Laurel Hill GIS Utility Network Data Assessment

Version 1.0

November 2022

307 Bross Street ● Longmont, Colorado 80501 ● Voice: 303.772.6856 ● Cell: 303.775.6740 ● Email: matt@lhgis.com

Table of Contents

Introduction

This whitepaper will discuss the key elements of a Utility Network Data Assessment in preparation to migration from the Geometric Network to the Utility Network. The term strategic is used because there will always be some level of error in a geodatabase, choosing the right strategy to detect and correct the most important errors that will provide the highest return on investment for the migration. Additionally, there is always limited resources and time to perform the migration therefore we must target the highest value corrections possible to make sure the migration goes off as planned. this paper will discuss the use of geodata sentry, an automated QA/QC product for testing the geodatabase, as the strategy to quickly configure tests and detect errors to support strategic planning, analysis, and remediation prior to migrating data to the utility network.

Planning

Planning for migration requires identification of what is the most important part of the network and the attribution that must be correct to comply with the new rules and architecture of the utility network. key items such as geometric connectivity and logical connectivity with respect to the connectivity rules of the utility network as well as key attributes such as open or closed status for valves and switches or phasing and MAOP design, must be identified early in the process and validated to ensure that their values and their connections are correct. One of the Significant differences between the geometric network and the utility network is the ability to perform actions with data errors the geometric network is for forgiving in the sense that it will allow tracing and other operations on the network while there are data errors. These may be connectivity errors where the data does not comply conform to the connectivity rules or low-level geometry errors that allow for these operations. If these types of errors are not corrected in the geometric network in the source data for the migration, then the result will be dirty areas in the utility network these dirty areas must be cleared to perform the building of subnetworks for instance for gas, the system subnetwork cannot be built with any geometric errors or any of the connectivity errors. If the open and closed status of certain valves is not correct, then the building of pressure systems subnetworks will not work correctly. These types of errors are categorized as severe and must be corrected before the migration can begin.

Strategic Validation

Once the planning is complete for the validation the tests can be created with geodata century each test type is identified with severity ranging from 1 being least severe to 3 being the most severe the most severe errors detected will cause dirty areas to be created in the utility network because of this these errors must be corrected prior to migration or during the migration process. Each test group describe below we'll have a specific severity assigned.

Data Model Conformance Validation

Data Model Conformance is defined as how well the data conforms to the geodatabase rules and properties There are several tests that can be configured such as subtypes and domains. These tests detect errors with the data or errors in the geodatabase data model properties. In either case the errors that are detected need to be evaluated for Utility Network migration impact.

Subtype Validation

This test suite tests feature classes and related tables with subtype control to ensure that all subtype values are valid. The test suite is automatically generated from the subtype rules in the geodatabase. The subtype is the most critical attribute for ArcGIS and is the basis for the Asset Types for the feature classes in the Utility Network. Correct feature migration relies on valid Subtypes in the source data.

For these reasons the Severity for Subtype Validation is the highest at 3. Prior to migration all subtype errors need to be corrected.

Domain Validation

This test suite tests feature level coded and range domain values. Ensuring that where a column is domain controlled, that the values within the column are valid. The test suite is automatically generated from the domain rules in the geodatabase. Domain errors may affect migration to the Utility Network. Values specifically related to valve or switch status, active or abandoned, pressure or phasing will likely impact the creation of subnetworks in the Utility Network. Others may have little to no affect to migration, however, could affect symbology or definition queries in the ArcGIS Pro Project for the Utility Network.

For these reasons the Severity for Domain Validation is 2. Prior to migration all Domain errors need to be reviewed for Utility Network and ArcGIS Pro Project impacts.

Subtype Domain Validation

This test suite tests subtype level coded and range domain testing. Ensuring that where a column is domain controlled, that the values within the column are valid. The test suite is automatically generated from the subtype and domain rules in the geodatabase. This test is essentially a domain test with a subtype filter with similar importance as the domain tests described above. The Utility Network aggregates multiple subtypes into line, junction, and device features classes. The assignment of domains by Subtype is a critical aspect of migration, and the subtype dependent domain errors are likely to have an impact on the migration to the Utility Network.

