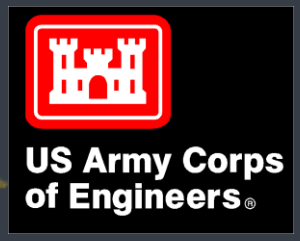
DATA MANAGEMENT BEST PRACTICES

JOELLE CARNEY & JON HALL

JAN. 17, 2023



Team of Professionals Making Tomorrow Better

WHAT IS DATA MANAGEMENT?

DATA MANAGEMENT

BUILDING STRONG

Data Management is a critical business driver used to ensure data is acquired, validated, stored, and protected in a standardized way. It is essential to develop and deploy the right processes so end users are confident their data is reliable, accessible, and up to date. To make sure that your data is managed most effectively and efficiently.

****Is data management the same as database management? What are the differences between these concepts?****

Data Management *is the standards by which you follow to collect, maintain/update, reference, secure, and archive the data.*

Database Management *is the standards by which you maintain and manage the "BOX" that houses that data.*

https://www.tableau.com/learn/articles/data-management-best-practices

DATA STANDARDS

Create standard file naming and cataloging conventions

Implement metadata/documentation such as Standard Operating Procedures for data sets (to help answer questions like why data exists and how it can be utilized)

Data Storage – Hardware (ex. Computer), network drives/infrastructure, cloud, and external drives

Commitment to data culture – Making sure there is an understanding of what data management is and the importance of implementing it

Typically, data management practices make use of a combination of processes, including:

- **Collection/Documentation**
- **Review/Record Tracking Process**
- **Access**
- **Storage**
- **Availability**
- **Security and Privacy**

