES

Key tenets of the ADRF security philosophy include the following:

- Implementing Defense in Depth – The ADRF production environment has been designed and built using a "defense in depth" strategy. See the section below for additional information.

- Maintaining System Availability Through Redundancy and Resiliency:

○ To enhance availability, the ADRF is hosted at multiple AWS GovCloud FedRAMP-authorized facilities. The utilization of AWS GovCloud FedRAMP IaaS facilities allows the ADRF to be architected to use AWS facilities as the primary processing and storage center for the ADRF production environment. Additionally, AWS hosting also provides for reliable operations and a readily available alternate processing and data storage site.

○ Separate virtual systems residing in different AWS GovCloud availability zones allow the ADRF to operate in a Hot (Active-Active) site mode. The separation of these systems provides redundancy and failover capabilities in case of a disaster, unavailable processing functionality or unpredicted outage.

- Providing Data Confidentiality and Integrity – The ADRF uses cryptography to protect the confidentiality and integrity of data at rest and during transmission. The confidentiality and integrity of data at the ADRF alternate processing/storage site are safeguarded in a manner identical to the data in the ADRF primary processing/storage site. Data stored at both sites is encrypted to safeguard the data using AWS Server-Side Encryption (SSE). FIPS-140-2 validated encryption is also used to protect data during transmission using VPNs and encrypted links.

SECURITY REQUIREMENTS

As a system originally designed to meet the strict compliance and security needs of Government data clients, the FedRAMP and National Institute of Standards and Technology ("**NIST**") Special Publication (SP) 800-53, Revision 4, *Security and Privacy Controls for Federal Information Systems and Organizations*, PL-8 (Information Security Architecture) has been applied to the ADRF. This requires that the developer of an information system must also develop a security architecture that describes the approach taken with regard to protecting the confidentiality, integrity, and availability of hosted information.

The ADRF has been categorized as a FIPS 199 Moderate Impact level system. As a result, the ADRF security requirements are defined by the FedRAMP Moderate Impact overlay to the NIST SP 800-53 Moderate Impact minimum security controls. Additional NIST SP 800-53 security control requirements defined by the ADRF Authorizing Official (AO) have also been incorporated into the ADRF information security architecture.

SECURITY APPROACH

The ADRF security approach is based on a "defense in depth" implementation model. The first level is protected at the VPC level. At the next level, the ADRF production environment is implemented using AWS security groups that segregate customer data, management traffic, backup information and security functionality. Finally, at the system-component level, the configuration settings for the components within the ADRF production environment reflect the most restrictive mode consistent with operational requirements.

The ADRF is designed so that each major system function is isolated through strictly regulated AWS GovCloud Availability Zones, the implementation of distinct authentication mechanisms, the implementation of distinct Management and Service VPCs, and the employment of encryption within the ADRF environment. Application user interfaces are further separated from application system services and functionality through the implementation of strict role-based access and a segregated environment that enforces logical access.

Save for duly authorized customer users, there is no public inbound access to the ADRF network. A comprehensive suite of AWS security tools is also deployed within the boundary to ensure advanced threat protection. These include, but are not limited to, the AWS GovCloud Security Tools.

The Coleridge Initiative conducts vulnerability scans on all servers before placing them into the ADRF production environment and subsequently conducts routine scans thereafter to identify potential risks and to develop appropriate mitigations. Vulnerability scans are also performed on any server when the operating system is reinstalled or reconfigured, a new service or application is added, or patches are applied. The operating framework requires that high-risk vulnerabilities are to be mitigated within thirty (30) days, moderate risk vulnerabilities are to be mitigated within ninety days (90) and low risk vulnerabilities are to be mitigated within 180 days. All vulnerability mitigations will be tracked in the ADRF support system.

ADRF CONTINUOUS MONITORING PROGRAM

The FedRAMP program has developed an ongoing assessment and authorization program for the purpose of maintaining the authorization of its authorized Cloud Service Providers ("**CSPs**"). After a CSP system receives a FedRAMP authorization, it is recognized that the security profile of the system could change over time due to changes in the deployed hardware or software on the cloud service offering, or also due to the discovery and propagation of new exploits over time as the state of the art changes. Ongoing assessment and authorization procedures provide customers using CSP hosted cloud services with a way of detecting changes to the security profile of a system for the purpose of making risk-based decisions.

The ADRF applies the FedRAMP continuous monitoring program, which is based on the continuous monitoring process described in NIST SP 800-137, *Information Security Continuous Monitoring for Federal Information Systems and Organizations* and the *FedRAMP Continuous Monitoring Strategy Guide*. As described more closely below, the goal of the ADRF FedRAMP continuous monitoring program is to provide: (i) operational visibility; (ii) managed change control; and (iii) attendance to incident response duties.

