FAI

DATA QUALITY MONITORING STANDARD OPERATING PROCEDURE (SOP)

AUGUST, 2024

Data Quality Monitoring Standard Operating Procedure (SOP)

INTRODUCTION

This Standard Operating Procedure (SOP) outlines the processes and responsibilities for monitoring and taking corrective actions to ensure data quality at the Federal Acquisition Institute (FAI) for User Profile data within FAI's instance of Cornerstone on Demand (FAI CSOD). These procedure will ensure FAI's in compliance with federal data quality policies (e.g., the Information Quality Act[1] and the Foundations for Evidence-Based Policymaking Act of 2018[2]), which emphasize data quality, relevance, accuracy, and integrity in federal data assets.

Data quality is crucial for data-driven decision-making and maintaining the credibility of data products. High-quality data enables effective decision-making, supports operational efficiency, and builds public trust. This SOP provides detailed guidance for FAI staff and stakeholders (e.g., Acquisition Career Managers [ACMs] and Small Agency Representatives [SARs]) to understand, monitor, and improve data quality. The Evidence Act contains several provisions to improve the quality of federal data assets. One provision requires agencies to "develop and maintain a strategic information resources management plan that . . . implements a process to evaluate and improve the timeliness, completeness, consistency, accuracy, usefulness, and availability of open Government data assets" (Pub. L. No. 115-435, Title II, Sec. 202(c), 2018, 4174-8).

FAI manages the data quality of the FAI CSOD system by categorizing the data into two primary types: system-managed transactional records and user-initiated inputs.

System-managed transactional records are data created and managed by the FAI CSOD system based on internal processes and rules. The FAI CSOD system creates this data during normal operations, such as tracking login histories, course attendance records, transcripts, and other system interactions. FAI relies on internal system rules to maintain the quality of these records, including automated checks and validations performed by the FAI CSOD system. FAI assumes that all system-managed transactional records meet fitness-for-purpose requirements, meaning that the data is suitable for its intended use and adheres to internal system data processing rules.

User-initiated inputs are data entries made directly by users based on their knowledge outside the control of the FAI CSOD system. User-initiated inputs include user data entries, often involving personal information, qualifications, and other user data. Users enter and manage this data in the FAI CSOD user profile. FAI monitors data quality for

---

[1] [insert reference]
[2] Pub. L. No. 115-435, 132 Stat. 5529, 2018

these types of inputs through data standards and Key Performance Indicators (KPIs) (i.e., Data Targets and Data Thresholds). This SOP focuses on monitoring the data quality of user-initiated inputs and implementing corrective actions to ensure the accuracy of the data within the FAI CSOD system.

DEFINITIONS

Data quality refers to the usefulness and credibility of data and its products, such as reports, analyses, and visualizations. High-quality data accurately represents reality and sustains public trust. FAI evaluates data quality using three main components:

1. **Utility:** The extent to which information is well-targeted to identified and anticipated needs, reflecting its usefulness to intended users.
2. **Objectivity:** The accuracy, reliability, and unbiased presentation of information
3. **Integrity:** Maintaining rigorous quality standards and protecting information from manipulation, influence, or unauthorized access