For these reasons the Risk Level for Domain Validation is 3. Prior to migration all Domain errors need to be reviewed for Utility Network and ArcGIS Pro Project impacts.

Referential Integrity

Referential Integrity test detects mismatches between two tables based on the primary key/ foreign key relationship. All rows in one table are related to all rows in a second table based on a common key item. This test may be run in both directions to ensure that orphan records do not exist in either table. Relationship errors must be reviewed for impact to migration to the Utility Network. Feature associated in the Utility Network such as metersetting to meter may be impacted if there are Referential Integrity errors in the source data.

For these reasons the Severity for Referential Integrity Validation is 2. Prior to migration all Referential Integrity errors need to be reviewed for impacts to the Utility Network.

Annotation Validation

The Annotation Validation tests the integrity of feature linked annotation. A test is created for each annotation expression. The annotation expression is evaluated based and compared to the actual annotation text strings reporting differences. If the database values differ from the annotation, it can signal that the annotation has been edited, rather than the database being updated, and the annotation automatically being updated. It may also be the case where the relationship link between the OBJECTID of the feature and the FEATUREID of the annotation is invalid. While the database may be up to date, the annotation cannot be updated due to the faulty linkage.

If replacement of annotation with dynamic labeling is part of the Utility Network migration plan, this test is important. Any differences will need to be reviewed and evaluated for corrections to the annotation, or the database values that drive the expressions.

For these reasons the Severity for Annotation Validation is 2. Prior to migration all Annotation Validation errors need to be reviewed for Utility Network impacts.

Logical Connectivity Validation

Logical Connectivity tests are generated from the Geometric Network Connectivity rules. There are three categories of tests that validation connections between junctions and edges based on utility business rules. A geometric network for an electric utility may require that an overhead device only be connected to an overhead conductor, or that a service meter may only be connected to a service lateral. A water utility may require that all types of fire hydrants must be connected to fire hydrant laterals. These business rules control data at the subtype level and are stored in the geometric network connectivity rules.

The Geometric Network architecture will allow for tracing and analysis with junction/edge connection errors. On the other hand, the Utility Network architecture creates a Dirty Area for each invalid network connection. All Dirty Areas must be cleared in the network topology to create Utility Network subnetworks. This restriction makes feature to feature connections based on asset group and asset type critical.

For these reasons the Severity for All Logical Connectivity Validation is the highest at 5. Prior to migration all Logical Connectivity errors need to be corrected.

Junction Edge Validation

The Junction to Edge Test detects where a junction is connected to an edge where it should not be, based on the connectivity rules from the geometric network. Examples of junction to edge rules are: ServiceLocation junctions can only connect to Service Pipes, or Overhead Service Taps can only connect to Overhead Conductors.

Edge Edge Junction Validation

The Edge to Edge with Junction Filter Test detects where two edges are connected at a junction where they should not be, based on the connectivity rules from the geometric network. Examples of Edge Edge Junction rules are Steel Distribution Main can connect to Steel Distribution Main at a Regulator, or Overhead Conductors can connect to Underground Conductors at Riser junctions.

Edge Edge Validation

The Edge Edge Test detects where two edges are connected that should not be, based on the connectivity rules from the geometric network. An example of Edge Edge rule is A service conductor cannot connect to a transmission conductor regardless of the connecting junction.

Geometric Validation

The Geometric Network architecture will allow for tracing and analysis with low level geometry errors. On the other hand, the Utility Network architecture creates a Dirty Area for most of the geometry errors. All Dirty Areas must be cleared in the network topology to create Utility Network subnetworks. This restriction makes detecting and correcting geometry errors critical to migration to the Utility Network.

For these reasons the Severity for Geometry Validation is the highest at 3. Prior to migration all Logical Connectivity errors need to be corrected

Duplicate Geometry Validation

The Duplicate Geometry test detects duplicated geometries for points, lines and polygons. Duplicate features can exist in the geometric network and still allow for tracing and analysis. Stacked features should not occur in the Geometric Network, however if they are necessary for the Utility Network, then a Z value will need to be assigned to clear Dirty Areas.

Duplicate Vertex Validation

The Duplicate Vertex test detects duplicated vertices for edges. Duplicate vertices can exist in the geometric network and still allow for tracing and analysis. Each edge with a duplicated vertex will create a Dirty Area that will need to be cleared.