FIVE SAFES FRAMEWORK

The ADRF follows the "Five Safes" framework to provide controlled research access to sensitive or confidential data:

Safe Projects

The ADRF contains only customer-approved projects or enclaves that have been proposed and agreed upon by project owners and their organization's data stewards. Approved projects require signed agreements and only ADRF customers and their duly authorized users can access the relevant customer project workspace or enclave within the ADRF. Project workspaces or enclaves are isolated from one another with controlled access to resources, delineated by individual and group memberships. This ensures there is no shared environment between projects or enclaves and resources unless the data owner has determined that this will be the case.

All standard tools that are pre-installed as part of the Microsoft Windows 10 operating system are available.

In addition, the ADRF customer enclaves can include the following, as standard or at additional cost, as specified in the customer agreement:

Project Workspace Tools	
Database Access Tools	**Statistical Tools**
DBeaver	PyCharm
Microsoft SQL Server Management Studio	Python
	Jupyterlab
	R
	RStudio
	RTools
	Stata
Productivity Tools	**Other Tools**
LibreOffice	EmEditor
Firefox (*Offline)	Notepad++

Google Chrome (*Offline)	PDF Reader WordNet Java
Version Control Tool Tortoise SVNWinRaR	

Please note: We regularly add (and sometime remove) elements of the standard libraries and packages, including as part of our response to feedback received from customers and users from time to time. Additional cost services will be provided in accordance with the customer agreement.

Safe People

Only customer approved users are permitted to access project workspaces or enclaves and related resources. These individuals are required to go through our on-boarding process, as part of which they must sign Data Use Agreements, Terms of Use, and complete an approved security training module before being granted access to the relevant project or enclave. Data resource access is specifically granted based on project requirements and are provided strictly in a read-only mode to ensure the integrity of the source data. Researchers are required to provide a username/password in addition to a second factor when attempting to login. The security protocols (user identity, password complexity and expiration, etc) follow the strict FedRAMP guidelines and industry best practices. All user access activity is logged and monitored.

Safe Settings

The ADRF is designed to provide secure methods of data transfer for customer micro-data, including datasets that include Personally Identifiable Information (PII). Only customer-identified and duly authorized personnel are invited to perform data transfers. The transfer of data into the ADRF uses the FedRAMP Authorized FIPS 140-2 validated Kiteworks Secure Environment. It is restricted to upload operations only. Files do not need to be encrypted or password protected in advance of initiating the transfer. Additional security protocols include regular system and application vulnerability scanning and third-party penetration testing.

Safe Data

Hashing: In collaboration with our data curators and technology partners, the Coleridge Initiative has developed a stand-alone Windows-based application to help simplify and facilitate the hashing of data prior to transmission to the ADRF. The hashing application can be downloaded directly to the operator’s desktop and has no dependencies on external resources. It guides the user through the identification of the source file (with un-hashed data), selection of fields to hash, selection of basic data validation and identification of the target file to create (with hashed data). The default ADRF “salt” may be used or a custom salt can also be provided by the user.
Data Stewardship Application: The Data Stewardship web-based application, if included in a customer agreement, is positioned primarily as a management and monitoring console for project and data stewards. It provides detailed insight on project or enclave configurations, user activity, user onboarding status, and a snapshot of the overall cost of a project on the ADRF.

Safe Exports

The ADRF inhibits the unauthorized removal of information from within the secure enclave environment by users. Users seeking to export their work (e.g. summary data, analysis output, supporting code, etc.) must do so through the export module within the ADRF. The export module allows each customer’s users to verify that they are not requesting intermediate output and to provide the documentation needed for relevant Coleridge Initiative staff and/or customer data stewards to conduct a thorough disclosure review of the requested materials. Once an export request is initiated and depending upon the service levels being provided as stated in the customer agreement, Coleridge Initiative staff and/or customer data stewards perform an initial review in accordance with any given guidelines for each dataset used to generate output. The review will usually focus upon ensuring that proper cell suppression has been applied, that there are no complementary disclosures, that rounding and noise have been applied where appropriate, and that there are no references to specific observations or counts that would be inadvertently disclosive. Depending on the customer’s preferred approach, after the export request has passed initial review, it may then be given a final review by the appropriate customer data stewards before being released to the researcher. Coleridge Initiative staff maintain a log of export requests for auditing purposes and to evaluate subsequent requests by the same user for complementary disclosure.