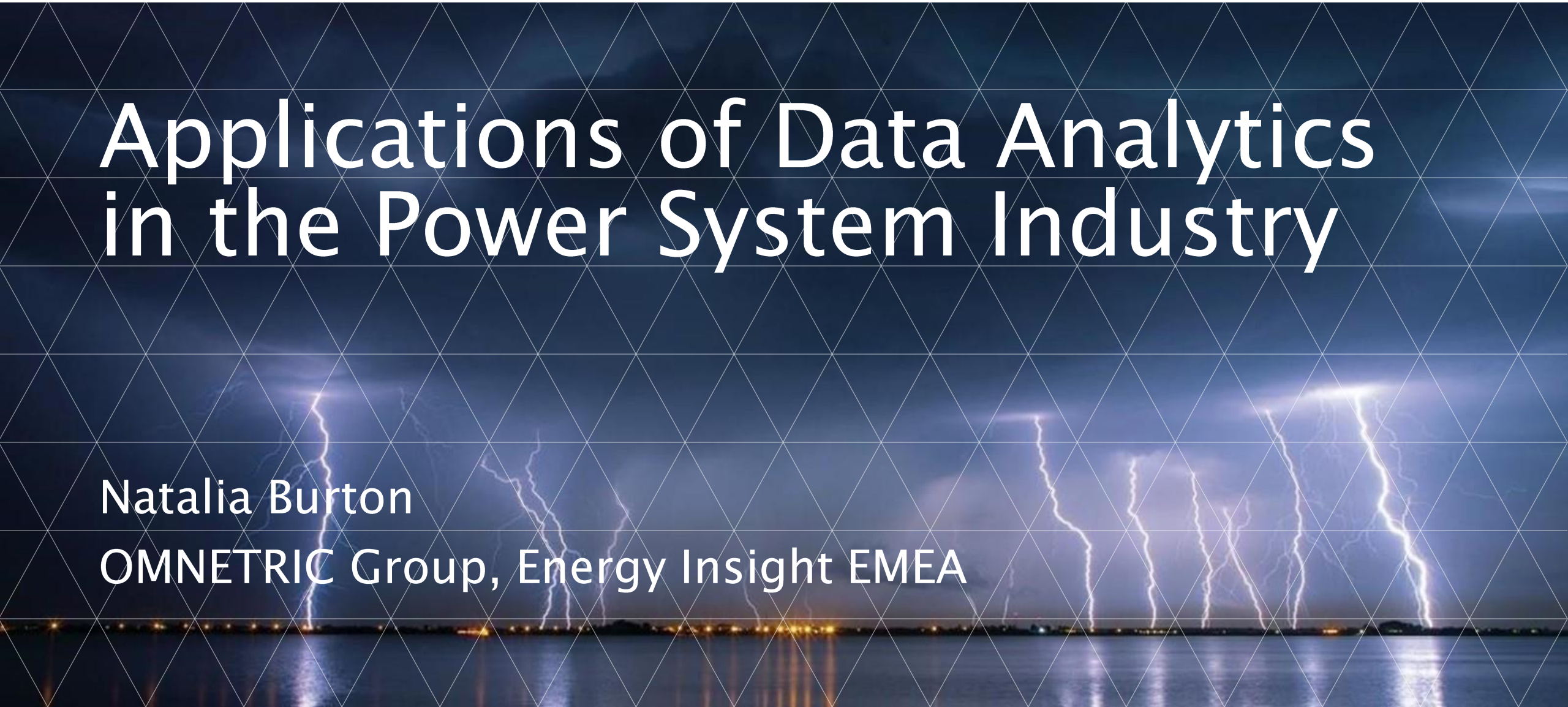
OMNETRIC Group

A Siemens & Accenture Company



Applications of Data Analytics in the Power System Industry

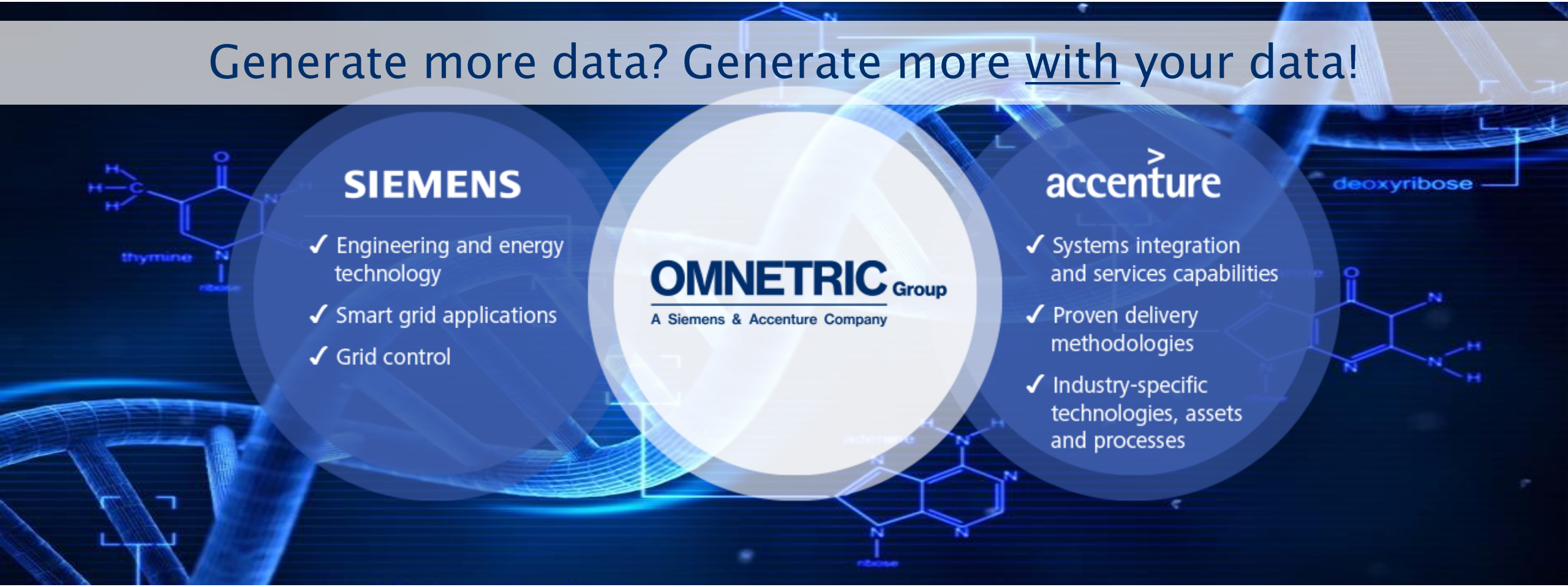
Natalia Burton

OMNETRIC Group, Energy Insight EMEA

Who is OMNETRIC Group

Out target in Energy Insight

Generate more data? Generate more <u>with</u> your data!

