

# Project Communication Report

## Data Quality Paper

February 5, 2016

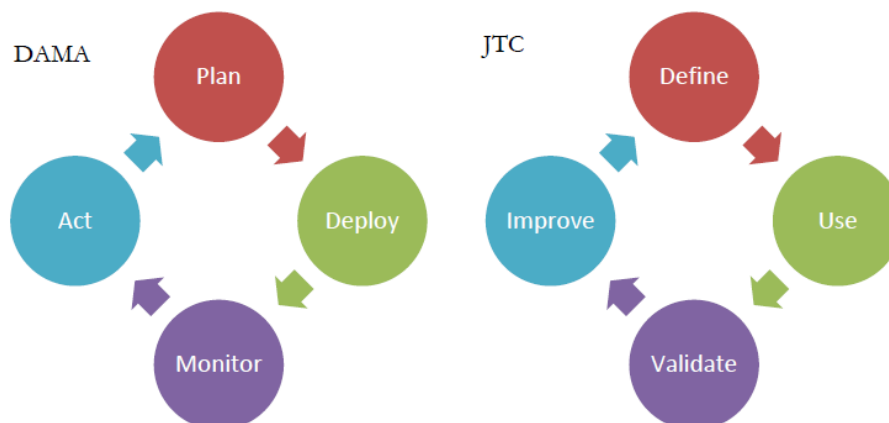
Data Quality Written Report

I have exchanged emails with the primary drafter of the paper, Richard Schauffler, and he has informed me that the timeline for an initial draft has been moved to the Summer JTC Meeting.

An outline was defined for the paper in September of 2015 which stemmed from a refocused approach for the workgroup. During the July 11<sup>th</sup> meeting in Louisville, KY, the JTC reviewed survey results regarding specific areas for “Data Management” and the topic was narrowed to “Data Quality”. The workgroup met several times and the following items were completed:

1. DropBox Folder was created to compile relevant papers and materials for review.
2. Various sources and papers were considered for content.
3. Coordinated with IJIS to incorporate their paper and resources.
4. Review documents and select relevant content related to “Data Quality”.
5. Outline was produced and presented at the September 16<sup>th</sup> JTC call.
6. Data Management Association (DAMA) resources were incorporated.
7. A second call was held on September 10<sup>th</sup>
8. Judge Trickey produced an initial outline which was shared in late Sept. 2015.
9. A primary writer, Richard Schauffler, was identified to produce the first draft.
10. The October draft was pushed off to December 2015.
11. December draft was pushed off until “early 2016”.
12. Latest update is that the draft will be delivered in the “summer JTC meeting”.

The following outline was the latest artifact(s) produced by this workgroup.



## Data Quality Management Cycle

1. Identify critical data issues to achieve business objectives
2. Identify quality dimensions
3. Define business rules critical to ensuring high quality data
  - a. Define tolerance levels for data entered.

### Plan

1. Scope of known issues
2. Cost / impact / risks of those issues
3. Potential alternatives for addressing them

### Deploy

1. Profile data
2. Institute inspections / monitoring
  - a. Correct errors / flaws
  - b. Ideally, earliest point in the data flow process

### Monitor

1. Apply business rules to evaluate data quality
2. Does data meet quality thresholds defined by rules?
3. Communicate evaluation outcomes with data stewards / owners

### Act

1. Take action to resolve emerging data quality issues
2. Ongoing, iterative process requiring management oversight
3. Data Governance – policy decision makers with authority

Data Quality: influenced by data governance

Synonymous with information quality

- Poor quality results in:
  - Inaccurate information
  - Poor business performance
- What we need our readers to know!!

Ideal Solution: identify root causes of data defects

- What we need our readers to do!!

## DAMA JTC

How??

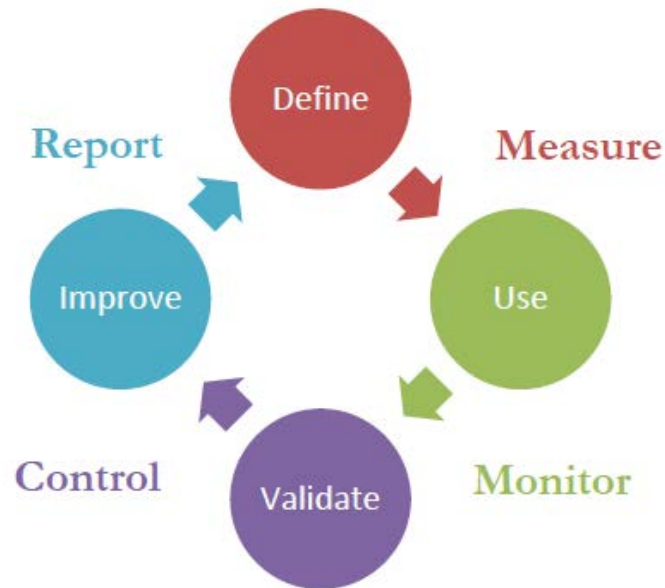
- Manage data life cycles – data meets the needs of all consumers / partners
  - Business, budget, legislative, executive, justice, corrections

1. Creation
2. Transformation
3. Transmission

Identify court needs:

Best way

1. Measure
2. Monitor
3. Control
4. Report



Process Results:

- Notification to data stewards to make corrective action

Removes:

- Root causes of data quality issues

Continuous Ongoing Management Process:

- Continuous communication
- Continuous monitoring
- Define and adjust acceptable levels of quality for business units

Questions and Answers – DAMA

Grid: - DAMA

Build – Business case – Harris County, TX ??

1. Define (Identify the problem - Business Requirement) ‘Five Why’s’

- a. Need all information on Customers in one location
- b. Inconsistent data in multiple systems
- c. Multiple data sources
- d. Multiple sources of authority
- e. Lack of trust in the data (Quality)
  - i. Inventory of data sources
  - ii. Data elements – Common
  - iii. Data definitions
  - iv. Impacted customers

2. Measures

- a. Define Business Rules
  - i. Measurability - Populated
  - ii. Business Relevance - Data Types
  - iii. Acceptability - Format
  - iv. Accountability/Stewardship/Ownership
  - v. Controllability - Accuracy
  - vi. Track ability - Flags

3. Use (Data Audit - Profiled)

Data Analysis against Business Rules - Data Dimensions

- i. Populated
  - 1. Completeness
  - 2. Required
- ii. Data Types
  - 1. Numerical
  - 2. Alpha / Text
  - 3. Date/Time
- iii. Format
  - 1. Consistency
    - a. Reasonableness
    - b. Timeliness
    - c. Uniqueness
- iv. Flags
  - 1. Privacy
  - 2. Checkboxes
  - 3. Radio
  - 4. True/False
- v. Accuracy
  - 1. Currency
  - 2. Validity
  - 3. Precision
  - 4. Referential Integrity

4. Monitor (Ongoing Considerations)

- a. Trends / Patterns

- b. Thresholds
  - c. Communication to Data Stewards / Owners
    - i. Training
    - ii. Systems
    - iii. Processes
5. Controls (Offset / Manage Patterns)
- a. Policy and Procedures
  - b. Development
  - c. Verification and Validation (IV&V)
6. Improve (Act – Ongoing iterative process)
- a. Ongoing Management Oversight
  - b. Data Governance (Decision Makers)
7. Report (Publish)
- a. Reviews
  - b. Websites