Informatica

PACIFIC DATA INTEGRATORS

Grid Health 360

Pacifc Data Integrators
5710 Lone Tree Blvd
Rocklin CA 95765
pacifcdataintegrators.com

Table of Contents

Grid Health 360 for Asset Management

Description of Analysis

This process is intended to show the potential output of an implementation of Informatica Master Data Management (MDM) in an environment with several databases with conflicting data and varying levels of "trustworthiness." The proceeding will validate the merging of a "Customer Information System" database with a "Meter Data Management" database. This portion of the showcase highlights how Informatica MDM uses trust scores to populate a "golden record of grid asset information". The goal of the analysis is to improve data quality and governance in order to increase the granularity of the analytics. The process works as follows:

- First, database objects are on-loaded and assigned trust scores bases on a series of attributes and dependencies.
- Next, Informatica MDM is used to bulk load and process the data into a golden record.
- After auto-merging most objects, "null score" objects are sent to a data steward for review.
- Finally, the record can be used for more accurate analytics e.g. high-resolution load planning, fraud mitigation, and resolution of missing pulses.

In this use-case, the "Customer Information System" (Oracle) database is merged with a "Meter Data Management" (SAP) database. Examples of possible object scenarios merged in this case include:

- Conflicting meter start dates
- Meter/transformer miss-allocations
- Conflicting usage data, etc.

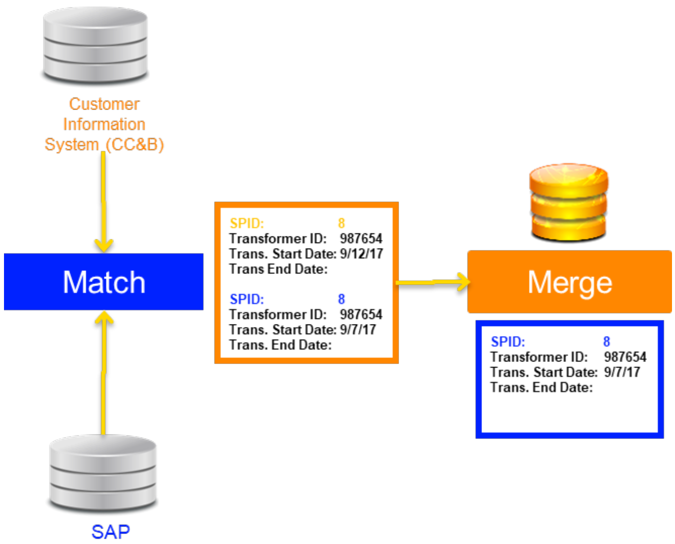

Figure 1. Auto-merge of Oracle and SAP databases based on trust scores

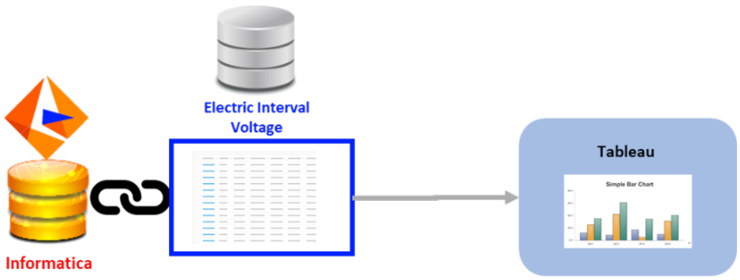

Figure 2. Informatica MDM "Golden Record" and Tableau Interaction

Informatica MDM "golden record" data (i.e. usage, meter manufacture, meter installer "technician" and geo-location) is run through an analytics environment to identify load patterns and anomolies such as missing pulses. In this case, Tableau in used as the trend analytics and mapping software [Fig. 2]. Mesh network issues can be isolated to manufacturers, technicians, or geographical or weather conditions. With the Informatica MDM "golden record," the following analytics are possible:

- Missing pulse resolution
- Identify need for supplier involvement or technician training
- Identification of unregistered solar panels or batteries
- Improved load planning
- Fraud mitigation