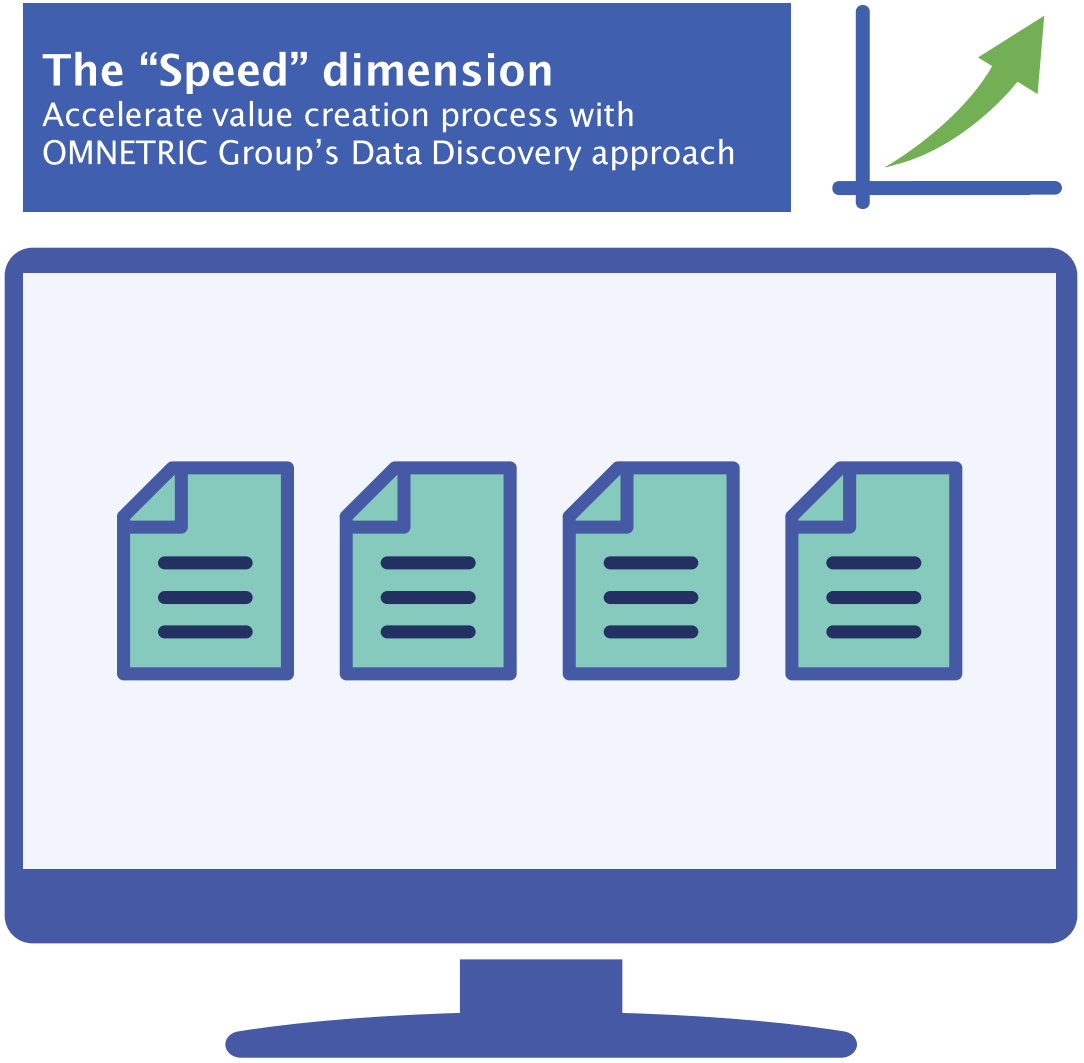


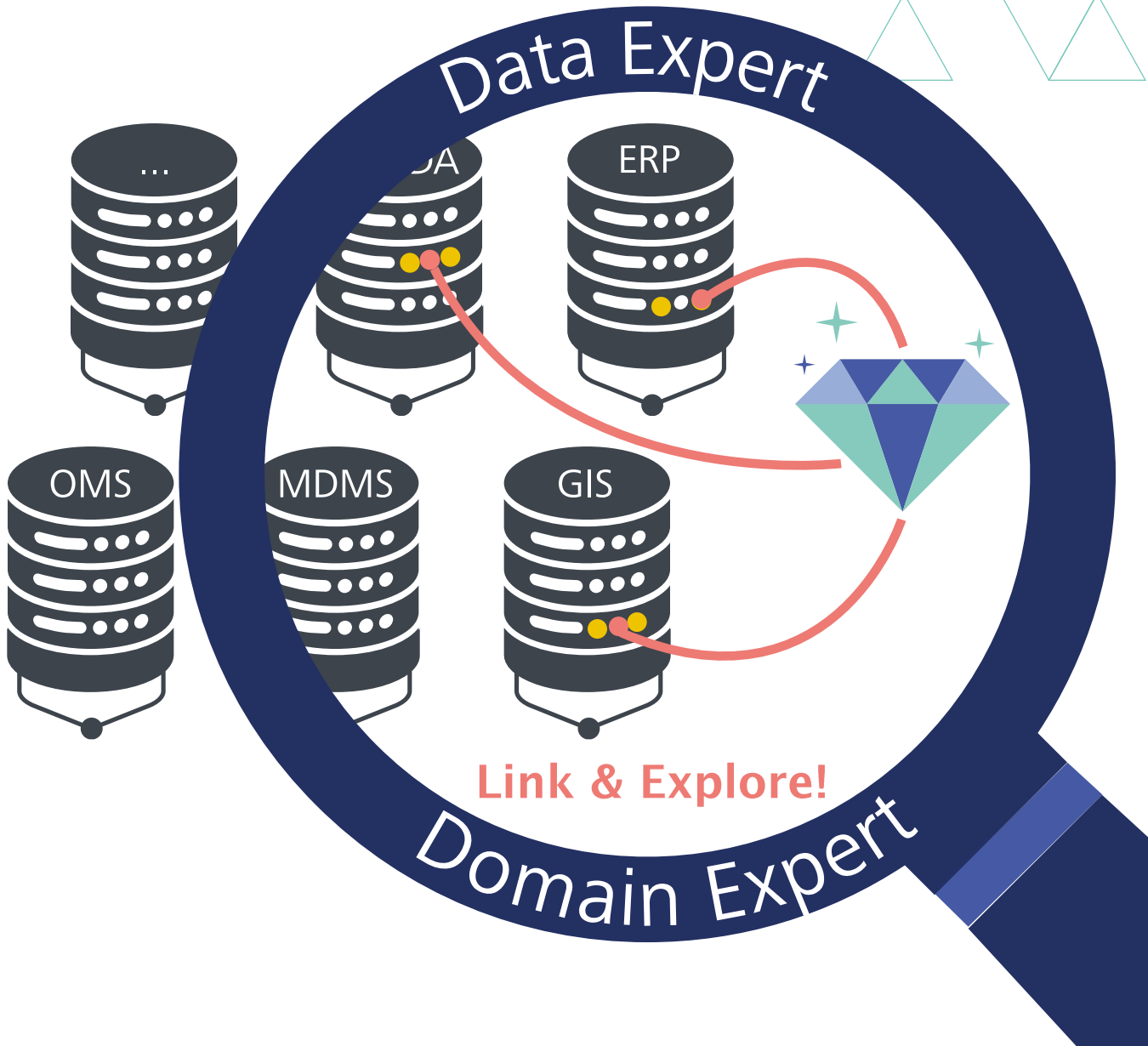
Two Worlds?

Operations & Analytics going hand-in-hand

“Diamond hunter” team: Combining IT and OT experts

The “Speed” dimension
Accelerate value creation process with OMNETRIC Group’s Data Discovery approach





For us „it‘s just data“ – analytics applications with basically all data sources in the energy sector

Data Expert

Domain Expert

1.1 AAA – Advanced Asset Analytics

Assets

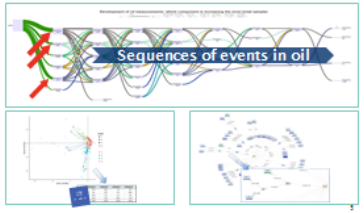
Detection and prediction of failure trends in measurement transformers

WHERE?
- European TSO

WHAT?
- Focus on measurement transformers: data from oil measurements is used to identify the health status of the assets
- Our approach: analyze data independent from the given assessment method
- Our aim: Secure insight into data
  - Identify influencing factors of the asset's health status
  - Predict how individual transformers behave in future

HOW?
- Data enrichment: weather data + external information
- Principle Component Analysis on oil and on gas measurements
- Event sequence analysis for error detection
- Decision tree classification and modelling

WHY?
- Insights into measurement transformers' disturbances
- Influencing factors for performance
- List of transformers that should be observed carefully due to their higher statistical risk profile



4.3 ACCP – Advanced Commercial Consumption Analysis

Customers

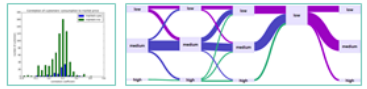
Understand what really matters to the big consumers: Weather or sports?

WHERE?
- Asian B2B energy retailer

WHAT?
- Analysis of business customer consumption
