



# Improving Confidence in Energy Distribution Network Connectivity Models

An advanced analytics approach



# Executive summary

Distribution network operators face an increasingly complex environment and typically manage multiple challenges daily, including safety, regulatory compliance, investor relations, critical infrastructure operation, supply quality, asset optimisation, use of system revenue, metering point administration, and market settlements. Of course, also of concern is the impact of increasing diversity in generation and demand, including renewables, distributed and embedded sources, demand-based energy sources, and energy storage technology.

To be fully effective in each of these areas it is becoming critically important for network operators to have more confidence in the data underlying systems and functions supporting areas of the business. The ability to bring data together to form a single federated view of the operation is now more urgent than ever.

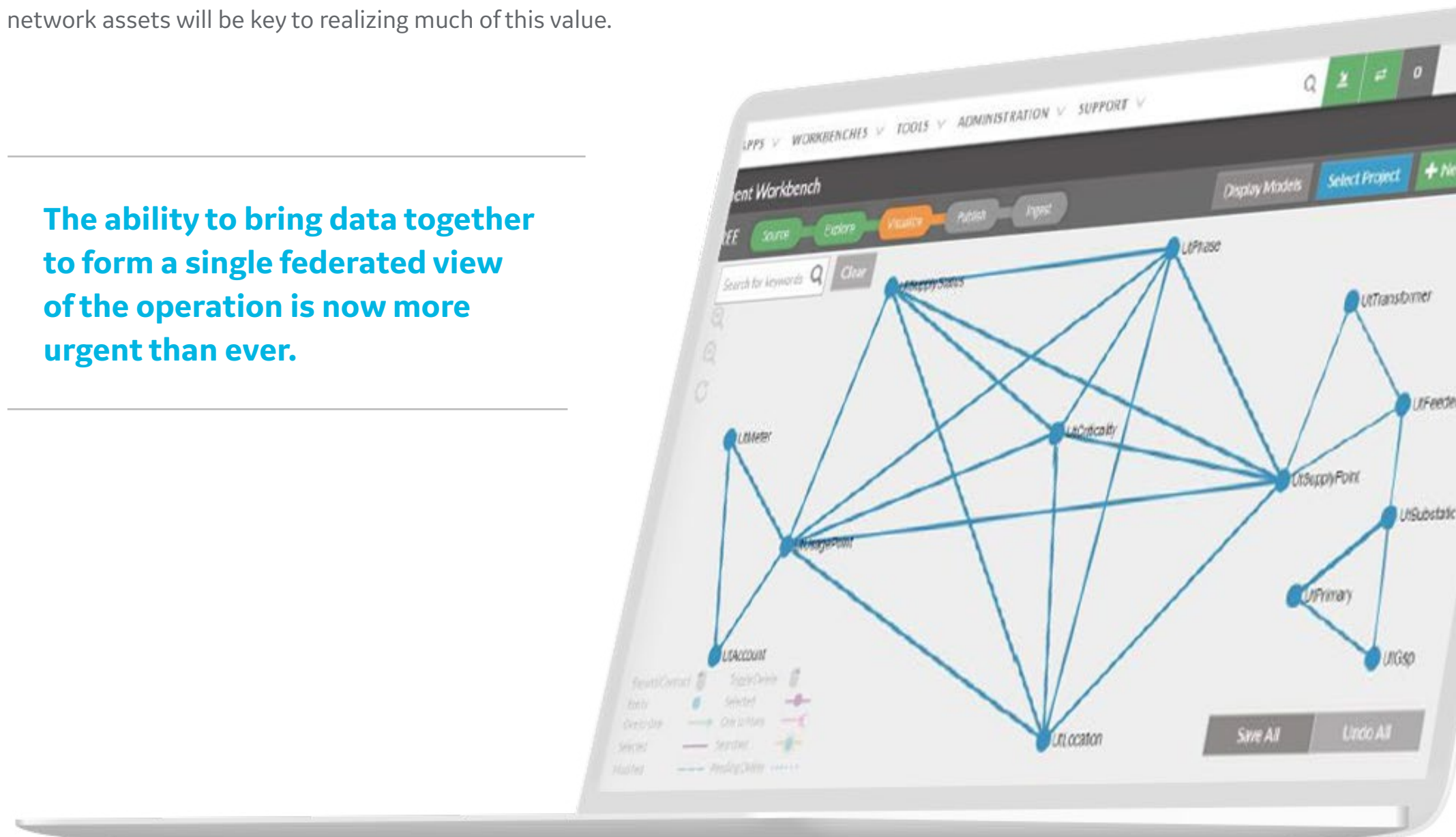
Of growing importance is the relationship between physical network assets and logical market-facing entities. This includes data in the form of measurements and events relating to assets; actual and virtual metering; logical aggregations; and complex product bundling for commercialization. The ability to reliably associate operational network data with energy retail identifiers is now important to many functions, including safety management, smart meter rollouts, network modelling, asset performance, maintenance, network operation, and supply quality reporting.

