System Security Assessment Plan Template

Template Version: 1.0

Template Date: Nov 15, 2025

Prepared by: Patricia Buendia, Patrick Shironoshita, Lars Eklund, David Molik, Natalie Meyers, Research Data Alliance (RDA) AI Data Visitation Working Group (AIDV WG) DV4RDA project

Description/Abstract

The RDA AIDV System Security Assessment Plan Template is a standardized framework developed by the Research Data Alliance's AI Data Visitation Working Group and others to evaluate and approve the security of data visitation systems (i.e., platforms that enable controlled, temporary access to sensitive research data without transfer). The template requires an outline of team composition, assessment scope, procedures, and deliverables for evaluating system code, security documentation, and AI model safeguards against standards such as NIST SP 800-171[1] and ISO/IEC 27018[2]. The template also emphasizes code review, authentication integrity, data leakage prevention, dependency and container security, and continuous monitoring. These measures support consistent approval decisions, remediation recommendations, and final sign-off to ensure that data visitation technologies meet rigorous cybersecurity and compliance requirements. This template is recommended for future Data Visitation System Security Assessments.

Table of Contents

[1] R. Ross and V. Pillitteri, "Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations," National Institute of Standards and Technology, NIST Special Publication (SP) 800-171 Rev. 3, May 2024. doi: 10.6028/NIST.SP.800-171r3.

[2] ISO/IEC 27018:2025," ISO. Accessed: Nov. 17, 2025. [Online]. Available: https://www.iso.org/standard/27018

System Security Assessment Plan

This document outlines the security assessment plan for a data visiting system.

System Name: _____

System Version: _____

Date Completed: _____

Prepared by: _____

1. Selection of an Assessment Team by Expertise

To ensure agility and efficient task distribution, a focused, small team should be formed for the assessment. Use the template to assign persons to essential roles according to expertise.

Identify System Security Evaluators for your assessment: A team of independent, certified security professionals should conduct the assessment. These experts can be identified through professional networks, cybersecurity firms specializing in disciplinary specific domains like healthcare, or referrals from trusted partners.

Objective: The primary objective is for the team to assess and approve the data visiting platform's system security implementations. Actionable recommendations may be provided to further enhance security.

- **Code Evaluators (Computer Science, Cybersecurity, and System Security Background):** Source code inspection should be restricted to individuals with a demonstrable background in computer science, including relevant degrees or substantial professional experience in software development and security analysis. Some systems require specific subspecialties like digital biosecurity, or cyberbiosecurity experts.

  Task: These individuals will perform code reviews, identifying vulnerabilities, and providing actionable recommendations.

  Requirements: These individuals must have a computer science background with system security specialization may need to sign an NDA to access proprietary computer code.

  - Name with link to Profile: _____
  - Name with link to Profile: _____
  - Name with link to Profile: _____

  Estimate Code Evaluators' Time Effort: (e.g. 2-3 hours and/or 2 weeks to complete task): _____

- **System Security Plan** (**SSP) Evaluators (System Security Background):** Individuals reviewing the SSP may possess a broader range of backgrounds, including security policy experts, compliance officers, and risk management professionals.

  Task: These individuals will evaluate the SSP for completeness, accuracy, and adherence to relevant standards.

  Requirements: These individuals must be either security policy experts, compliance officers, or risk management professionals.

  - Name with link to Profile: _____
  - Name with link to Profile: _____
  - Name with link to Profile: _____

  Estimate SSP Evaluators' Time Effort: (e.g. 2-3 hours and/or 2 weeks to complete task): _____

- **Tool Evaluators: Pilot the DV Tool (System Security Background):** Individuals will install and run the tool.

  Task: Install and run the data visitation tool with sample data.

  Requirements: These individuals must be IT proficient.

  - Name with link to Profile: _____
  - Name with link to Profile: _____
  - Name with link to Profile: _____

  Estimate ToolEvaluators' Time Effort: (e.g. 2-3 hours and/or 2 weeks to complete task): _____

- **Internal Liaison / Assessment Team Project Manager:** a designated representative from the Organization with a Data Visitation Technology to be tested.