- Combining several data sources
  - Consumption / meter / contract data
  - Market data
  - Weather data
  - Special events

HOW?
- Time series analysis
- Correlation histogram
- Event and consumption sequence analysis

WHY?
- Address specific business questions on customer behavior
- Do the customers react to the market price?
  - No, they don't (at this client)
- Does consumption change with strong weather events?
  - At first sight, yes, consumption appears to be reduced
  - It turns out that bad weather correlates with weekends in the observed period
  - With the weekend effect removed, no significant consumption change due to extreme weather events
- Influence of major sports events on consumption?
  - Traces for changed behavior patterns days before the sports even could be identified



6.2 CSI – Comprehensive SCADA Intelligence

Operation

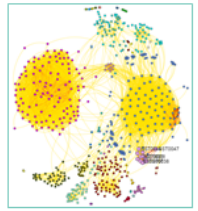
Facebook-style network analysis for SCADA data from electricity grids

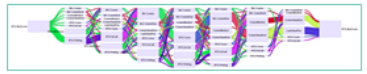
WHERE?
- European DSO
- National distribution grid

WHAT?
- Substation SCADA communication data
- Operational logbook
- Millions of status and error messages
- Communication sequences
- Network structure and relationships

HOW?
- Sequence pattern mining
- Event type and time as link between substations
- Collaborative filter
- Affinity chart

WHY?
- New insight into LV/MV grid correlations
- Strong path of regular communication patterns
- Large substations form a small, distinct group
- Two large groups of ordinary substations
- Several small groups of ordinary substations
- One small group acts as "incident transmitter" to the two big groups