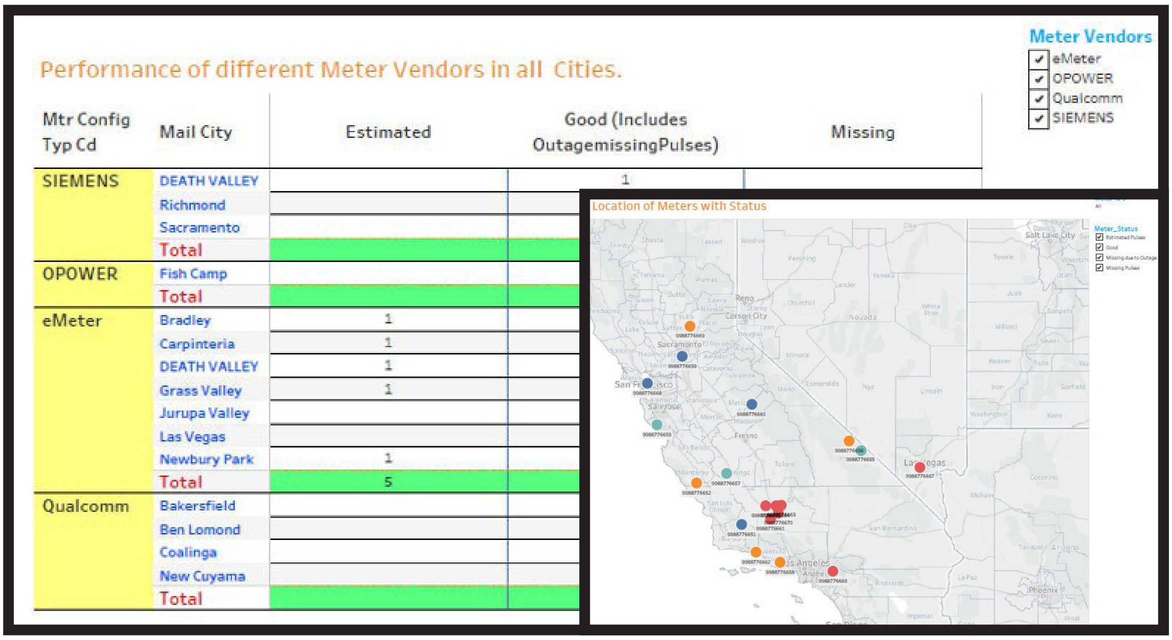

Figure 3. Missing Meter Pulse Resolution

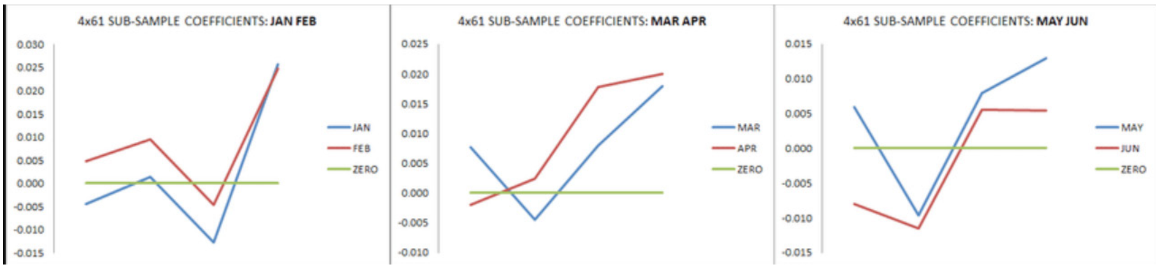

Figure 4. Highly Granular Analytics

More complex analytics become possible when public domain data sources are included. Social media data can be used to incorporate data objects such as demographics to understand how grid-connected devices (referred to as distributed energy resources "DERs") spread in a particular demographic or geographic service segment. Weather data is just one example incorporated in this showcase to normalize pulse data in order to identify suspected DER or battery storage installations.

Vegetation Management

LiDAR Transmission Lines vs. Tree Canopy

Description of Analysis

This process is intended to show a potential output of an analysis comparing the proximity of tree canopy to transmission lines. The showcase used four public domain raw .LAZ lidar pointcloud files in a forested region in Portland, Oregon. The lidar data was first run through a pipeline of LAStools tools in order to create a 1m2 resolution GIS-compatible raster dataset wherein each pixel represents the maximum height of trees in the cell. The final steps in this process were performed in FUSION LiDAR software. The process works as follows:

- First, the raw .laz file is converted to a local projection with meters as units.
- Next, the .laz is filtered for noise, removing spurious points.
- After calculating approximate height-above-ground, the .laz files are classified into 'trees,' 'buildings,' and 'ground' points.
- Finally, maximum elevation of all 'tree' points within a grid of 1m2 cells is created.