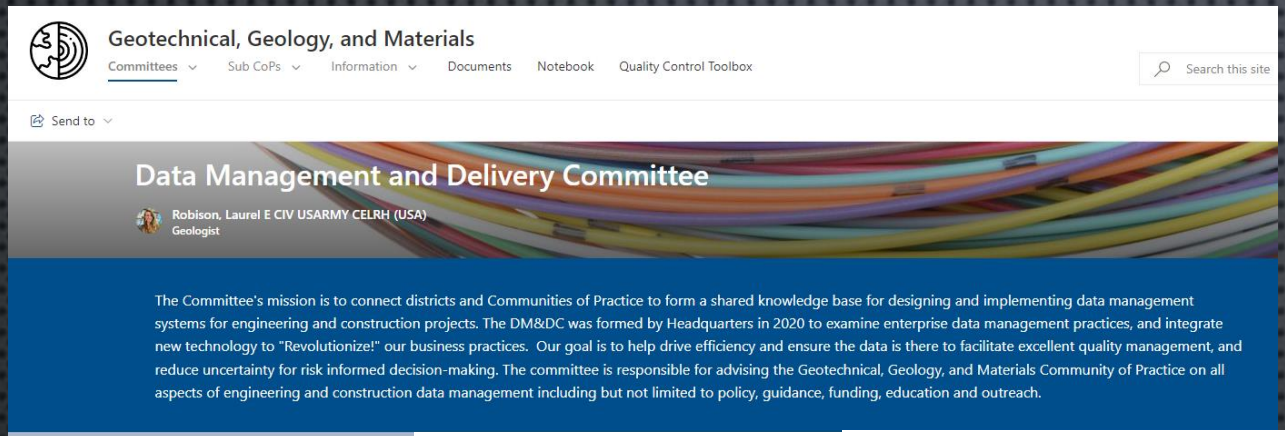
PRODUCTS

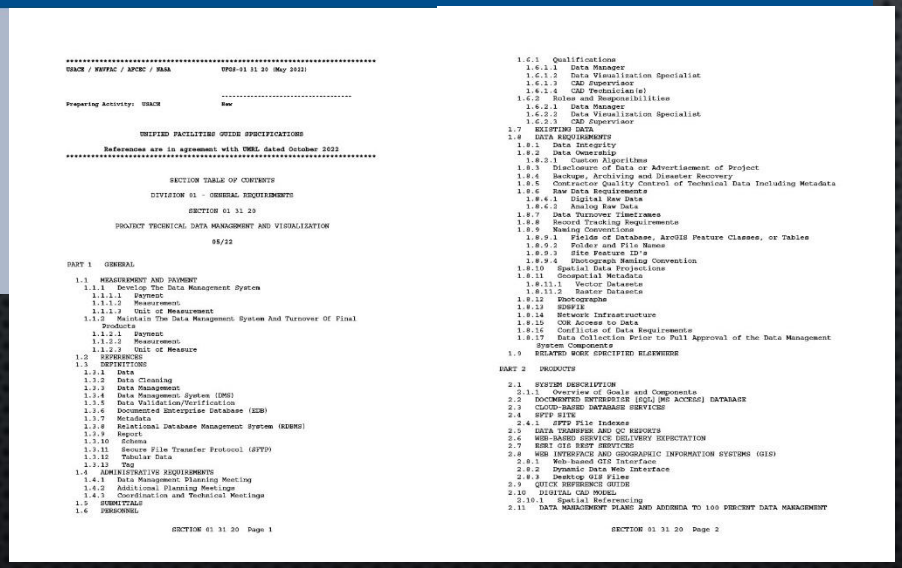
BUILDING STRONG

- Data Management Specification Guide
- Database
- Data Dictionary
- Entity-Relationship Diagram
- Online Components – Viewers
- Reports

DATA MANAGEMENT SPEC

BUILDING STRONG





USACE / NAVFAC / AFCEC / NASA UFGS-01 31 20 (May 2022)

Preparing Activity: USACE New

UNIFIED FACILITIES GUIDE SPECIFICATIONS

References are in agreement with UMRL dated October 2022

SECTION TABLE OF CONTENTS

DIVISION 01 - GENERAL REQUIREMENTS

SECTION 01 31 20

PROJECT TECHNICAL DATA MANAGEMENT AND VISUALIZATION

05/22

PART 1 GENERAL

1.1 MEASUREMENT AND PAYMENT
1.1.1 Develop The Data Management System
1.1.1.1 Payment
1.1.1.2 Measurement
1.1.1.3 Unit of Measurement
1.1.2 Maintain The Data Management System And Turnover Of Final Products
1.1.2.1 Payment
1.1.2.2 Measurement
1.1.2.3 Unit of Measure
1.2 REFERENCES
1.3 DEFINITIONS
1.3.1 Data
1.3.2 Data Cleaning
1.3.3 Data Management
1.3.4 Data Management System (DMS)
1.3.5 Data Validation/Verification
1.3.6 Documented Enterprise Database (EDB)
1.3.7 Metadata
1.3.8 Relational Database Management System (RDBMS)
1.3.9 Report
1.3.10 Schema
1.3.11 Secure File Transfer Protocol (SFTP)
1.3.12 Tabular Data
1.3.13 Tag
1.4 ADMINISTRATIVE REQUIREMENTS
1.4.1 Data Management Planning Meeting
1.4.2 Additional Planning Meetings
1.4.3 Coordination and Technical Meetings
1.5 SUBMITTALS
1.6 PERSONNEL