6.5 GTF – Grid Topology Fingerprinting

Operations

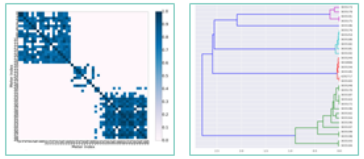
Derive and verify grid topology using data-driven fingerprint identification

WHERE?
- European research project
- European DSO

WHAT?
- Topology identification for low voltage networks
  - Network level
  - Feeder level
- Mapping of smart meters using their unique communication patterns or voltage asymmetry patterns
- "Fingerprint (or DNA) identification"

HOW?
- Correlation coefficient matrices
- Hierarchical clustering, dendrograms

WHY?
- Changes in static grid topology over time are often not well documented
- De-facto versus assumed grid topology can thus differ quite a lot over time
- Deduce actual topology using data-driven methods and check/update network plan



6.3 PnOI – Planning and Outage Intelligence

Operations

Superior predicting outages for smart maintenance and grid control planning

WHERE?
- European DSO

WHAT?
- Focus on outages in the grid: Which assets are involved in outages and what factors are most influential?
- Client's aim:
  - Combine all relevant internal & external data
  - Build a reliable statistical outage assessment model
  - Develop a fully-automated prediction model indicating the risk-status of assets and sections of the grid

HOW?
- Data integration: Multidimensional data enrichment (assets, maintenance, weather, GIS, elevation, SCADA, customer, etc.)
- Hadoop environment for large-scale data processing
- 2-stage logistic regression model for outage risk prediction

WHY?
- Help to reduce outages without increasing maintenance budget by focusing on highest-risk assets and grid sections
- Extensive data integration for in-depth analysis
- Reliable model with >80% prediction accuracy
- List of critical assets for maintenance prioritization

5.3 ESA – Enhanced Sensor Analytics

Sensors

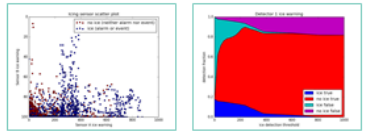
Comparative analysis of icing sensors using entirely different approaches

WHERE?
- European Utility
- Wind power plant in extreme weather conditions

WHAT?
- Initial field study on ice detection
- Two sensors using different technology
  - Sensor A: Change of vibration characteristics
  - Sensor B: Temperature measurement of heating cycles
- Non-linear scaling of icing indicator value completely different:
  - Sensor A: 0 to ~1000 (open end)
  - Sensor B: 100 to 0

HOW?
- Scatter plots
- Box plots
- True/False positive/negative dynamic threshold categorization

WHY?
- Due to harsh weather conditions, blade icing is the most serious impact on the performance of this wind park
- For better information on icing, two types of sensors were installed on one wind mill for testing
- In the future only one of these sensors shall be rolled our
- The gained insight will help to improve preventative shut down of turbines in case of icing
- Facilitates evaluation of blade heating installation business case



Planning and Outage Intelligence

Planning Intelligence – Scope

Baseline

Data Integration: Combine the large amount of diverse data

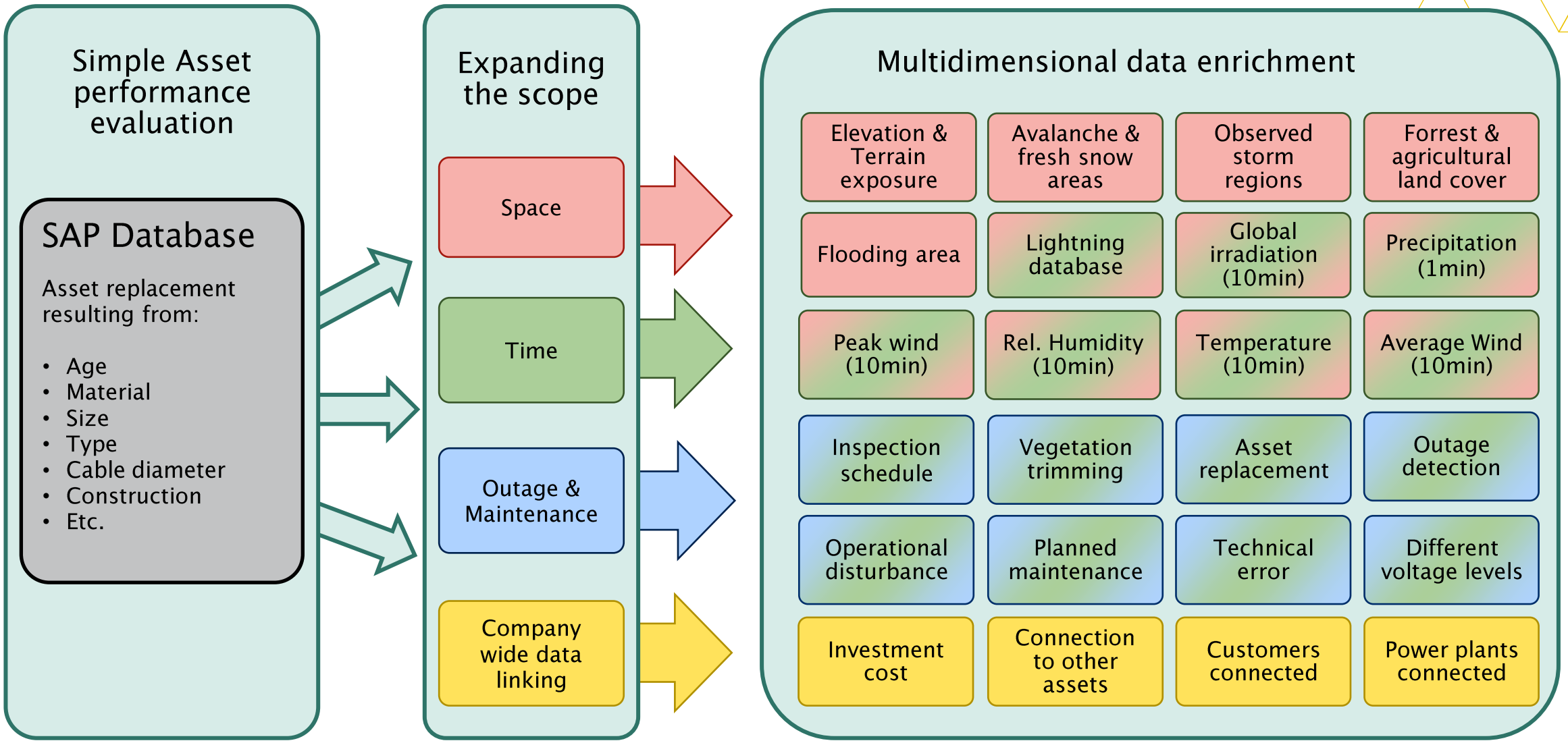
- ❑ Operational data (such as asset condition, inspections, outages history etc.)
- ❑ External data (environment, GIS, weather, etc.)