After creating this 'canopy-only elevation' dataset, a transmission line dataset was incorporated. In this case the transmission line information was derived from public domain maps via the Energy.gov website. This GIS vector file was converted into a raster, with elevation data approximated as 40 meters to the bare-earth elevation. Next, this transmission line raster was buffered by 50 meters, and this buffer was ultimately used to clip a subset of the canopy-only elevation raster. Both the transmission line raster and this new nearby canopy layer were converted to points, with 1 point created per raster cell. The 2D (Euclidean) distance was calculated between each nearby canopy and the closest transmission line point. This process also recorded the ID of each transmission line point. A table was created from these distances, wherein the height of each nearby canopy point was recorded alongside its corresponding transmission line point ID. A join allowed the transmission line heights to be appended to the table, based on ID. In order to calculate 3D (trigonometric) distance, the Pythagorean theorem was used on variables in this table:

$$Trigonometric\ Distance = \sqrt{(Canopy\ Height - Transmission\ Line\ Height)^2 + Euclidean\ Distance^2}$$

For nearby canopy points that were higher than their corresponding transmission line points, 2D distance was kept. These final distance values were inserted into a new raster, and the resulting distance surface was mapped.

Summary of Workflow:

1. Obtain .LAZ files
2. Process .LAZ into a canopy-only elevation raster
3. Merge transmission line data w/ elevation data
4. Extract and transform transmission line into raster with values corresponding to elevation
5. Transform transmission line raster to points
6. Transform nearby canopy [see 6] to points
7. Create table of nearest features and distances
8. Calculate new column in table [see 7] for 3D distance, using 2D distance if the tree is taller than the transmission line.
9. Create duplicate of nearby canopy raster, replacing values with the 3D distances found in [8]
10. Tile the distance raster and map.

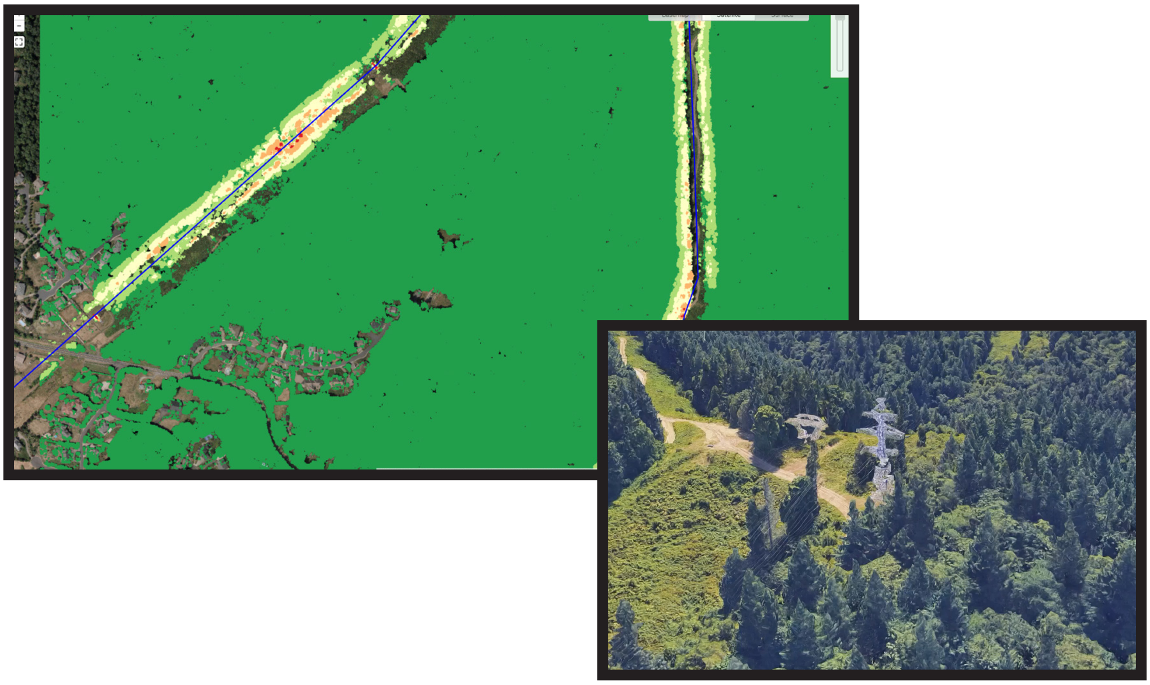

Figure 5: From Lidar data risk of arcing is interpolated. 3D viewing provides context and verification.