The global drive to install smart meters will also offer new opportunities to gain value from the richness of available data. The reliable association of measurement data and events to network supply point and low voltage (LV) network assets will be key to realizing much of this value.

---

**The ability to bring data together to form a single federated view of the operation is now more urgent than ever.**

---





## Supply quality KPIs and auditing

Supply quality is an area which has received more attention from regulators in recent years, and the progressive regulatory policies being applied now in deregulated markets in Europe are likely to be adopted in other markets in future. Now subject to KPIs, incentives, and penalties, supply quality measurement and reporting is a key area for improvement for many network operators. It is now possible that a proportion of a network operator's revenue up to 3%<sup>1</sup> may be at risk, depending on performance regarding customer interruptions (CI) and customer minutes lost (CML) KPIs. In addition, the systems and processes employed by network operators to report these KPIs, may be audited<sup>2</sup> from time to time.

Demonstrating confidence in the low voltage (LV) connectivity model used to report supply quality KPIs reduces the risk of revenue exposure affecting the bottom line for network operators.

Typically, network operators use a combination of data from multiple systems, such as GIS, OMS, CRM and customer call management systems, in conjunction with the LV connectivity model to calculate and report performance to supply quality KPIs. In many cases the integration challenge has yet to be met and resulting short-term workarounds have been employed to gather and process data for reporting. The automated modelling of the relationships between data held in the multiple source systems is key to future efficiency and effective reporting.

---

**Demonstrating confidence in the low voltage connectivity model used to report supply quality KPIs reduces the risk of revenue exposure affecting the bottom line for network operators.**

---

<sup>1</sup> <https://www.ofgem.gov.uk/electricity/distribution-networks/network-price-controls/quality-service/quality-service-incentives>

<sup>2</sup> <https://www.ofgem.gov.uk/publications-and-updates/information-and-incentives-project.-audit-incident-reporting-200102>

## Challenges affecting the reliability of LV connectivity models

Many existing utilities have been operating in one form or another for more than 60 years. These utilities typically have many millions of kilometres of physical infrastructure, much of which may have been installed prior to the existence of GIS and CAD applications, or indeed computers. This means that digital representations of today's network models are often composite views through various layers of history, including hand-drawn paper records, scanned raster images, modern GIS techniques, and geospatial tools and services. Errors in

the resulting network connectivity model can arise due to many reasons, including measurement accuracy, duplication, missing data, asset replacement, and new builds adding to and changing network records.

The potential for such errors to impact supply quality reporting and other processes is significant. To increase confidence in LV connectivity models, manual surveys of existing networks and connected metering points are now real considerations for some network operators. A better solution would be one which avoids the need for costly site visits.

---

**A better solution would be one which avoids the need for costly site visits.**

---



# Applying advanced analytics to the problem

Utilising the flexible plugin architecture provided by its Mlx Core Platform, GE Digital has developed and implemented advanced geospatial analytical applications for use in a broad range of network modelling applications, including low voltage and high voltage networks.

The solution uses machine learning for data integration, including data discovery, feature extraction, relationship modelling, and data ingestion. Using the Mlx Core Platform, existing data from utility systems such as GIS, SCADA, DMS, OMS, MDM, and WMS can be rapidly modelled, mapped, and ingested to provide a semantic data model representing the combined data set. This semantic model of the business and operations can be federated for use by many downstream analytical and operational applications.

The Mlx Core Platform also provides a framework for assertion-based logic, enabling the orchestration of analytical processes to be configured for each application.

GE Digital's network modelling applications enable the automated creation of a utility-specific network model using existing utility source data. The model is automatically built up through a process of orchestrated analytics which results in an estimation (including confidence score) of the connectivity between network assets referenced in the source data.

The Mlx Core Analytics Framework (MCAF) includes advanced geospatial and iterative analytics, which can be used to traverse asset records to estimate connectivity between assets. The system can operate in a fully autonomous mode to derive an estimate of asset connectivity based on a combination of rules defined in Mlx assertion logic and geospatial proximity analysis.