SECTION 01 31 20 Page 1

1.6.1 Qualifications
1.6.1.1 Data Manager
1.6.1.2 Data Visualization Specialist
1.6.1.3 CAD Supervisor
1.6.1.4 CAD Technician(s)
1.6.2 Roles and Responsibilities
1.6.2.1 Data Manager
1.6.2.2 Data Visualization Specialist
1.6.2.3 CAD Supervisor
1.7 EXISTING DATA
1.8 DATA REQUIREMENTS
1.8.1 Data Integrity
1.8.2 Data Ownership
1.8.2.1 Custom Algorithms
1.8.3 Disclosure of Data or Advertisement of Project
1.8.4 Backups, Archiving and Disaster Recovery
1.8.5 Contractor Quality Control of Technical Data Including Metadata
1.8.6 Raw Data Requirements
1.8.6.1 Digital Raw Data
1.8.6.2 Analog Raw Data
1.8.7 Data Turnover Timeframes
1.8.8 Record Tracking Requirements
1.8.9 Naming Conventions
1.8.9.1 Fields of Database, ArcGIS Feature Classes, or Tables
1.8.9.2 Folder and File Names
1.8.9.3 Site Feature ID's
1.8.9.4 Photograph Naming Convention
1.8.10 Spatial Data Projections
1.8.11 Geospatial Metadata
1.8.11.1 Vector Datasets
1.8.11.2 Raster Datasets
1.8.12 Photographs
1.8.13 SDSFIE
1.8.14 Network Infrastructure
1.8.15 COR Access to Data
1.8.16 Conflicts of Data Requirements
1.8.17 Data Collection Prior to Full Approval of the Data Management System Components
1.9 RELATED WORK SPECIFIED ELSEWHERE

PART 2 PRODUCTS

2.1 SYSTEM DESCRIPTION
2.1.1 Overview of Goals and Components
2.2 DOCUMENTED ENTERPRISE [EDB] [NO ACCESS] DATABASE
2.3 CLOUD-BASED DATABASE SERVICES
2.4 SFTP SITE
2.4.1 SFTP File Indexes
2.5 DATA TRANSFER AND QC REPORTS
2.6 WEB-BASED SERVICE DELIVERY EXPECTATION
2.7 ESRI GIS REST SERVICES
2.8 WEB INTERFACE AND GEOGRAPHIC INFORMATION SYSTEMS (GIS)
2.8.1 Web-based GIS Interface
2.8.2 Dynamic Data Web Interface
2.8.3 Desktop GIS Files
2.9 QUICK REFERENCE GUIDE
2.10 DIGITAL CAD MODEL
2.10.1 Spatial Referencing
2.11 DATA MANAGEMENT PLANS AND AGENDAS TO 100 PERCENT DATA MANAGEMENT

SECTION 01 31 20 Page 2

- NATIONAL INITIATIVE
- SECTION 01 31 20
- BROAD-BASED SPEC
- TAILOR THE SPEC TO THE PROJECT NEEDS

https://usace.dps.mil/sites/KMP-GGM/SitePages/Data-Management.aspx

Team of Professionals Making Tomorrow Better

DATA MANAGEMENT SPEC

BUILDING STRONG

Outlines:

- Scope
- Technical Definitions
- Data Integrity
- Data Ownership
- Backups/Archiving/Security
- Data Requirements
- Submittal Procedures
- Products
- References (Standards)

SCOPE

BUILDING STRONG

THE WORK COVERED BY THIS SECTION CONSISTS OF A DATA MANAGEMENT SYSTEM. THE DATA MANAGEMENT SYSTEM ENCOMPASSES EVERY ASPECT OF HOW THE DATA IS HANDLED FROM GENERATION TO TURNOVER, AND ITS CONSTITUENT COMPONENTS.

ELEMENTS OF THE DATA MANAGEMENT SYSTEM INCLUDE HARDWARE, SOFTWARE, NETWORK INFRASTRUCTURE, DATABASE ARCHITECTURE, A DATA MANAGEMENT PLAN, RAW AND PROCESSED TABULAR DATA, A WEB-BASED INTERFACE, AND FLAT FILE REPORTS AND DOCUMENTS INCIDENTAL TO ALL CONSTRUCTION OR VERIFICATION ACTIVITIES, COORDINATION, TRAINING, QUALITY CHECKING, AND ANY OTHER ASSOCIATED INFORMATION DELIVERABLES REQUIRED DURING THE CONSTRUCTION CONTRACT.