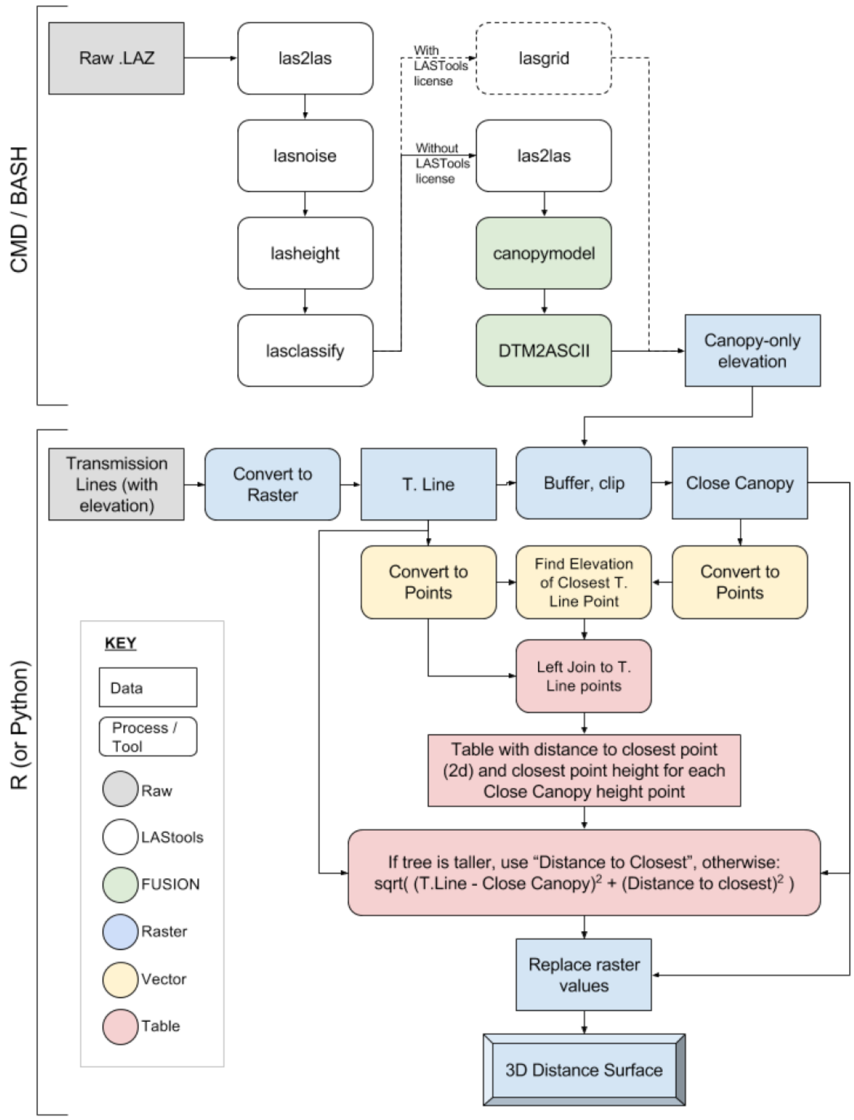

Figure 6: Workflow Diagram

About Informatica

Informatica is the only Enterprise Cloud Data Management leader that accelerates data-driven digital transformation. Informatica enables companies to unleash the power of data to become more agile, realize new growth opportunities, lead to new inventions resulting in intelligent market disruptions. With over 7,000 customers worldwide, Informatica is the trusted leader in Enterprise Cloud Data Management.

For more information, call +1 650-385-5000 (1-800-653-3871 in the U.S.), or visit www.informatica.com. Connect with Informatica on LinkedIn, Twitter and Facebook.

About Pacific Data Integrators

Pacific Data Integrators (PDI) is an Informatica partner with a strong emphasis in the utility industry. For more information about our MDM solution, call +1 800-403-5213, or visit pacificdataintegrators.com. You can also connect with PDI via LinkedIn, Twitter, Facebook and Google +.