  Task: Act as Assessment Team liaison, providing access to necessary documentation and facilitating project communication between the assessment team and internal stakeholders.
  Requirements: Background in project management or equivalent skillset

  - Name with link to Profile: _____

  Estimate Liaison's Time Effort: (e.g. 2-3 hours weekly): _____

2. Scope of Assessment and Approval Criteria

Data Visitation Code Review

A focused review of the platform's data visitation source code by qualified individuals should be undertaken to determine if the code meets security approval criteria. The below assessment elements are recommended.

- **Sensitive Data Exposure Analysis:**
  - **Focus:** Identify code segments that transmit, process, or store sensitive information (e.g., personally identifiable information, financial data, authentication credentials).
  - **Manual Review:** Conduct a thorough manual review of code related to:
    - Data input and output.
    - Data storage and retrieval.
    - API interactions.
    - AJAX like calls.
  - **Automated Tools:** (Time permitting) Utilize Static Application Security Testing (SAST)[3] tools to automate code analysis for potential sensitive data leaks.
- **Injection Vulnerability Assessment:**
  - **Focus:** Identify and assess the risk of common injection attacks (e.g., SQL injection, cross-site scripting) that could lead to data destruction or unauthorized access.
  - **Manual Review:**
    - Examine AJAX-like calls and database interactions for potential injection vulnerabilities.
    - Review input validation and sanitization routines.
  - **Automated Tools:**
    - Dynamic Application Security Testing (DAST)[4] tools to simulate real-world attacks.
    - Interactive Application Security Testing (IAST)[1] tools for runtime analysis.
- **Authentication and Authorization Security:**
  - **Focus:** Evaluate the security of user authentication and authorization mechanisms.
  - **Manual Review:**
    - Inspect password storage and hashing practices (e.g., use of salted hashes).
    - Analyze API call token generation and validation.
    - Review user role management and elevation of privilege vulnerabilities (e.g., plain text calls to rights database).
  - **Automated Tools:** (Time permitting) Utilize penetration testing tools to simulate authentication and authorization attacks.

---

[3] SAST, IAST: https://www.semanticscholar.org/paper/Interactive-Application-Security-Testing-Pan/4d3065d450a12b028b38a2e65fd5fde35537bdc8

[4] DAST: https://ieeexplore.ieee.org/abstract/document/10543484

- **Dependency Vulnerability Assessment:**
    - **Focus:** Identify and assess known vulnerabilities in third-party libraries and packages.
    - **Automated Tools:**
        - Software Composition Analysis (SCA)[5] tools to analyze dependency versions and identify known vulnerabilities.
        - Ensure that all dependency library/package versions are up to date.
- **Security Monitoring and Alerting:**
    - **Focus:** Evaluate the effectiveness of existing security monitoring and alerting systems.
    - **Manual Review:**
        - Review security logs and monitoring dashboards.
        - Assess the timeliness and accuracy of security alerts.
        - Confirm the presence of alerting systems, and monitoring of the systems.
-
- **Consult NIST Community Profiles and CSWP (cybersecurity white papers)[6] for disciplinary and community specific guidance**
    - Community Profiles provide a way for communities, (i.e., group of organizations that share a common context and an interest in their cybersecurity posture) to describe a consensus point of view about cybersecurity risk management. The NCCoE provides examples of Community Profiles and other resources to help communities understand and develop Community Profiles, e.g. Cybersecurity Framework Profile 3 for Genomic Data (NIST IR 8467 ipd)[7]
    - Disciplinary-specific cybersecurity white papers (CSWP) offer nuanced guidance e.g. “Cybersecurity Threat Modeling the Genomic Data Sequencing Workflow” (CSWP) 35. This Draft NIST Cybersecurity White Paper evaluates potential threats in a genomic data processing environment using an iterative methodology. It provides an example use case and demonstrates an approach that organizations can adapt to identify cybersecurity threats and mitigations in their environments.
- **Restricted Data Access for AI Models:**
    - **Principle:** Assess controls that prevent AI models from accessing or transferring extraneous, secondary data beyond the explicitly authorized scope
    - **Action:**
        - Inspect data access policies that limit AI models to only the necessary data for their intended function.
        - Develop automated testing tools to monitor and audit data streams accessed by AI models, ensuring they adhere to defined access controls.
        - Verify input and output validation.
        - Verify data sandboxing.
- **AI Model Security Assessment & Data Leakage Prevention:**

[5] https://arxiv.org/pdf/1909.00973
[6] https://csrc.nist.gov/publications/cswp
[7] https://doi.org/10.6028/NIST.IR.8467.2pd