Overshoot/Undershoot Validation

The Overshoot/Undershoot Test detects features from one feature class that do not connect with feature of the same feature class or an optional second feature class. This test is critical to detecting spatial problems within and between feature classes. Small gaps or undershoot or small extensions of edges or overshoots may result in poor tracing within a feature class or between feature classes. An example of this is a service later line feature class that does not fully connect to the main feature class. This may result in poor spatial analysis when tracing to custom meters. These errors may be detected using standard tracing methods, but the process is a visual inspection that can be labor intensive. Additionally, these undershoot/overshoot feature may not be visible with tracing if there is a back-feed situation.

These errors may not create Dirty Areas in the Utility Network topology, and therefore do not have an automated method for detection other than this test.

Intersect Validation

The Intersect Test detects features from one feature class that do not intersect with features from a second feature class. This test is critical to detecting spatial problems between feature classes. The quality of the Utility Network may be compromised if features that are supposed to participate in the network are not coincident with the network. Not all Intersect errors create Dirty Areas in the Utility Network. If the junction or device feature is far enough away from the edge it should be intersecting, then a Dirty Area may not be created. Intersect errors can affect feature to feature associations, and potentially remain in the data until the features are interacted with, unless they are detected by this test.

For these reasons the Severity for Intersect Validation is the highest at 3. Prior to migration all Intersect errors need to be corrected

Invalid Geometry Validation

The Invalid Geometry Test detects features with invalid geometry with lines, points, polygons and annotation. There are 13 different types of geometry problems that this test identifies, with the most common being NULL Geometry. Each of the features identified by this test need to be reviewed and the geometry issue corrected.

For these reasons the Severity for Invalid Geometry Validation is the highest at 5. Prior to migration all Intersect errors need to be corrected

Overlapping Edge Validation

The Overlapping Edge test detects where there is overlap between edges within a feature class and between feature classes. Overlapping geometries may not create Dirty Areas, however they may cause errors when creating subnetworks. Each of the features identified by this test need to be reviewed and the overlapping issues corrected.

For these reasons the Severity for Overlapping Edge Validation is the highest at 3. Prior to migration all Overlapping Edge errors need to be corrected

Cutback Validation

The cutback test detects where an edge cuts back are an acute angle. Cutback can create jagged geometries in the data, but are not necessarily errors, however generally they are not desirable and usually are incorrect.
For these reasons the Severity for Cutback Validation is the highest at 1. Prior to migration all Cutback errors need to be reviewed and corrected if possible

Disconnected Edge

The disconnected edge test detects edge that do not connect to any other edge in the geometric network. These edges do not create dirty areas in the Utility Network; however these edges will not trace, and will not be assigned to subnetworks, potentially causing downstream impacts.

For these reasons the Severity for Disconnected Edge Validation is the highest at 3. Prior to migration all Disconnected Edge errors need to be corrected

Minimum Feature Length Validation

The Minimum Feature Length test detects edges that are shorter than a user defined feature length. Short segments are can potentially cause dirty areas in the Utility Network and could be nearly zero length. Nearly Zero Length edges need to be deleted.

For these reasons the Severity for Minimum Feature Length Validation is the highest at 3. Prior to migration all Disconnected Edge errors need to be corrected.

Critical Business Rule Validation

Pressure Validation

Below are two critical pressure validations that must be run against the source data prior to migration to help streamline the Gas Pressure System Subnetwork creation since these errors do not cause Dirty Areas but can be time consuming to identify with the Utility Network.

For these reasons the Severity for Pressure Validation is the highest at 3. Prior to migration all Logical Connectivity errors need to be corrected.

MAOP or Pressure Class Connections – This test is an Edge Edge test that compares the MAOP or Pressure Class of Service Pipe to the MAOP or Pressure Class of the Distribution Main. This test may find issues with attribution of the edges, or that there are disconnects at service taps. Either way the errors must be corrected before migration to ensure proper Pressure Subnetwork creation.