in order to get new insights for planning and operational decisions.

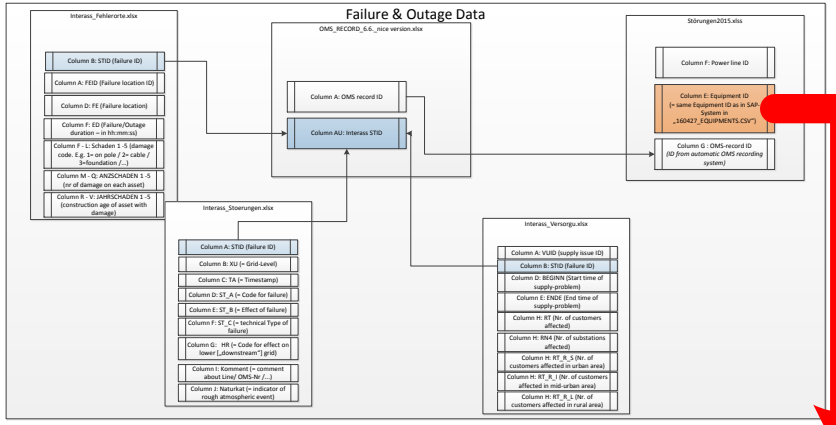
Use case

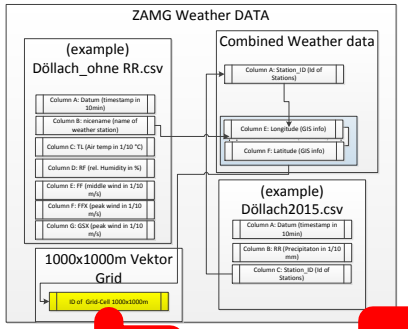
- Which assets are at risk and need to be maintained first?
- How can we optimize maintenance with given budget?
- May we have a day-ahead warning based on weather forecasts?