TECHNICAL DEFINITIONS

BUILDING STRONG

- Data
- Tabular Data
- Data Management
- Data Management System
- Data Cleaning
- Data Verification
- Raw Data
- Appended raw data
- GIS
- Metadata
- Report
- Tag
- Schema
- SFTP
- Documented Enterprise Database *(EDB)*

DATA INTEGRITY AND OWNERSHIP

- **Integrity:** The accuracy and consistency of the data stored within the database.

- **Ownership:** Who ultimately owns the data?

BACKUPS / ARCHIVING / SECURITY

- Ensuring that there will be an adequate procedure for data back-up, generally in a separate location from the main data repository.

- The data has sufficient digital and physical security.

DATA REQUIREMENTS & STANDARDS

- What raw data requirements are needed?
- What testing standards are needed?

  Examples of standard tests: ASTM, API, ACI, USACE, etc.

WHAT IS A DATABASE?

- Collection of structured information
- Controlled by a Relational Database Management System (RDBMS), e.g., Oracle.
- The data, RDBMS, and associated programs together are the database system

RDBMS VS ACCESS VS EXCEL

BUILDING STRONG

- Single to small user base
- Small Data Collections
- No complex data

RDBMS VS **ACCESS** VS EXCEL

BUILDING STRONG

- SINGLE TO SMALL USER BASE
- SMALL TO MODERATE DATA COLLECTIONS
- SMALL AMOUNT COMPLEX DATA

RDBMS VS ACCESS VS EXCEL

BUILDING STRONG

- Multiple Simultaneous Users
- Large Collections of Data
- Highly Complex Logic and Language

RELATIONAL DATABASES

- Based on Relational Model
- Database Objects – Tables
- Unique IDs – Keys
- Relationships
- Conceptual, Logical, and Physical Models

DATABASE OBJECTS - TABLES

BUILDING STRONG

ELEMENTS OF A RELATIONAL DATABASE TABLE:

1. NAMED VERTICAL COLUMNS
2. HORIZONTAL ROWS
3. CELL
4. PRIMARY KEY COLUMN

THINK EXCEL / SPREADSHEET!

PRIMARY KEYS

- Ensures unique values within a column
- Unique ID
- One per table
- No duplicate values
- Does **not allow** NULL values
- Cannot be deleted

FOREIGN KEYS

- Provides link between two tables
- Refers to the primary key of another table
- Multiple foreign keys per table
- Duplicate values
- NULL values allowed
- Can be deleted

RELATIONSHIPS

Relationship (Cardinality) Types:

- One-to-One (1-1)
- One-to-Many (1-M)
- Many-to-Many (M-N)

DATABASE SCHEMA

BUILDING STRONG

WHAT IS A SCHEMA?

- DATA DICTIONARY
- ENTITY RELATIONSHIP DIAGRAM *(ERD)*

DATA DICTIONARY

BUILDING STRONG

- A LISTING OF DATA OBJECTS (NAMES AND DEFINITIONS)
- DETAILED PROPERTIES OF DATA ELEMENTS (DATA TYPE, SIZE, NULLABILITY, OPTIONALITY, INDEXES)
- ENTITY-RELATIONSHIP (ER)
- REFERENCE DATA (CLASSIFICATION AND DESCRIPTIVE DOMAINS)
- BUSINESS RULES, SUCH AS FOR VALIDATION OF A SCHEMA OR DATA QUALITY

DATA DICTIONARY COMPONENTS

BUILDING STRONG

- **Spreadsheet Tab:** Each tab generally will be used for a single table.
- **Object Name:** The name of the table and associated columns (or other database object).
- **Object Description:** A brief description of the table and associated columns (or other database object).
- **Column Data Type:** Varchar2, number, timestamp

DATA DICTIONARY COMPONENTS

BUILDING STRONG

- **Character Length:** For varchar2 (4000 bytes / characters)
- **Acceptable Values:** List of all acceptable constraint values.
- **Required:** Is the column required to run calculations.
- **Accepts NULL Value:** The column allows NULL values.