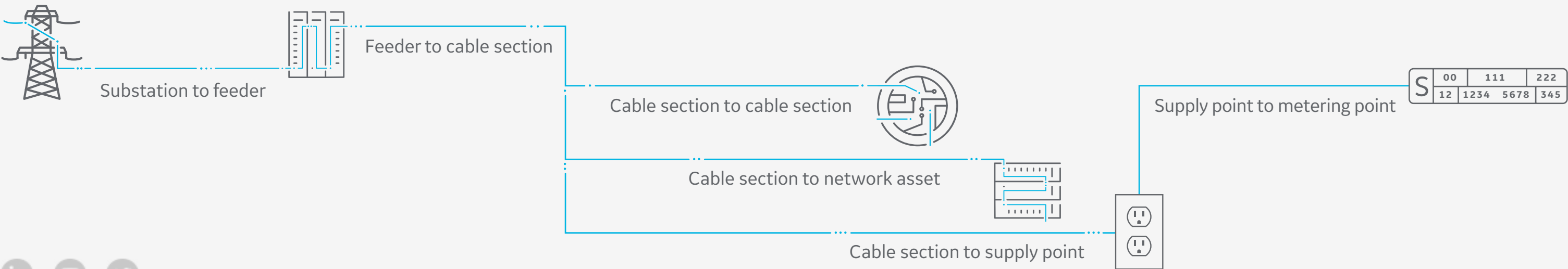
To make a reliable connection between the network entity supply point and retail market-identifier metering point, it is sometimes necessary to enrich address-based information with geo-coordinates integrated from an external source or service.

Once established, the connectivity model can be further enriched through the application of further analytical plugins. As an example, GE Digital's sophisticated methods for smart metering phase detection based on event monitoring can be used to record the phase to which each energy supply is connected.

The approach designed by GE Digital has also proved to be valuable in highlighting errors and inconsistencies in customer data. The estimation of connectivity by the system is subject to configurable tolerances and, as such, the system can automatically identify problems in the source data and flag these for investigation and improvement.

The system provides online tools to enable utility staff to interact with the model and decide on the corrective action to be implemented for each error identified.

## Associations and confidence scoring can include:





## The benefits of a reliable connectivity model

By application of an analytics approach, it is possible to quickly establish a working network connectivity model which supports immediate benefits such as better fault management and supply quality reporting.

Though projects should target and plan for further benefits stretching into the future, in the early days it is likely the model will highlight some limitations in the source data used. Tuning of confidence ratings will inform the optimum configuration of the model where control data is available. Further evolution of the model through incremental improvements will enable it to provide broader applicability and further benefits as reliability improves and adoption by other areas of business increases. Functions likely to benefit from the greater reliability offered include network fault diagnosis and restoration planning; smart meter event and consumption data management; workforce management; technical and non-technical losses; unmetered supplies management; the use of system charging, distribution network management, demand management, and curtailment; and asset optimization. Network investment

planning can also benefit, including assessing the effects of technological and commercial progression in the use of renewable energy, energy storage, virtual power plants, and demand management initiatives. Any downstream applications which rely on the aggregation of granular consumption data from smart meters, or have increased reliance on a "joined up" view of utility and market, will benefit from improved confidence in the network connectivity model.

---

**Any downstream applications which rely on the aggregation of granular consumption data from smart meters, or have increased reliance on a "joined up" view of utility and market, will benefit from improved confidence in the network connectivity model.**