    - **Principle:** Conduct thorough security assessments of AI models, particularly those interacting with external APIs (e.g., OpenAI), to identify and mitigate potential vulnerabilities
    - **Action:**
        - Implement rigorous data leakage testing to verify that sensitive data is not being transmitted or stored by the AI model.
        - Assess the model's resilience against adversarial attacks, including prompt injection and data poisoning.
        - Establish incident response plans for potential security breaches, including ransomware attacks.
        - Implement rate limiting to prevent abuse and reduce the risk of data exfiltration through repeated or automated queries.
- **Securing Intelligent Agents**
    - As AI systems evolve from simple assistants to fully autonomous agents, they introduce increasingly complex security risks that data visitation platforms using such agents must actively address. Each level of autonomy, from observing and acting under human guidance to making independent decisions, creates new vulnerabilities that traditional incident response strategies are not equipped to manage. Given the significant computational demands of AI agents, data visitation platforms are generally expected to rely on no more than a single agent. However, even one agent, when granted decision-making authority and tool access, presents a delicate balance between utility and risk.
    - The Coalition for Secure AI (CoSAI)'s AI Incident Response Framework[8] is addressing these unique, dynamic risks posed by intelligent, autonomous systems with support from major technology companies, including Google, which recently donated data from its secure AI framework (SAIF)[9].
    - Key assessment procedures recommended by CoSAI include:
        - Agent Behavior Profiling: Establish baselines for normal agent behavior to detect anomalies
        - Tool Invocation Auditing: Log and review all tool calls made by agents for misuse or escalation
        - Prompt and Output Screening: Analyze inputs and outputs for signs of injection, hallucination, or leakage
        - Cross-Team Incident Playbooks: Develop shared protocols across security, ML, and product teams for coordinated response
        - Red Teaming and Simulation: Regularly test agent resilience through adversarial scenarios and stress testing
        - Post-Incident Forensics: Conduct root cause analysis and update guardrails and policies based on findings

---

8 https://github.com/cosai-oasis/ws2-defenders/blob/main/incident-response/AI%20Incident%20Response.md

9 https://www.oasis-open.org/2025/09/16/google-donates-secure-ai-framework-saif-data-to-coalition-for-secure-ai/

        - Continuous Risk Mapping: Use dynamic threat models that evolve with agent capabilities and deployment contexts
- **Data Corruption Prevention & Model Versioning:**
    - **Principle:** Evaluate measures used to prevent data corruption and ensure data integrity, particularly for time-sensitive or highly sensitive data (e.g., genomic data).
    - **Action:**
        - Verify that robust data integrity checks have been implemented, including checksums and data validation routines, through code inspection and review.
        - Utilize "Save Our State" (SOS) tests to identify and prevent data corruption.
        - Verify that strict model versioning and locking mechanisms have been implemented, to prevent unauthorized modifications or data drift.
        - Verify that data provenance tracking has been implemented, and verify that tracking is accurate
- **Temporality and Time Drift Management:**
    - **Principle:** Address the challenges of data temporality and time drift, which can impact the accuracy and reliability of AI models.
    - **Action:**
        - Verify that timestamping and versioning mechanisms for all data have been implemented.
        - Assess data drift detection and mitigation strategies.
        - Assess time-sensitive data validation techniques.

Data Visitation System Security Plan (SSP) Review

This Assessment plan template depends on attention to the relevant NIH required **NIST SP 800-171** security controls implemented in the data visiting system(s) being tested.

Relevant sections of the SSP should be evaluated for approval based on their completeness, accuracy, and adherence to standards specific to data visitation security. The SSP should address the protection of **Controlled Unclassified Information (CUI)** as defined by NIST SP 800-171, where applicable, considering the limited scope of data transmission.

Applicable Controls: For data visitation scenarios, NIST SP 800-171 rules concerning access control, authentication, and authorization are highly applicable as they govern who can access the data visitation tool and the underlying data. Audit and accountability rules are also relevant for tracking actions performed during data visitation sessions. Rules regarding system and information integrity, including protection against malware, are crucial to ensure the data visiting tool doesn't compromise the host system or the data itself.