DATA DICTIONARY

BUILDING STRONG

Table: CONSTRUCTION_INFO					
Table Information					
Definition:	Table containing general information about the construction project.				
Attribute Information					
Model Name	Type	Domain	Data Required	Nullable	Attribute Definition
id	number(38,0)		yes	no	The primary key of the table.
contractNumber	varchar2(50)		yes	yes	The contract number of the construction project.
projectName	varchar2(100)		yes	yes	The name of the contracted project.
usaceContact	varchar2(50)		yes	yes	The USACE contact for the construction.
division	number(38,0)	d_DIVISION	yes	yes	The division in which the construction is taking place.
district	number(38,0)	d_DISTRICT	yes	yes	The district in which the construction is taking place.
constStartDate	timestamp		yes	yes	The start date of the construction project.
constEndDate	timestamp		yes	yes	The end date of the constrcution project.
comments	varchar2(2000)		yes	yes	Any additional comments about the table.

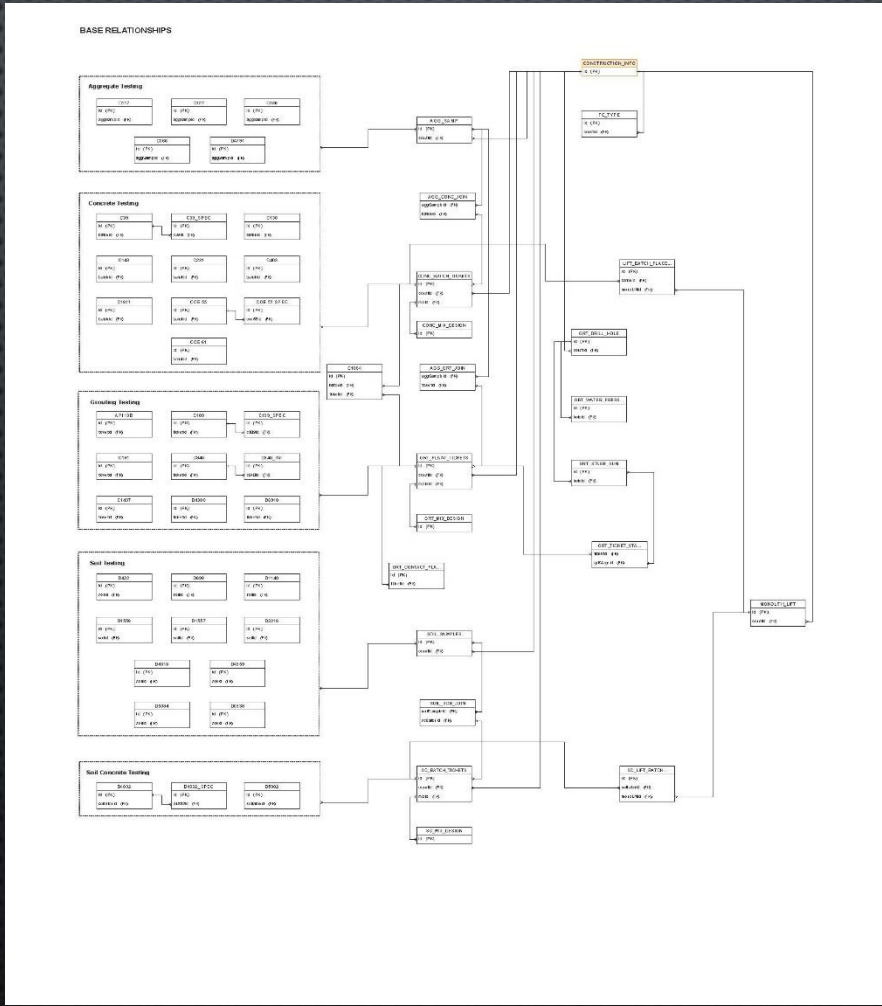
ENTITY RELATIONSHIP DIAGRAM *(ERD)*

Basic Components:

- Entities / Tables
- Keys (Primary and Foreign)
- Relationships
- Attributes / Columns

ENTITY RELATIONSHIP DIAGRAM *(ERD)*

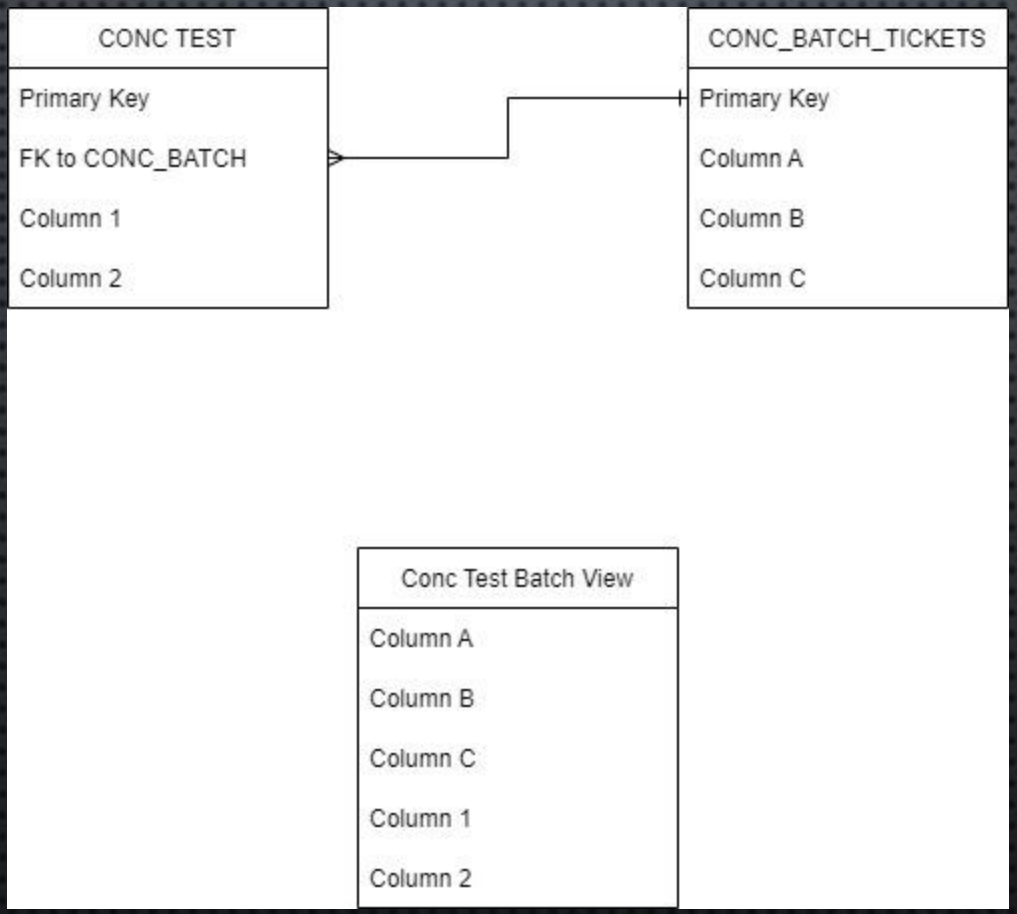
BUILDING STRONG



Team of Professionals Making Tomorrow Better

VIEWS / REPORTS

BUILDING STRONG



DURING & AFTER CONSTRUCTION

BUILDING STRONG

Import Data

Export Data

Merge Data Sets

Verify & Enrich

Rebuild Missing Data

De-Duplicate

Standardize Data

Normalize Data

USACE Data Manager

Team of Professionals Making Tomorrow Better

SUMMARY

BUILDING STRONG

- Data Management Concept
  - Integrated into other programs
- SAJ Implementation
  - C-43
  - HHD Cutoff Walls
  - C-23/24 STA
  - EAA

QUESTIONS

BUILDING STRONG

Team of Professionals Making Tomorrow Better