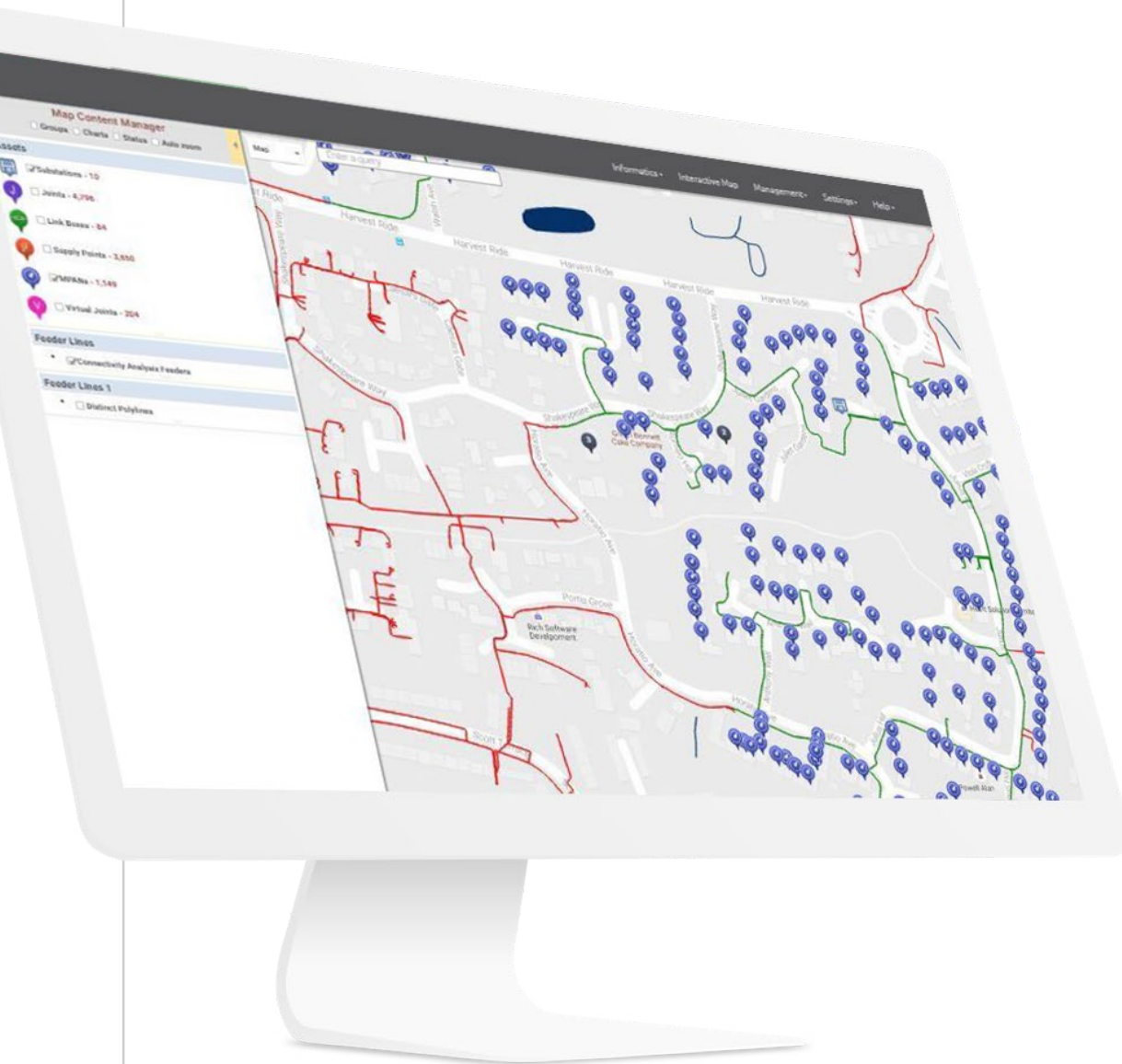
---





# Implemented solutions

GE Digital has used this approach to successfully implement network connectivity modelling solutions for customers in North America and Europe. The following section includes some examples and related functionality developed and implemented by GE Digital.



## LV connectivity model

In this project, geospatial algorithms were used to configure the LV connectivity model for a customer in Europe. The algorithms act on each distribution asset recorded in GIS source data to connect upstream and downstream entities. The algorithms estimate the connectivity between assets, with each connection assigned a confidence level. A further algorithm is used to establish connectivity between each 11KV/415V secondary substation and one or more cable section polylines in proximity of the substation.

The combination of results from this analysis enabled the model to estimate the start of each medium voltage feeder radiating from the substations. Having established the initial assets related to each feeder, the model then uses the iterative traversal algorithm to estimate the connectivity as it radiates to further branches of the LV feeder. The algorithm then follows the connected assets to a terminating asset in each case. In addition, the model includes rules which enable it to automatically correct some common data quality issues, such as inserting a "virtual joint" at a point where the model has estimated a structural joint to be missing from the source data. This innovative traversal model approach results in all assets (cable sections, joints, link boxes, distribution pillars, supply points) downstream of the substation being associated to a feeder ID and secondary substation. Further associations with HV network assets, primary substation, and grid supply point are then possible using existing HV and EHV network connectivity data.

Having established the connectivity model from substation and feeder down to supply point, the model then uses an algorithm to estimate the connection between each metering point and feeder. This first requires that each metering point record is supplemented with a pair of geo-coordinates that the model can use for its geo-spatial analysis. The solution implemented incorporates a software plugin to obtain the geo-coordinates for the metering point based on limited site address information available.

On completion of the modelling process, the solution provides interactive tools to enable staff at the utility to visualize and investigate the results. The solution provides an interactive map, metering point analysis, and LV connectivity analysis workbenches, including data grids and graphical chart indicators to provide flexible and intuitive ways to examine and interact with the model.

A geo-spatial representation of the model is provided as a map overlay onto the geographic area modelled. This is useful in examining the results of the model to identify and investigate issues in underlying source data.



# Extending the connectivity model

In this project, GE Digital implemented a solution for a customer in North America. The project used similar geo-spatial algorithms and techniques to deliver a rich set of functionality, including the modelling of the relationship between distribution network, supply point, and associated metering point.