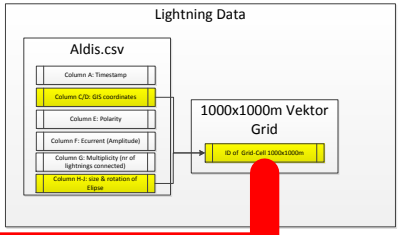
Planning Intelligence – Data Enrichment



Planning Intelligence – Example Of Data Modelling

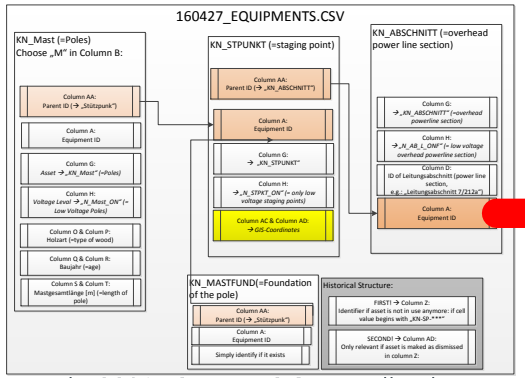


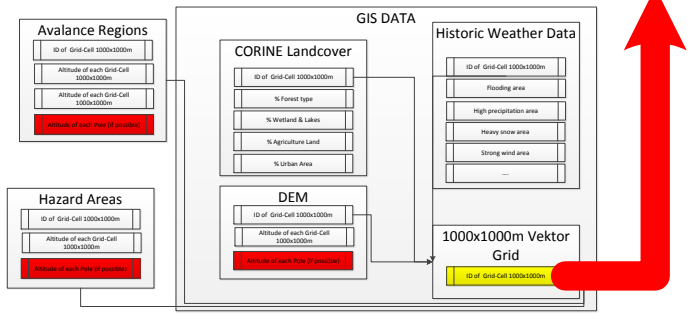


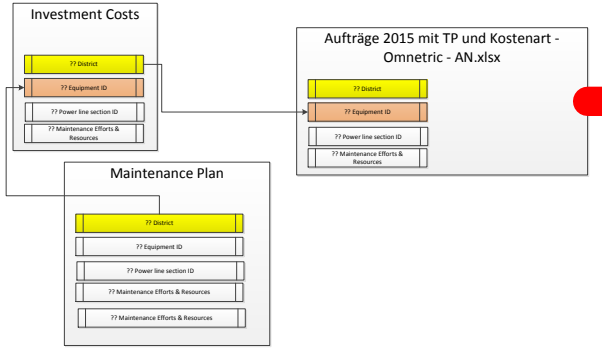


„BIG Table“

TIME	IDs			Poles (from SAP)					Failure & Outages							GIS Data					Weather data				Lightning Data		Maintenance Data		
Timestamp	*Equipment ID (N_Mast_ON)*	*Equipment ID(N_AB_L_ONF)*	*power line section (N_AB_L_ONF)*	*Holzart*	*Baujahr*	*Mastgesamtl änge*	*Foundation (KN_MASTFUND)*	*Coordinates (KN_STPUNKT)*	*STID*	*OMS Record*	*XU*	*TA*	*ST_A*	*RT*	...	*GIS 1000x1000m Grid Cell*	*Share Mixed forest*	*Hazard flooding area*	*Altitude*	*District ID*	...	*GIS 1000x1000m Grid Cell*	*Timestamp*	*Precipitation*	*Wind*	...	*GIS 1000x1000m Grid Cell*	*Lightnig amplitude*	*District ID*	*Equipment ID (N_Mast_ON)*	*Financial effort*
(2015-2016 in ~20min	ID of the pole	ID of the power line section	Identifier for power line section	Type of wood	Building age	Length of the pole	Does Pole have a foundation?	GIS coordinates of the Pole / staging point	Outage ID	ID from OMS System	Voltage Grid Level	Time stamp (Interass_störungen. xlsx)	Failure Code	Nr of customers affected	...	identifier for GIS-Data	from CORIN E in % of cell	historic data on hazard/	altitude of the asset DEM	district belonging	...	weather station mapped with cell-ID	time aggregated to 20min	aggregated to 20min	average of these 20min	...	weather station mapped cell-ID	provide amplitude	district belonging	ID of the pole	-	*Additional Data*
23.june2014, 14:20	0000111	1100012	7/719a	F	1994	10	1	1.4678E+16 : 4.6319E+16	-	-	-	-	-	-	...	33	54	1	1012	18	...	33	23.june2014, 14:20	8	23	...	33	0	18	0000111	20000	...
23.june2014, 14:20	0000112	1100012	7/719a	F	1994	11	1	1.4678E+16 : 4.6319E+16	-	-	-	-	-	-	...	33	54	1	1004	18	...	33	23.june2014, 14:20	8	23	...	33	0	18	0000112	20000	...
23.june2014, 14:20	0000113	1100012	7/719a	K	1994	10	0	1.4678E+16 : 4.6319E+16	-	-	-	-	-	-	...	34	13	0	995	18	...	34	23.june2014, 14:20	8	23	...	34	0	18	0000113	20000	...
23.june2014, 14:20	0000114	1100012	7/719a	K	1994	12	0	1.4678E+16 : 4.6319E+16	100012	AH00046743	9	23.june2014, 14:20	20	0	...	34	13	0	993	18	...	34	23.june2014, 14:20	8	23	...	34	21	18	0000114	20000	...
23.june2014, 14:20	0000115	1100012	7/719a	K	1994	10	0	1.4678E+16 : 4.6319E+16	-	-	-	-	-	-	...	35	24	0	1001	18	...	35	23.june2014, 14:20	8	23	...	35	0	18	0000115	20000	...
23.june2014, 14:20	0000116	1100231	6/20/13cb	K	1987	8	1	1.4682E+16 : 4.6241E+16	-	-	-	-	-	-	...	47	32	0	998	19	...	47	23.june2014, 14:20	8	23	...	47	0	19	0000116	20000	...
23.june2014, 14:20	0000117	1100231	6/20/13cb	L	1987	8	1	1.4682E+16 : 4.6241E+16	-	-	-	-	-	-	...	48	65	1	989	19	...	48	23.june2014, 14:20	8	23	...	48	0	19	0000117	20000	...
23.june2014, 14:20	0000118	1100231	6/20/13cb	L	1987	9	1	1.4682E+16 : 4.6241E+16	100013	AH00046743	9	23.june2014, 14:20	22	23	...	48	65	1	988	19	...	48	23.june2014, 14:20	8	23	...	48	39	19	0000118	20000	...
23.june2014, 14:20	0000119	1100231	6/20/13cb	L	1987	10	1	1.4682E+16 : 4.6241E+16	-	-	-	-	-	-	...	48	65	1	984	19	...	48	23.june2014, 14:20	8	23	...	48	0	19	0000119	20000	...