Not applicable Controls: Conversely, rules around physical protection, media protection (pertaining to physical media), and arguably hardware maintenance have limited direct applicability to the data visitation process itself, especially when the data visitation tool is a temporary software layer installed on hardware not owned or maintained by the data visitor; the security of the underlying hardware remains the responsibility of the data owner. However, the

data owner's adherence to all NIST SP 800-171 controls is still paramount for the overall security of the CUI being visited.

For further information, consult Appendix 1: Applicability of NIST SP 800-171 controls to data visitation systems.

Data Visitation Technology Review

The core technologies enabling data visitation can be tested with sample data to assess security robustness and compliance with relevant standards using the below Assessment Guidelines and Standards, and following the recommended Assessment Procedures.

3. Assessment Guidelines & Standards

- **NIH Guidance for NIH Controlled-Access Data:** NIH uses the **NIST SP 800-171** guidance. The assessment should verify compliance with the controls outlined in NIST SP 800-171, particularly those related to data access and security, as a key requirement for data visitation feature approval.
- **Data Visitation Industry Best Practices:** The assessment should consider industry best practices for secure data access and visitation as benchmarks for approval.
- **ISO/IEC standard regarding Cloud and PII:** Information technology – Security techniques – Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors (ISO/IEC 27018:2019)
- **The Coalition for Secure AI (CoSAI)** proposes a new, AI-specific incident response framework that emphasizes continuous monitoring, cross-functional collaboration, and rapid containment strategies tailored to the fluid nature of AI agent threats.

4. Assessment Procedures

- **Data Visitation Code Review:** Qualified computer scientists will conduct focused code analysis of the data visitation feature to inform the approval decision.
- **Data Visitation Policy and Procedure Review :** The relevant sections of the SSP and related documentation will be reviewed for data visitation approval compliance.
- **Data Visitation Data Flow Analysis :** Data flow analysis specific to the data visitation feature will inform the approval decision.
- **Data Visitation Security Configuration Review:** Security configuration reviews of the data visitation infrastructure will inform the approval decision.

5. Deliverables & Reporting (Data Visitation Approval Documentation)

For each list item below assign a responsible party and record when the report/activity was completed and how to access it:

Activity	Responsible Party	Completion Date	File location
Data Visitation Feature Approval Decision Report: A report documenting the approval decision specifically for the data visitation feature and the rationale behind it.			
Data Visitation Remediation Recommendations: Actionable recommendations for security enhancements to the data visitation feature, if any, provided separately from the approval decision.			
Data Visitation Feature Approval Decision Report: A report documenting the approval decision specifically for the data visitation feature and the rationale behind it.			

6. Timeline

For each activity/milestone below set target dates and record completion dates

Activity	Responsible Party	Completion Date	File location
Sign NDA for code review			

Finalize Evaluation Plan			
Code Review in parallel with System Security and Tool installation and testing			
Write and publish report(s)*			

*see above section #5 on Deliverables & Reporting for recommended report(s)

7. User Access and Data Flow Diagram of the System to be Evaluated

- For every system to be evaluated, customize a system security flowchart with descriptions of the steps in the process.
- Appendix 3 Example Flow Diagram: User Access and Data Flow shows an example flowchart for the FAIRlyz platform that can be re-used & customized

8. Post-Assessment Findings and Recommendations

Use this section to summarize and record findings and recommendations.

Appendices

Appendix 1: Applicability of NIST SP 800-171 controls to data visitation systems

This section outlines the applicability of NIST SP 800-171 controls to data visitation systems:

Applicable NIST SP 800-171 Controls to Data Visitation Systems:

- Access Control (AC): Controls governing who is authorized to access the data visitation tool, the data being visited, and the actions they can perform (e.g., read-only, analysis). This includes establishing access requirements, controlling internal and remote access, and limiting data access based on the principle of least privilege.
- Identification and Authentication (IA): Controls ensuring that users accessing the data visitation system are uniquely identified and their identities are verified (e.g., using strong passwords, multi-factor authentication) before access to CUI (which includes PHI under certain conditions) is granted.
- Audit and Accountability (AU): Controls for creating and maintaining audit logs of system and data access, user activity within the data visitation session, and any modifications made. This helps in tracking actions and identifying potential security incidents.
- System and Information Integrity (SI): Controls focused on protecting the system and information from unauthorized modification, damage, or destruction. This includes measures to detect and prevent malware on the data visitation tool and the environment it operates within (to the extent controllable by the visitation process).
- Awareness and Training (AT): Ensuring that users of the data visitation tool are aware of security risks and their responsibilities in protecting the data being visited.
- Incident Response (IR): Establishing procedures to handle security incidents that may occur during or related to data visitation activities.
- Risk Assessment (RA): Identifying and assessing potential risks associated with the data visitation process.
- Security Assessment (CA): Periodically assessing the security controls implemented for the data visitation system.
- System and Communications Protection (SC): Controls related to securing the communication channels used during the data visitation session to protect the confidentiality and integrity of the data in transit.
- Planning (PL): Developing and maintaining security plans that outline how CUI (including PHI, if applicable) will be protected during data visitation.
- Configuration Management (CM): Establishing and maintaining baseline configurations for the data visitation tool and the environment it operates in (to the extent controllable), and managing changes to these configurations.