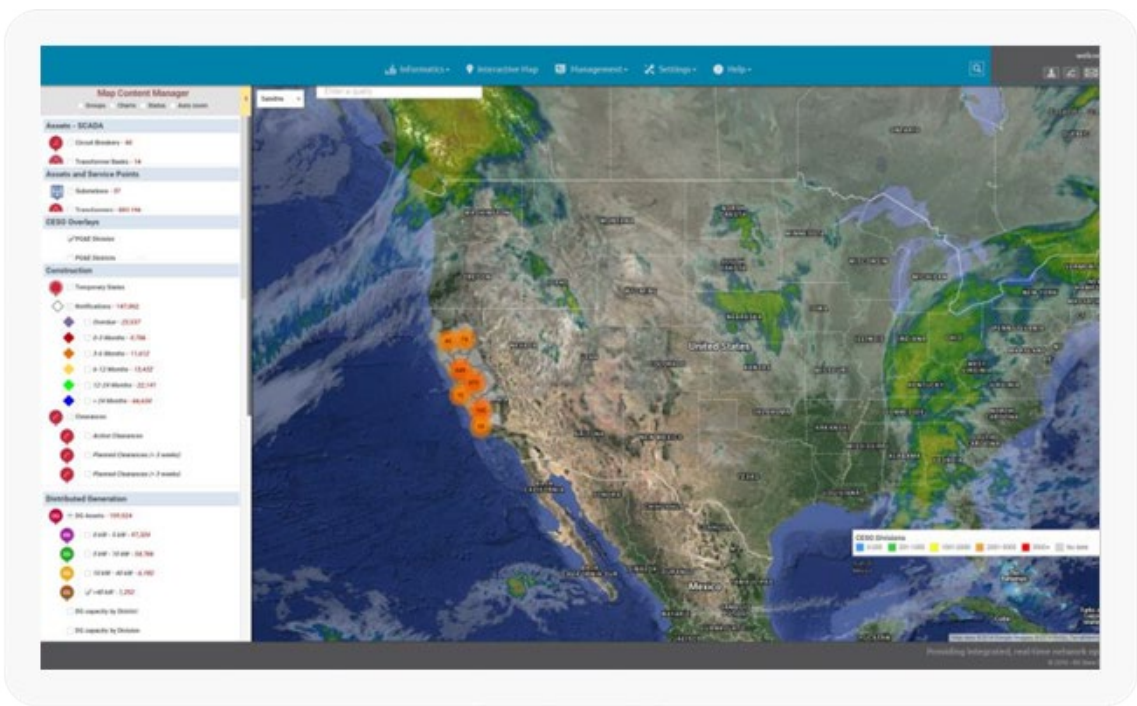
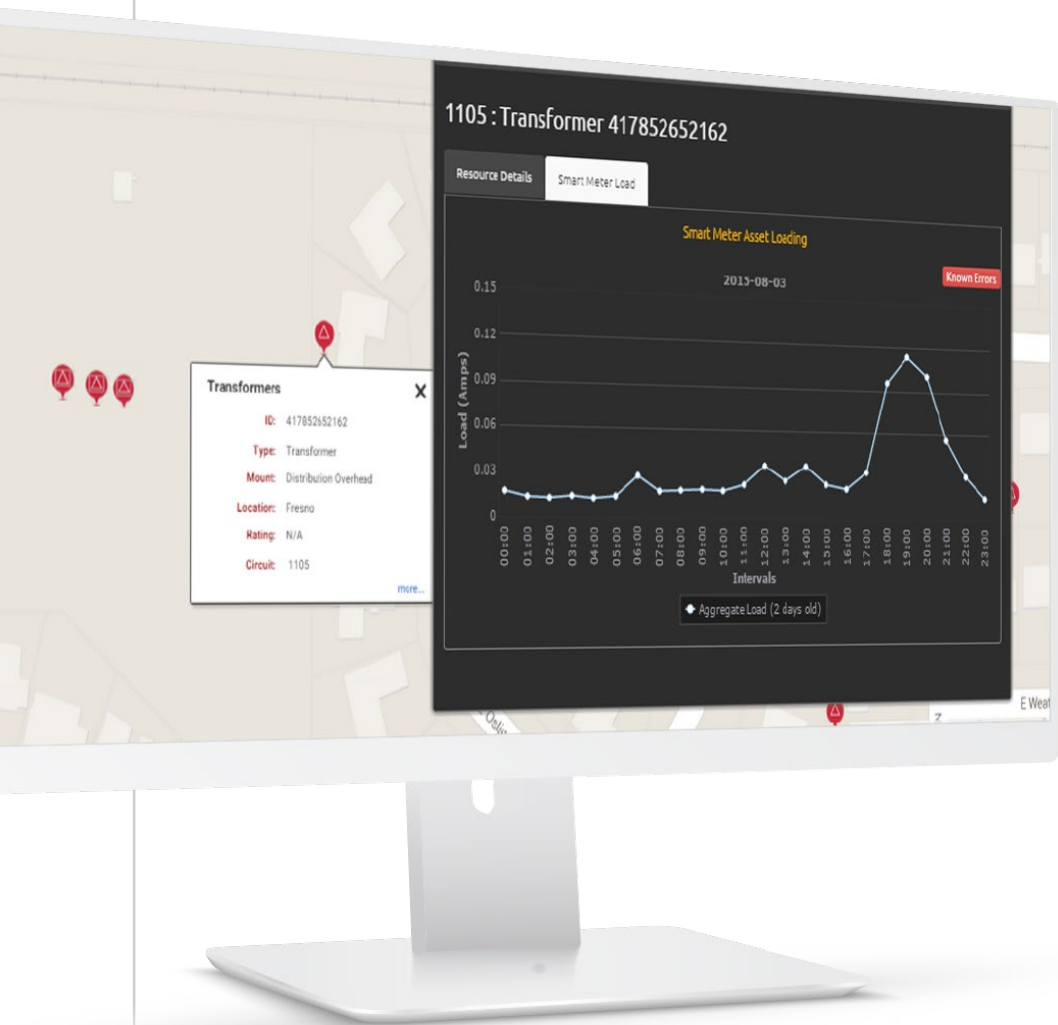
With the availability of granular time series data from smart meters, the system uses its connectivity model to aggregate the load at connected sites, and thereafter to indicate the loading on specific network assets.