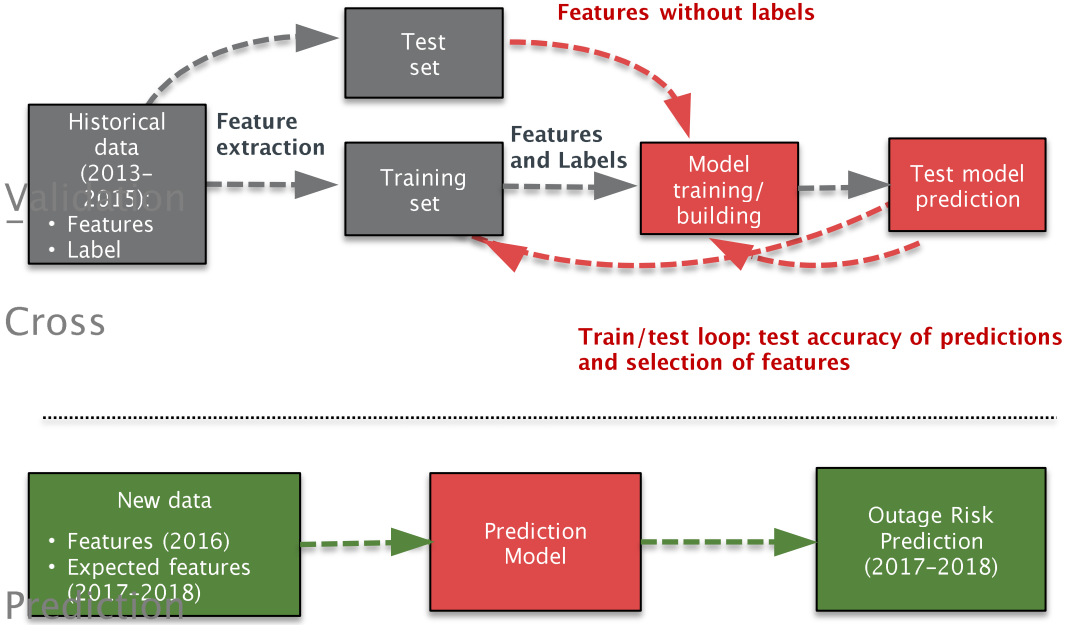




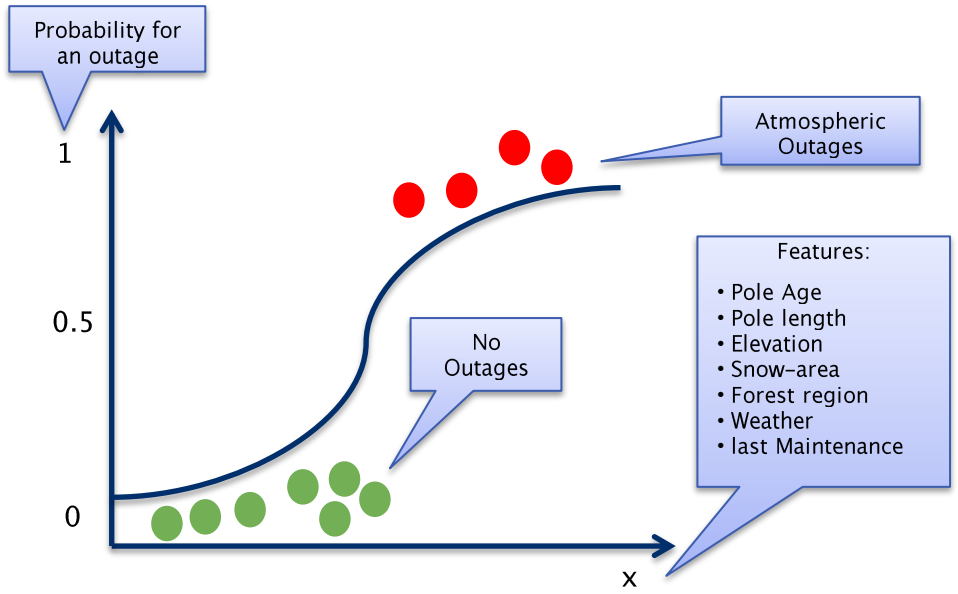


Applying machine learning methods

Machine Learning Process



Binary decision: Logistic Regression



With a **Logistic Regression** we measure the relationship between the Y "Label" (= Outage yes/no) and the X "Feature-Set" by estimating probabilities using a logistic function, which is then used to predict the label class.

Simulate potential impact of maintenance-efforts:

(1) Include options to decrease disturbance risk:

What if: median age is 10 years instead of 25 years ?

$$f(disturbance) = ß0 + ß1 * year_{dummy} + ß2 * \boldsymbol{pole_{charact}} + ß3 * environment_{data} + ß4 * \boldsymbol{Vegtrim} + ß5 * \boldsymbol{inspection}$$

(2) Estimate 'roughly' costs of maintenance actions:

What will be the costs associated with a potential vegetation trimming?

$$f(vegetation_{trim}) = ß0 + ß1 * year_{dummy} + ß2 * share_{forest} + ß3 * altitude + ß4 * last_{Vegtrim} + ß5 * soil_{type} + \ ...$$

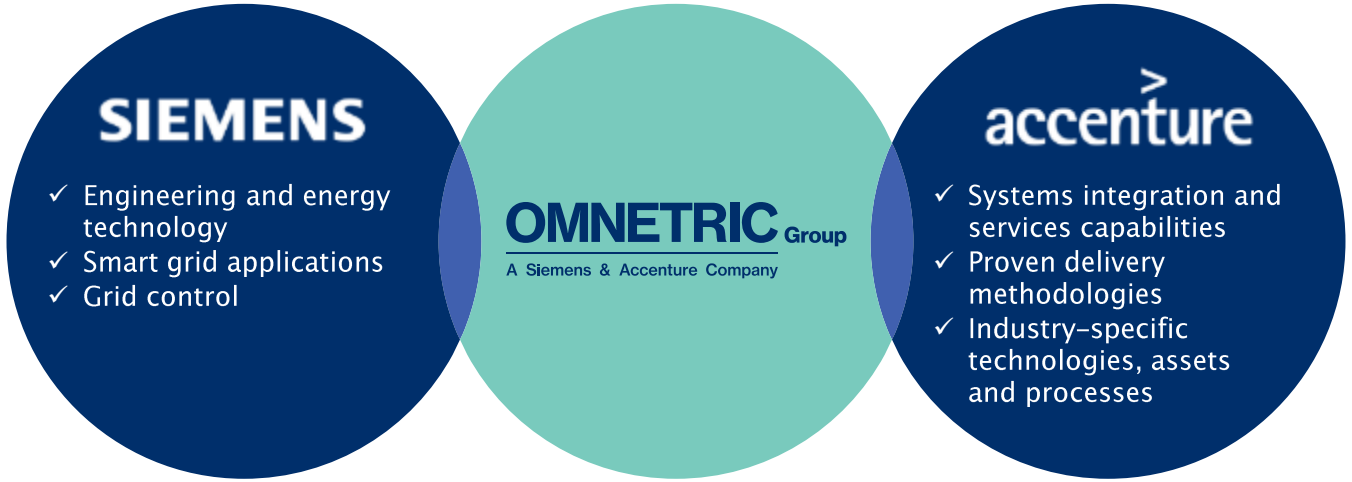
Scenarios

- Analyze the Impact of the maintenance effort on the disturbance risk
- Predict Associated Costs of maintenance actions or modification of the assets
- Select powerline sections for maintenance by a given budget or by specific risk-threshold

