UMS Group Inc.
Substation Data Validation, Cleansing, De-cluttering, And New Data Creation

OVERVIEW

UMS Group was engaged to examine and repair data on substation assets in SAP:

- What information is stored in SAP for each asset type?
- Where are there errors or gaps that can be repaired now?
- What pull down lists are available to users when characterizing each asset?
- Define new standards for each characteristic
- Define a correction strategy for moving to the new data standards

METHODOLOGY

Data Fields:

Equipment	Technical IdentNo.	Oil (Gals)	Nominal Voltage
Description	Main work center	Operator	Owner
Functional Location	Created on	Breaker PCB Level	Owner-Business Unit
Functional location description	Created by	Breaker PCB Test Date	PO
Status Text	Changed on	Rated Voltage (kV)	Pressure Vessel Permit
User Text	Changed by	Is Brkr a Tert Reac 500KV	Purchased
Equipment type	GUID	Total Weight (lbs)	Purchased Job Nbr
ABC Indicator	Long text exists	Clearance Window End	Removed
Plannergroup	Current work order	Clearance Window Start	Removed Job Nbr
Manufacturer	Asset #	Contract	SCADA Control
Manufactured date	Bill Third Party	Control Center	SCADA Status
Manufacturer Model no.	Breaker analyzer installed	Cost	SLIC Res ID
Manufacturer Serial no.	Transmission Line Section	Compliance Requirements	SLIC Res Type
Address	Used for Cap Bank Switching	Is Breaker DC Trip	Spec Nbr
Zipcode	Continuous Current (Amps)	Disposal Auth	Transmission Line Section 2
AuthorizGroup	FD Class	In Service Date	Equipment Sub Type
Inventory number	Breaker FPT Possible	Installed	Warranty End Date
Construction year	Gas Weight (lbs)	Installed Job Nbr	Warranty Start Date
Delivery date	Interrupt Capacity (Amps)	Intertie	Latitude
Start-up date	Interrupt Capacity (kVA)	ISO Controlled	Longitude
Location	Interrupt Medium	ISO Index Nbr	MANUFACTURER
Plant section	Breaker Lubrication	Functional location	MANUFACTURER MODEL NUMBER
Work center	Pre-BT MFR VALUE	Manufactured	MANUFACTURER PART NUMBER
Superord. Equipment	Maximum Micro Ohms	Mnt Dept	MANUFACTURER SERIAL NUMBER
Position	Number of Tanks	Special Protection Schemes	YEAR OF MANUFACTURE

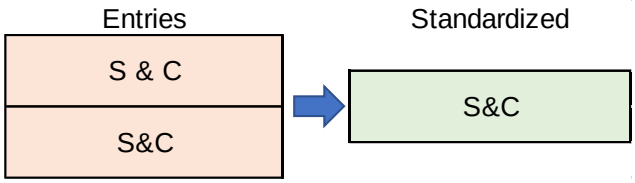
UMS group

www.umsgroup.com
info@umsgroup.com
Americas +1.973.335.3555
Europe +31.20.561.7033

METHODOLOGY (continued)

Example 1: Data cleansing process for the "Manufacturer" field:

1. Compiled a list of unique entries
2. Standardized the naming convention:



3. Updated the Manufacturer name to reflect mergers and acquisitions through research, e.g.

Original Entries	Current Entities
COOPER	Eaton
COOPER POWER	Eaton
Cutler Hammer	Eaton
EATON	Eaton
EATON COOPER POWER	Eaton
Eaton Cutler-Hammer	Eaton
KYLE	Eaton
MCGRAW EDISON1	Eaton
MCGRAW EDISON2	Eaton
SQUARE D2	Eaton
Westinghouse	Eaton

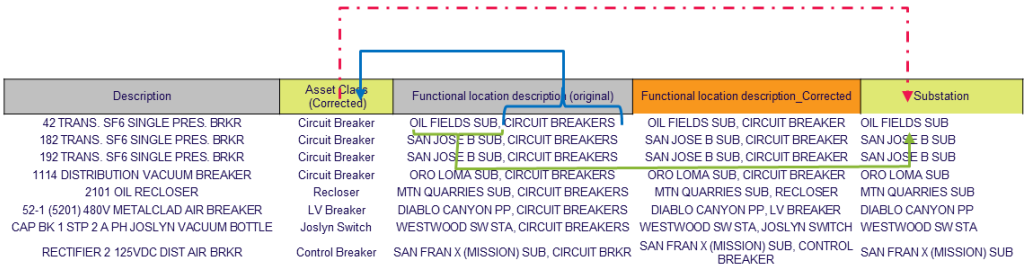
In 1947 Kyle merged with Line Material Company, which merged with McGraw in 1949, and then became part of Cooper Industries in 1985, and finally part of Eaton in 2012.