The effects of loading on individual network assets can be monitored and staff can use the workbenches provided to nominate assets for closer monitoring by adding them to a watch list. The system also uses outage information in conjunction with the connectivity model to calculate and report supply quality KPIs.

Further uses of the system include the integration of real time environmental data feeds (weather, wildfire, earthquake) to further inform the model and equip engineering teams to minimize supply interruptions.

The data relative to each asset in the model can be quickly searched via powerful data grids with detailed drilldown facilities, giving access to deeper layers of distribution asset-related formation.

Having rich asset data and historic metrics relating to the utility asset base modelled, indexed, and federated by the platform provides an excellent base for the implementation of future value generating applications, such as asset-class performance management and predictive asset failure and maintenance programs.





# Learning from LV connectivity modelling

As mentioned earlier, the historic asset and geo-spatial information residing in utility systems can be very variable in quality and may pose many challenges to network modelling projects and operational systems. It is essential that any modelling approach implemented can be used to identify issues and provide methods to improve or work with flawed data. The ability to deliver meaningful results by working within the constraints of the existing data is key to the success of a network modelling solution. By delivering early results and enabling improvements to data quality over time, the solution will provide both immediate and long-term value to the business.

A key feature of the platform is its ability to create a semantic data model representing the combined data sources and processed inputs. By ingesting and processing source data, the semantic data model can provide an optimized perspective on the information, which best represents the network connectivity model in the real world. While subsequent correction of all source data is desirable, the time to achieve this may be significant. An alternative for both the short term (and potentially the longer term) may be to leverage the semantic data and provide federated access as required to other applications that may benefit.



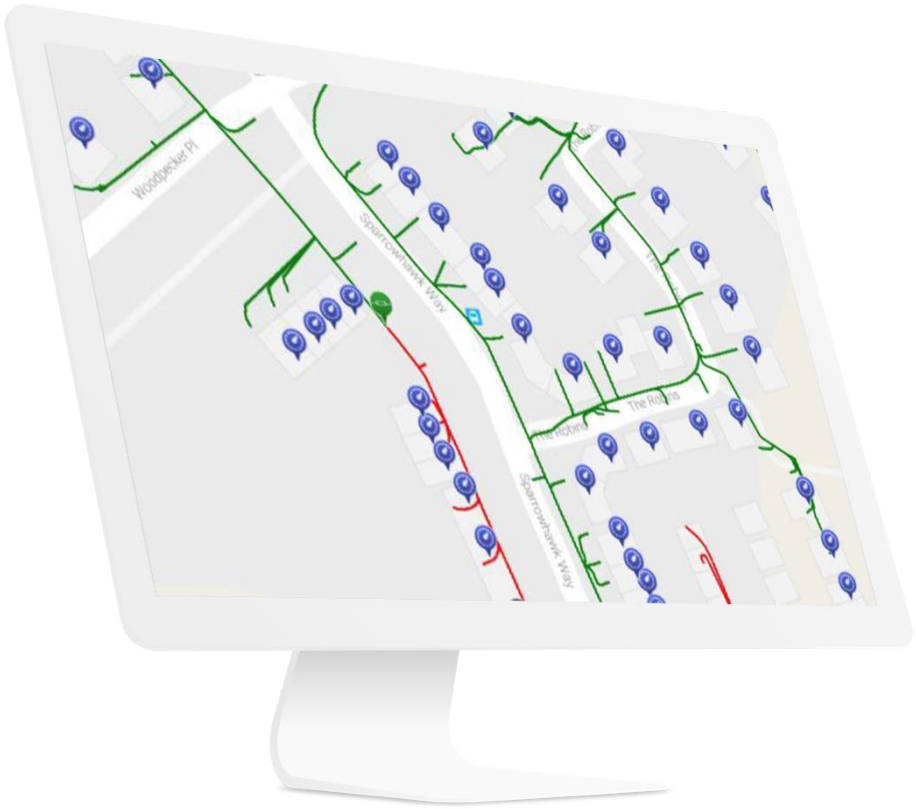


## Metering point and network asset to feeder association

With the GE Digital implementation for our European customer, we were successful in using the analytical model to automatically associate 90% of 1149 metering points to 32 calculated LV feeders.