NIST SP 800-171 Controls Less Directly Applicable to the Data Visitation *Process* Itself (with caveats):

- Maintenance (MA): While the *data visitor* might not be responsible for the physical hardware maintenance of the system hosting the data, they are responsible for ensuring the security of the *data visitation tool* they are using. This could involve ensuring the tool

is up-to-date and free of vulnerabilities. The data owner remains responsible for the underlying hardware maintenance.

- Media Protection (MP): This control focuses on the protection of physical and digital media. While the data visitor doesn't typically handle the storage media of the data owner, they are responsible for protecting any temporary media they might use during the visitation (if any) and ensuring no unauthorized data is retained after the session.
- Physical Protection (PE): These controls relate to securing the physical environment where the systems processing CUI reside. The data visitor has limited control over the physical security of the data owner's premises. However, the data visitor is responsible for the physical security of their own devices used for data visitation.
- Personnel Security (PS): While the data owner is responsible for the background checks and trustworthiness of their personnel, the data visitor's organization should also have their own personnel security measures in place for individuals performing data visitation.
- System and Services Acquisition (SA): This control relates to the security considerations during the acquisition of systems and services. It's primarily the responsibility of the data owner when acquiring their infrastructure. However, the selection of a secure data visitation tool would fall under this domain for the data visitor's organization.

Appendix 2: Example Recommendations

The security evaluation recommendations focus on clarity, actionability, and a more automated cybersecurity (Software Audit Operationalization).

Continuous Software Composition Analysis (SCA) & Vulnerability Management:

- **Action:** Integrate SCA tools into GitOps/CI pipelines to perform dependency mapping, vulnerability scanning, and license compliance checks after every GitHub commit.
- **Rationale:** Proactively identify and mitigate known security vulnerabilities and license risks in third-party dependencies.
- **Example Tools:**
    - bomber[10] (GitHub Action): For comprehensive Bill of Materials (BOM) analysis.
    - npm audit[11] (Node.js): Continue to utilize for Node.js dependency vulnerability scanning.
    - snyk[12] or OWASP Dependency-Check[13]: Consider these for broader dependency scanning capabilities across multiple languages and ecosystems.

Container Image Security Hardening & Scanning:

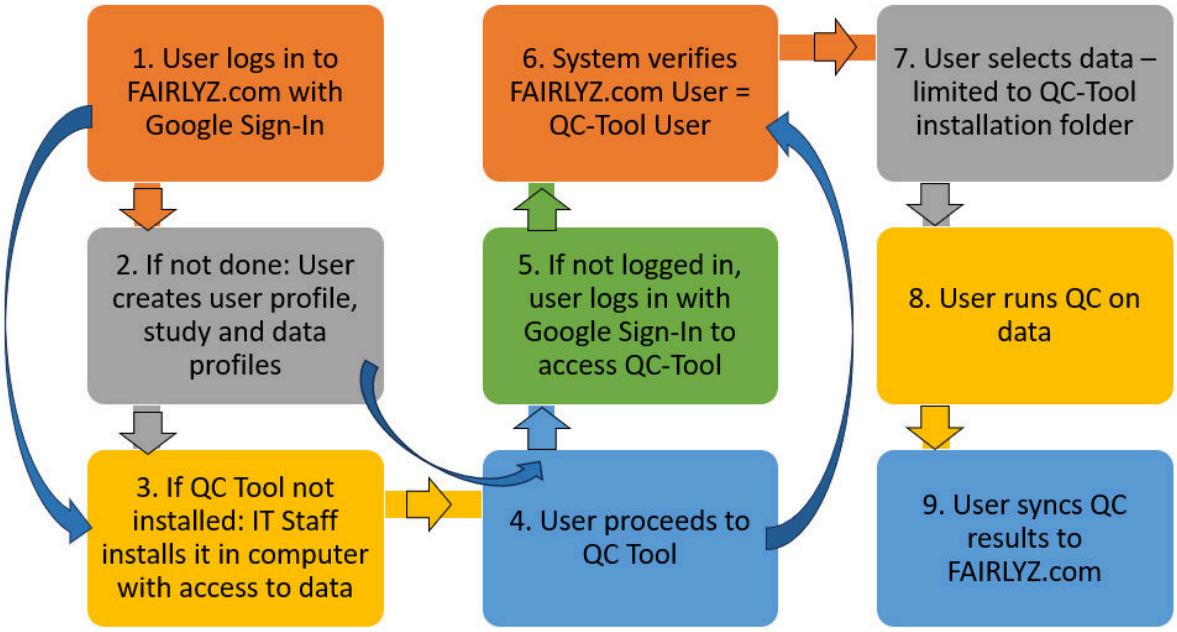
- **Action:** Implement automated security scanning of Docker container images within the CI/CD pipeline and during runtime. Enforce image hardening best practices.
- **Rationale:** Prevent deployment of vulnerable container images, ensure runtime security.
- **Example Tools:**
    - Trivy[14] or Anchore Grype[15]: For scanning container images for vulnerabilities.
    - Docker Bench for Security[16]: To assess and enforce container hardening configurations.

Automated Stale Issue/Ticket Management:

- **Action:** Implement automated actions for managing stale issues and tickets in GitHub or your issue tracking system.
- **Rationale:** Improve issue/ticket hygiene and ensure timely resolution.
- **Implementation:**
    - Utilize GitHub Actions or similar CI/CD tools to automate the process.
    - Define clear criteria for identifying stale issues/tickets (e.g., inactivity for a specified period).
    - Automate actions like adding labels, sending reminders, or closing stale issues/tickets.

[10] https://github.com/devops-kung-fu/bomber
[11] https://docs.npmjs.com/cli/v11/commands/npm-audit
[12] https://snyk.io/platform/snyk-cli/
[13] https://owasp.org/www-project-dependency-check/
[14] https://trivy.dev/
[15] https://github.com/anchore/grype
[16] https://github.com/docker/docker-bench-security

Appendix 3: Example Flow Diagram: User Access and Data Flow



User Interaction with The System, Data Access, And System Security Controls

Step 1. A user logs in using Google Sign-In and starts a session in FAIRLYZ.com

Step 2. If a user profile was not created, the user creates one. If a study was not registered, the user enters metadata for a study. If data for the study was not registered the user enters metadata for the data.

a. If the system has found an investigator ID for the user, the Investigator ID is stored in the session storage

b. Note: Login is not required if already logged in. The session expires after 24 hours.

Step 3. The QC Tool is installed in the user's computer/compute environment with access to the data.

Step 4. The user navigates to the QC Tool from FAIRLYZ.com by opening another tab with the localhost or the IP URL from the machine where the QC Tool is installed.

a. The study ID and data profile ID are provided with the URL.

b. The system stores this in local storage. Study ID and data profile ID are publicly known IDs that do not need to be protected

c. Investigator ID needs to be protected and not stored in local storage.

Step 5. The user login to QC-App requires another Google Sign-In

a. The system finds an Investigator ID for the user. The Investigator ID is stored in the session storage of the localhost or IP-Address of the QC-App.

b. Note: Login is required for the QC-App IP address or localhost separately from FAIRLYZ.com
c. Note: login is not required if already logged in. The session expires after 24 hours.

Step 6. The system verifies that the user who logs in is the same user who owns the study and data profile in FAIRLYZ.com
a. If verified, the system proceeds with data visitation and QC
b. If not verified, the access is blocked

Step 7. Data access is restricted to the Docker installation folder.

Step 8. The user runs QC on the data

Step 9. The user syncs QC results to FAIRLYZ.com

Please cite this document as: Patricia Buendia, Patrick Shironoshita, Lars Eklund, David Molik and Natalie Meyers. “System Security Assessment Plan Template.” DV4RDA Project of the AIDV-WG. Research Data Alliance. November 17, 2025.