Gas Pressure System Validations – This test is an Edge Edge Junction test that detects where there are pressure differences at junctions other than closed valves or regulators. The Gas Pressure System Subnetwork are built based on these rules, and if the data does not conform to these rules, then errors with be encountered when creating the Pressure Subnetworks.

Phase Validation

Below are two critical phase validations that must be run against the source data prior to migration to help streamline the Circuit Subnetwork creation since these errors do not cause Dirty Areas but can be time consuming to identify with the Utility Network.

For these reasons the Severity for Phase Validation is the highest at 3. Prior to migration all Logical Connectivity errors need to be corrected.

Phase Mismatch – This is a junction edge test that detects where junction and device feature phasing do not match the connected conductor phasing. These mismatches will cause issues when building Circuit subnetworks as well as outage management analysis.

Tie Switch FeederID and Phasing – This is an edge edge junction test that detects when different FeederIDs and phases are connected at closed switches. Ensuring that switch status is correct in the source data will streamline the creation of Circuit subnetworks.

ArcFM Required Field Validation

Source data that utilizes ArcFM will have required fields identified as Not NULL = TRUE. These columns must be populated or there will be an ArcFM QA/QC error thrown. Examples of required fields are MAOP, Size, Material, Gas Pressure System for Gas Data, and Phasing, Voltage, Size, Material, FeederID for electric data. This information is important and is likely to be relied upon as populated during migration. NULL data in these important columns should be reviewed and corrected to the highest extent possible prior to migration.

For these reasons the Severity for ArcFM Required Validation is the highest at 3. Prior to migration all ArcFM Required errors need to be corrected.

General Business Rule Validation

Unique Field Validation

The Unique test detects duplication and NULL values in columns that are expected to be unique. Devices identified as Subnetwork Controllers with specific attribute for the Subnetwork Controller name must be unique, Examples of these might be Gas Regulators, or Electric Feeder Circuit breakers. The existing naming conventions for these features are likely to be the source for the Subnetwork Controller names.

For these reasons the Severity for Unique Validation for subnetwork naming is the highest at 3. Prior to migration all duplicate value errors need to be corrected

Example Report Analysis

Utility Network Preparedness Risk Dashboard

A Utility Network Preparedness Dashboard can be created from GeoData Sentry reporting. Organizing the report based on the severity of the issue, helps prioritize error corrections. Below is an example of a dashboard that is created from the raw Sentry reports.

	Migration Impact		
	Acceptable - Severity 1	**Recommend Corrections - Severity 2**	**Must Correct - Severity 3**
Number of Tests	***11***	***21***	***17***
Number of Errors	***4292***	***4687***	***446***

Test Type	Errors
Cutback	113
Duplicate	154
Duplicate Vertex	1
Geometry	3301
Minimum Length	723
	4292

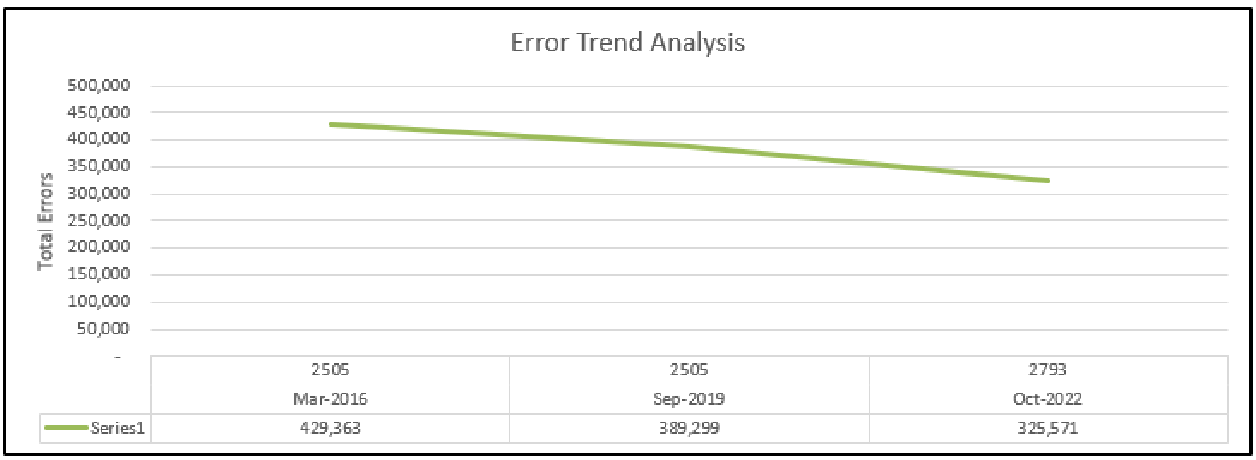
Test Type	Errors
Coded Domain	339
Cutback	6
Subtype Dependency	300
Subtype Null	3882
Unique Single Column	160
	4687