Substation to LV feeder connections were established for 100% of feeders at a connection confidence score of 90% or higher.

Supply point to cable section connections were established for 99% of supply points with a connection confidence score of 90% or greater.



*Closed link box incorrectly terminating feeder before connected metering points.*



*"Virtual joint" automatically added by the model.*

## Limitations identified

As seen above, the initial network connectivity results produced by the analytical model were very encouraging. Once the model was populated, the tools provided by the system were used to investigate exceptions where the results were not as good as expected. The powerful visualizations and search features of the system enabled the team to quickly identify several data exception types, which were the root cause of most cases where the model had failed to associate assets as expected. The ability to rapidly detect these issues and quickly decide on corrective action enabled very agile improvement and evolution of the system.

Some examples of the types of exceptions found in initial test result include:

- Underground cable link box status unidentified in source data
- Missing service cable joints in the source data
- Sub-populated substation site operating as MV distribution pillar
- Historic source data erroneously recording terminating joints in locations where feeders have been extended by later developments

In the case of missing service cable joints, the introduction of a new rule through Mlx assertion logic enabled the model to automatically add a "virtual joint" to manage this exception type.

In addition to the limitations identified, some substation and feeder to metering point connections modelled were found to be different to the control set. In one such example the survey data for one substation showed feeder 2 and feeder 3 as having 21 and 0 MPANs connected respectively. The LV connectivity model results returned the same total, but different distribution with feeder 2 and 3 having 16 and 5 MPANs connected. While the survey data was thought to be correct, the modelled results were examined in detail and the project team concluded that it was possible the modelled results were more accurate than the surveyed feeder identification achieved by signal injection method.

# Mlx Core Platform from GE Digital

The methods and implementations described in this paper are powered by Mlx Core Platform from GE Digital.

Mlx Core is the premier platform for automated handling of complex data integration, data analysis, and predictive automation for connected devices on the Industrial Internet.

Mlx Core is designed from the ground up to provide automated complex IIoT data management with machine intelligence by creating relationships across industrial data regardless of format, frequency, or type. Mlx Core applies schema to data as it is ingested to speed integration. An optimized NoSQL architecture and ability to dynamically adapt to sources as they come online allows projects to ingest and integrate data six-times faster than traditional ETL.

## Semantic model and data classification

Information modelling is one of the most important aspects in the design of the Mlx Core technology. Information is at the heart of operational management, and utility businesses can now capture a vast amount of information related to events, assets, sensor data, meter reads, location details, and connectivity models. It is expected that the utility will collect more information than has ever been collected before, and this will stress existing capabilities including transmission, synchronization, storage, analysis, and reporting. The information received will be rich in content and can be used to drive new business processes, capabilities, and value.

A utility can expect to receive as much as 300 million data points daily for every 1 million smart meters installed on the network. The impact of information on the overall ecosystem cannot be understated, and utility businesses will need to design and build not just data management plans, but an information management plan.

The many factors to consider include:

- Common data models
- Canonical documents
- Governance and stewardship
- Quality management and integrity
- Operational data store
- Confidentiality and privacy
- Discovery, aggregation, and presentation
- Federation, integration, and synchronization

Mlx Core technology includes a comprehensive data model based on IEC CIM along with strict data handling services and embedded rules for ensuring the integrity, consistency, and quality of data.

## Key features of the platform include:

- **Rapid data integration**  
Fast, dynamic, and intelligently automated data ingestion from all operational and enterprise systems.
- **Machine intelligence**  
Analytic ensembles use machine learning algorithms to create context, learn patterns, detect anomalies, and store new intelligence into Mlx Core's knowledge repository.
- **Semantic modelling**  
Intelligently automate integration to model and map your data enterprise-wide.
- **Pluggable architecture**  
The "pluggable" nature of the platform gives you the ability to slot-in new repositories, new persistence methods, new access methods, and even new algorithms.
- **Dynamic adapters**  
Regardless of whether a data source speaks a proprietary language or not, the platform can integrate its data with a universal adapter.
- **Edge analytics**  
No more data latency. Analytics applied to data from control systems at the edge-to-cloud environments.