METHODOLOGY (continued)

Example 2: Parsing Out and Correcting Individual Data Columns:

Functional Location Description field

Equipment
Description
Functional Location
Functional location description
Status Text
User Text
Equipment type
ABC Indicator
Plannergroup
Manufacturer
Manufactured date
Manufacturer Model no.
Manufacturer Serial no.
Address
Zipcode
AuthorizGroup
Inventory number
Construction year
Delivery date
Start-up date
Location
Plant section
Work center
Superord. Equipment
Position



Description	Asset Class (Corrected)	Functional location description (original)	Functional location description_Corrected	Substation
42 TRANS. SF6 SINGLE PRES. BRKR	Circuit Breaker	OIL FIELDS SUB, CIRCUIT BREAKERS	OIL FIELDS SUB, CIRCUIT BREAKER	OIL FIELDS SUB
182 TRANS. SF6 SINGLE PRES. BRKR	Circuit Breaker	SAN JOSE B SUB, CIRCUIT BREAKERS	SAN JOSE B SUB, CIRCUIT BREAKER	SAN JOSE B SUB
192 TRANS. SF6 SINGLE PRES. BRKR	Circuit Breaker	SAN JOSE B SUB, CIRCUIT BREAKERS	SAN JOSE B SUB, CIRCUIT BREAKER	SAN JOSE B SUB
1114 DISTRIBUTION VACUUM BREAKER	Circuit Breaker	ORO LOMA SUB, CIRCUIT BREAKERS	ORO LOMA SUB, CIRCUIT BREAKER	ORO LOMA SUB
2101 OIL RECLOSER	Recloser	MTN QUARRIES SUB, CIRCUIT BREAKERS	MTN QUARRIES SUB, RECLOSER	MTN QUARRIES SUB
52-1 (5201) 480V METALCLAD AIR BREAKER	LV Breaker	DIABLO CANYON PP, CIRCUIT BREAKERS	DIABLO CANYON PP, LV BREAKER	DIABLO CANYON PP
CAP BK 1 STP 2 A PH JOSLYN VACUUM BOTTLE	Joslyn Switch	WESTWOOD SW STA, CIRCUIT BREAKERS	WESTWOOD SW STA, JOSLYN SWITCH	WESTWOOD SW STA
RECTIFIER 2 125VDC DIST AIR BRKR	Control Breaker	SAN FRAN X (MISSION) SUB, CIRCUIT BRKR	SAN FRAN X (MISSION) SUB, CONTROL BREAKER	SAN FRAN X (MISSION) SUB

Broke down the ***Functional location description*** field into two fields:

1. Asset Class
2. Substation

Used these fields to correct the ***Functional location description*** field

Standardized the classification of Circuit Breakers into a consistent approach

Broke out differences between Circuit Breakers, LV Breakers, Reclosers, and Joslyn Switches

Cleaned up Substation names

Example 3: Added new fields for asset attributes that were parsed out of existing fields

1. Metal Clad (Y/N) – based on ***Description***
2. Transmission or Distribution – based on ***FD Class***
3. Interrupting Medium – based on ***Interrupt Medium***
4. Single or Multi-Pressure - based on ***Description***
5. Single or Multi-Tank – for Oil Breakers, based on ***Number of Tanks***

METHODOLOGY (continued)

Example 4: Filled in missing fields from correlating to other sources:

1. Substation Name – 5,506 rows added
2. Zip Code – 250 rows added
3. Substation Address – 82 rows added, – now 98.3% complete

RESULTS

We identified 15 (of the original 100) fields are empty or have only garbage data in them, and recommended deletion to clean up the data - decluttering it by 15% and making room for new more useful fields.

We looked for opportunities to develop and fill in missing data …

Field Name	Initial Completion	Current Completion	Method	Notes
Manufacturers	99.7%	99.7%	Research, Standardization	Reduced number of categories by half from 44 to 22
Position, Detailed Asset Description, Asset Class	[New]	100%	Text/String Analysis, Manipulation, and Extraction	Derived the Fields from Description; Standardized Asset Class into 5 categories
Asset Class, Substation	[New]	100%	Text/String Analysis, Manipulation, and Extraction	Derived the Fields from Functional location description; Standardized Asset Class into 5 categories; Cleaned up Substation names
Functional location description-Corrected	[New]	100%	String concatenation	Based on standardized and corrected fields, generated the corrected field for Functional location description
Metal Clad (Y/N)	[New]	100%		Derived based on Description
Transmission or Distribution	[New]	99.7%		Derived based on FD Class
Single or Multi-Pressure	[New]	100%		Derived based on Description
Single or Multi-Tank	[New]	99.2% (Oil only)		for Oil Breakers, derived from Number of Tanks
Substation Name	[New]	100%	Text/String Analysis, Manipulation, and Extraction	Derived from Functional location description
Zip Code	69.9%	72.7%	Research	

The project ultimately resulted in a more accurate, complete, and robust data set with the confidence of the client, including new data standards and strategies for improved data governance going forward.

www.umsgroup.com
info@umsgroup.com
Americas +1.973.335.3555
Europe +31.20.561.7033