FAI data standards set data element definitions, define acceptable values and prescribe data collection requirements. In addition, FAI uses data standards to determine the data quality requirements at agency or government-wide levels. FAI monitors data quality for most CSOD data elements on a government-wide scale, with two exceptions. Individual agencies manage the data quality of the Acquisition Workforce (AWF) and Contracting Officer Representative (COR) affiliation data elements. According to Policy Letter 05-01, individual agencies define AWF and COR affiliation data elements, which can vary from one agency to another. FAI has no government-wide standards for determining data quality targets for these two data elements. For AWF and COR affiliation data elements, FAI offers KPIs for monitoring data quality. Still, FAI will not set AWF and COR affiliation quality thresholds on a government-wide scale, leaving the determination of quality targets and thresholds to individual agencies.
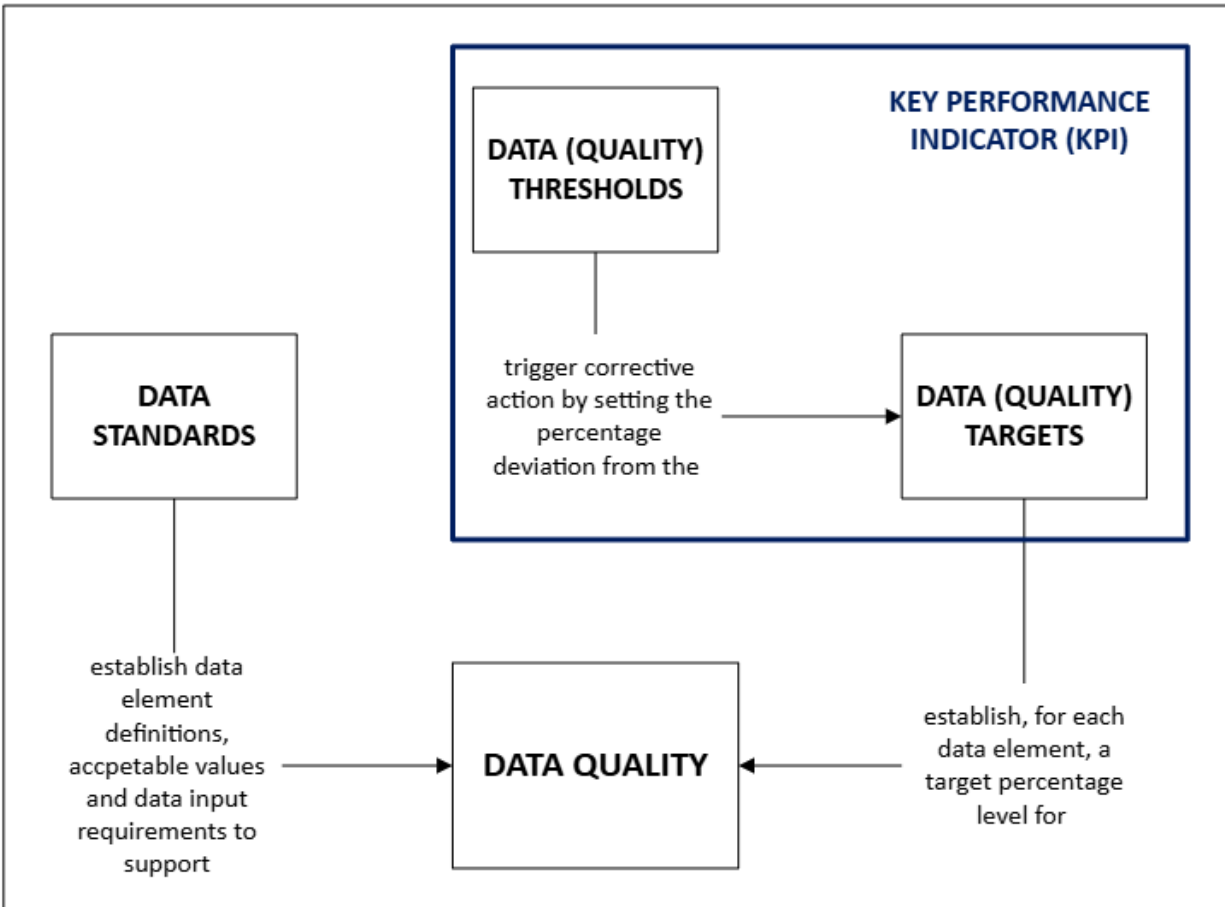
Figure 1. How Data Standards, Data Targets, and Data Thresholds relate to Data Quality

ROLES AND RESPONSIBILITIES

Maintaining data quality within the FAI CSOD system is a dual responsibility shared between FAI and the ACMs/SARs, each playing a crucial role at different levels. FAI oversees data quality acquisition workforce-wide, establishing and enforcing standards that ensure consistency and accuracy across all agencies. ACMs and SARs are responsible for the upkeep of data quality at the agency level. They are responsible for educating and directing agency AWF members to apply FAI's data standards, ensuring the FAI CSOD data meets defined quality requirements. This collaborative approach ensures that data quality is maintained broadly across the workforce and within each agency.

FAI acts as the FAI CSOD data custodian focused on evaluating data accuracy (which is a dimension of data quality) since data accuracy can be defined and measured. FAI defines accuracy expectations to ensure data products and services are free of deficiencies that may interfere with their use. FAI operationalizes fitness for use as

meeting adequately defined accuracy specifications. In addition, FAI is concerned with meeting or exceeding customer expectations. FAI assumes that more than simply conforming to accuracy specifications is needed, and data products or services must meet or exceed customer expectations by implementing data quality corrective actions when required.

Agencies own the data they create and manage. FAI cannot single-handedly improve the quality of agency data. ACMs/SARs are responsible for determining data's fitness for purpose, applying FAI's data standards and data quality thresholds, and maintaining an acceptable level of data quality. ACMs/SARs are accountable for instructing agency AFW members to follow FAI data standards requirements and monitoring the agency data quality levels based on the available KPIs. Refer to Appendix A for the definition of fitness for purpose.