## Assertion-based logic

The assertion-based logic processing approach used within the Mlx Core technology is a key concept behind the artificial intelligence and machine learning. An assertion at its most basic level is a discrete processing rule that can operate independently or in combination with other rules, or even be executed in sequence or through a decision-tree process.

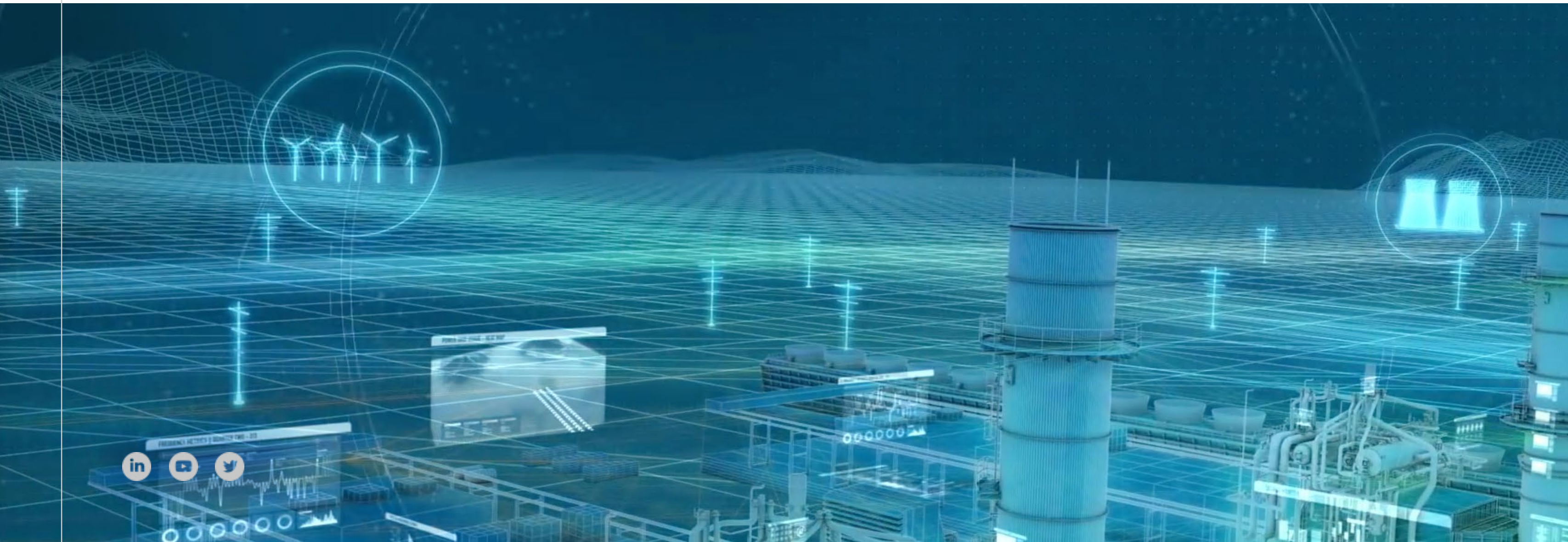
The assertions are used by the platform on a continuous basis to test data at rest and data in motion.

## Summary

Today, GE Digital is a driving force helping businesses realize value and maximize profit through the Industrial Internet of Things (IIoT). We continue to disrupt manual workflows and inefficient models of the past to instantaneously unify information and derive intelligence from billions of traditionally siloed data points. Daily, Mlx Core processes and analyses greater volumes of data than most of the largest social networks in the world.

Through data-driven intelligence and software-defined automation, GE Digital is enabling utilities to improve efficiencies, optimize results, and discover new business opportunities that increase revenue.

Through our work in utilities, we have become synonymous with successful software defined operations. GE Digital solutions have turned massive industrial data sets into data-driven automation for better operational efficiency, improved asset performance, and greater uptime. GE Digital automates complex operations by transforming disjointed data streams from control systems, connected devices, and industrial sensors into actionable intelligence. By uniting centrally managed data with connected devices at the edge of the network, GE Digital's solutions combine complex analysis and artificial intelligence to increase awareness, discover new business insights, and improve operations.





## About GE

GE (NYSE: GE) is the world's Digital Industrial Company, transforming industry with software-defined machines and solutions that are connected, responsive and predictive. GE is organized around a global exchange of knowledge, the "GE Store," through which each business shares and accesses the same technology, markets, structure and intellect. Each invention further fuels innovation and application across our industrial sectors. With people, services, technology and scale, GE delivers better outcomes for customers by speaking the language of industry.

## Contact Information

Americas: 1-855-YOUR1GE (1-855-968-7143)  
gedigital@ge.com

[www.ge.com/digital](http://www.ge.com/digital)