Test Type	Errors
Disconnected Edge	3
Duplicate	56
Edge - Edge with Junction Filter	6
Geometry	1
Intersect	176
Junction - Edge	186
Overlap	12
Overshoot/Undershoot	6
	446

Error Trend Analysis

Schedule validation of data allows for trend analysis over time. As errors are prioritized, and then corrected and re-running tests gives the opportunity to understand the trend of the corrections. Below are examples of testing over a 5 year period.

Error Trend March 2016-October 2022			
	Mar-2016	Sep-2019	Oct-2022
Total Tests Run	2505	2505	2793
Total Tests Passed	2,068	2,165	2,181
Total Tests Failed	368	272	202
Total Tests Failed due to errors	56	55	109
Total Rows Tested	116,766,406	118,688,607	120,336,457
Total Errors	429,363	389,299	325,571
Error Delta Period	-	(40,064)	(63,728)
Error Delta Cumulative	-	(40,064)	(103,792)
Error Rate	0.36771%	0.32800%	0.27055%



LaurelHillGIS
Geodatabase Integrity Solutions℠

Additionally, the errors can be categorized by the feature class and quantifying deltas over time. Below are errors counts shown between September 2019 and October 2022. The Red signifies increases in errors, and the green signifies a decrease.

Failures by Feature Class September 2019

Failures	(Multiple Items)
Row Labels	**Sum of Failures**
CapacitorAnno	13
CapacitorBank	25
CONDUCTORINFO	175485
DownGuy	5753
DynamicProDevice_Anno	12
DynamicProtectiveDeviceBank	21
ElectricDataset_Net_Junctions	4708
FuseAnno	14
FuseAnno_Tie	11
FuseBank	91
JOINTUSE	32956
Light	10154
Light_Anno	2
METER	1973
Misc_Anno	2512
MiscellaneousFeature	262
MiscNetworkFeature	52
MiscNetworkFeature_Anno	13
Pole	2593
PoleAnno	194
POLEASSEMBLY	71414
PrimaryOHLine_Anno	140
PrimaryOHLine_Feeder_Anno	97
PrimaryOHLine_Phase_Anno	116
PrimaryOHLine_Phase_Anno_C	2
PrimaryOHLine_Tie_Anno	31
PrimaryOHLineSection	796
PrimaryUGLine_Anno	282
PrimaryUGLine_Feeder_Anno	4
PrimaryUGLine_Phase_Anno	7

Failures by Feature Class October 2022

Failures	(Multiple Items)	
Row Labels	**Sum of Failures**	Delta
CapacitorAnno	11	-2
CapacitorBank	4	-21
CONDUCTORINFO	171234	-4251
DownGuy	5297	-456
DynamicProDevice_Anno	19	7
DynamicProtectiveDeviceBank	15	-6
ElectricDataset_Net_Junctions	6021	1313
FuseAnno	36	22
FuseAnno_Tie	5	-6
FuseBank	83	-8
JOINTUSE	32512	-444
Light	8777	-1377
METER	388	-1585
Misc_Anno	3660	1148
MiscellaneousFeature	30	-232
MiscNetworkFeature	7	-45
Pole	2546	-47
PoleAnno	216	22
POLEASSEMBLY	12753	-58661
PrimaryOHLine_Anno	119	-21
PrimaryOHLine_Feeder_Anno	328	231
PrimaryOHLine_Phase_Anno	109	-7
PrimaryOHLine_Tie_Anno	25	23
PrimaryOHLineSection	557	-239
PrimaryUGLine_Anno	269	-13
PrimaryUGLine_Feeder_Anno	30	26
PrimaryUGLine_Phase_Anno	7	0