FAI Data Program Support Staff

- **Responsibilities:**
  - Establish data standards and oversee the accuracy of learning management data within FAI CSOD.
  - Establish data quality thresholds for each data standard and conduct regular reviews of these KPIs.
  - Identify and communicate data quality issues to relevant stakeholders.
  - Support stakeholders implementing corrective actions.

ACMs and SARs:

- **Responsibilities:**
  - Conduct regular reviews of agency-specific data quality KPIs.
  - Execute agency data quality assessments and corrective actions.
  - Ensure data meets fitness for purpose requirements.
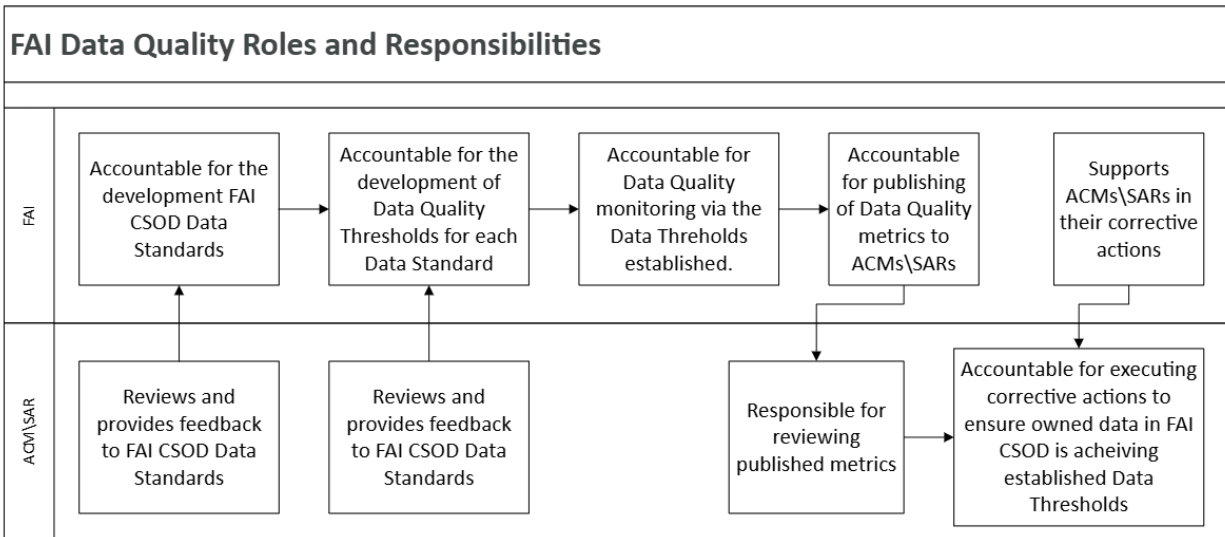 - Educating agency workforce on data quality expectations.

FAI Data Quality Roles and Responsibilities				
FAI Accountable for the development FAI CSOD Data Standards	Accountable for the development of Data Quality Thresholds for each Data Standard	Accountable for Data Quality monitoring via the Data Threholds established.	Accountable for publishing of Data Quality metrics to ACMs\SARs	Supports ACMs\SARs in their corrective actions
ACM\SAR Reviews and provides feedback to FAI CSOD Data Standards	Reviews and provides feedback to FAI CSOD Data Standards		Responsible for reviewing published metrics	Accountable for executing corrective actions to ensure owned data in FAI CSOD is acheiving established Data Thresholds

Figure 2. Process Breakdown by Role

DATA QUALITY MONITORING PROCESS

Data quality monitoring focuses on applying the data quality standards at the point of information collection, displaying KPIs, and examining the approaches for collecting data to determine whether they are likely to produce high-quality data over time. The commitment of agencies to following data standard requirements is integral to the production of good-quality data.

Data quality monitoring involves ongoing, systematic evaluation of KPIs to determine if data meets stakeholder requirements. The process answers two key questions:

1. Can a decision based on the data be made with the desired level of certainty?
2. Would the data support the same analysis with similar certainty if used again?

The initial set of KPIs (see Appendix A) focuses on presenting data quality assessments of essential FAI CSOD user profile data elements. FAI plans to expand the KPI inventory to include additional data quality assessments. FAI assumes the ownership for maintaining the data quality monitoring process (via the D2D Portal) for FAI CSOD data elements used by all agencies and making the following decisions:

- Is the data sufficiently complete and accurate to support business process needs (i.e., "fit for purpose");
- What is the desired state of specific attributes (i.e., "targets");
- What is the minimum level of quality acceptable (i.e., "thresholds");
- What are the best measures and metrics to track improvements?

Continuous Monitoring

FAI monitors FAI CSOD data to meet the required data quality thresholds. Continuous tracking helps identify and address data quality issues promptly. FAI uses the D2D Business Intelligence (BI) platform to host data quality KPIs and set event-driven notifications to signal if a KPI indicates that data quality is falling below minimum data threshold values at an agency level. Once FAI receives a message indicating that data quality drops below the expected data threshold on one or more KPIs, FAI will notify the agency about data discrepancies and request corrective actions.

Corrective Actions

ACMs/SARs are responsible for understanding how KPIs indicate increases or reductions in FAI CSOD data quality and applying proactive and reactive measures to preserve agency FAI CSOD data quality at or higher levels prescribed by data quality thresholds. ACMs/SARs should implement corrective actions aimed at improving data quality under the following circumstances:

- KPI data deviates from the target and is below the threshold, indicating poor data quality.
- Stakeholders report potential issues with data quality.

Most data quality issues result from past or present incorrect application of data standards when entering data into FAI CSOD. For example, a user might enter the wrong email address or forget to update organizational affiliation data after moving to a different agency. To prevent future data quality issues, agencies can take proactive steps to ensure that users are familiar with FAI CSOD data standards and understand data entry requirements. To mitigate existing data quality issues, agencies can either investigate data quality issues by downloading FAI CSOD reports and analyzing the data discrepancies or by working with FAI data specialists to develop a plan for fixing data issues. Typically, isolated data issues should be identified and handled by ACMs/SARs. FAI can assist with large-volume data quality issues that require bulk updates.

FAI can support ACMs/SARs in developing an approach for correcting data issues on a case-by-case basis. FAI can also help agencies develop 'lessons learned' assessments and recommendations for implementing future mitigation actions to address potential data quality concerns.

APPENDIX A: RESOURCES AND DEFINITIONS

Data Quality Targets and Thresholds:

This table outlines the minimum data targets and thresholds for FAI CSOD Data Elements. Each may be set higher as warranted on respective data elements as determined by FAI to achieve the objectives of this SOP.

Mandatory/Optional	Freeform/System constrained	Data (Quality) Target	Data (Quality) Threshold
Optional	Freeform	80%	75%
Optional	System constrained	90%	85%
Mandatory	Freeform	95%	90%
Mandatory	System constrained	98%	93%

List of Data Quality KPIs:

FAI CSOD Data Element	Data Target	Data Threshold
User Record: Work Email	95%	90%
User Record: Status	98%	93%
User Record: Job Series	98%	93%
User Record: Citizen Type	98%	93%
User Record: Grade	98%	93%
User Record: Organization	98%	93%
User Record: Organization Designation	98%	93%
User Record: City	95%	90%
User Record: State	95%	90%
User Record: ZIP	95%	90%
User Record: Training Manager	98%	93%

User Record: Training Manager Email	95%	90%

Why FAI Monitors Data Quality

<u>Data quality refers to how data accurately represents real-world values or events.</u> Ensuring quality involves identifying and correcting errors in CSOD data, such as incorrect entries or misrepresentations. FAI improves quality by implementing data standards, which help prevent inaccurate information from entering CSOD.

Monitoring data quality is essential because accurate data ensures that decisions made based on the data are reliable and valid while reducing inefficiencies and preventing loss of credibility. At FAI, the accuracy of user profile data within the FAI CSOD system is critical because:

1. **Decision-Making**: Quality data support sound decisions regarding learning management and resource allocation.
2. **Operational Efficiency**: Quality data streamline operations and reduce errors.
3. **Public Trust**: Maintaining data quality builds trust with stakeholders and the public, demonstrating FAI's commitment to high standards.

Federal guidance for implementing the Information Quality Act (Pub. L. No. 106-554, § 515(a), 2000) emphasizes the concept of "fitness for purpose" as a touchstone for evaluating and communicating the quality of data made available to the public. FAI treats "fitness for purpose" or "fitness for use" as a broad definition of data quality. FAI's data quality assessment evaluates how the data meets its intended purpose in FAI's operations, decision-making, and planning. In this view, we assess data quality within the context of the intended uses of the data. Fitness for use is the extent to which data successfully serves the purposes of consumers. While fitness for use captures the essence of quality, measuring quality using this broad definition is difficult.