Demo Planning and Outage Intelligence

Thank you for your attention and please feel free contact us

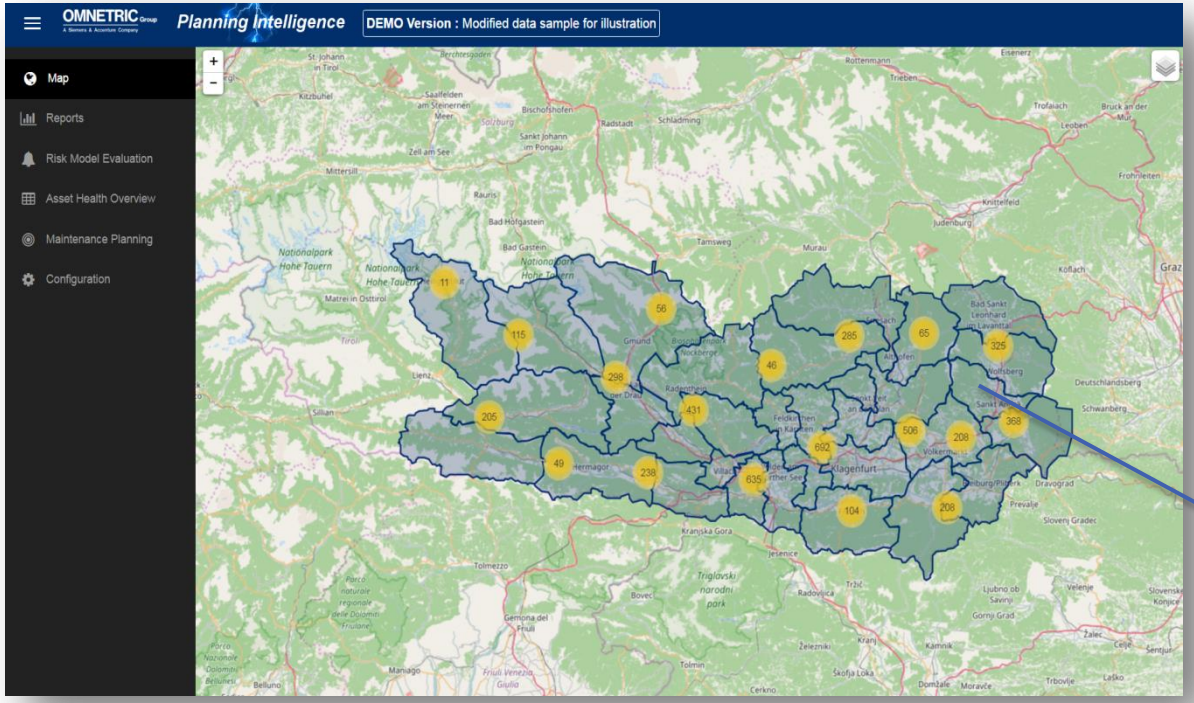
OMNETRIC Group is dedicated to the global delivery of integrated information technology and operational technology solutions and services, helping utility companies to achieve greater grid reliability and efficiency. OMNETRIC Group can support clients with innovative solutions wherever they may be on their path to a smarter grid. For more information, visit www.omnetricgroup.com



Demo Backup

Map: geographical Interface

Using the map allows having a quick overview on asset



Zoom in & click on assets (poles) to see:

- Characteristics (age, Material, etc.)
- Environment & exact location (Streetmap, Satellite, etc.)
- In Progress: search field to quickly find relevant assets

Customized Reports

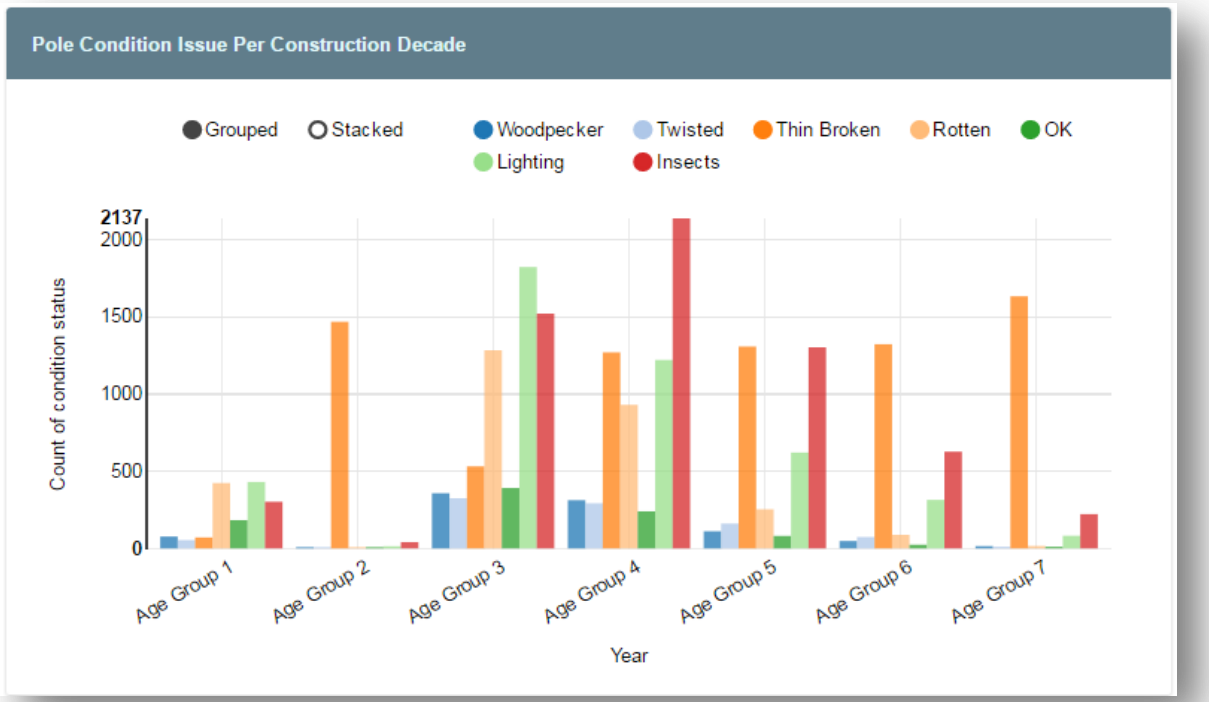
Reports are created on customer's demand

What do you need to know for your decisions?

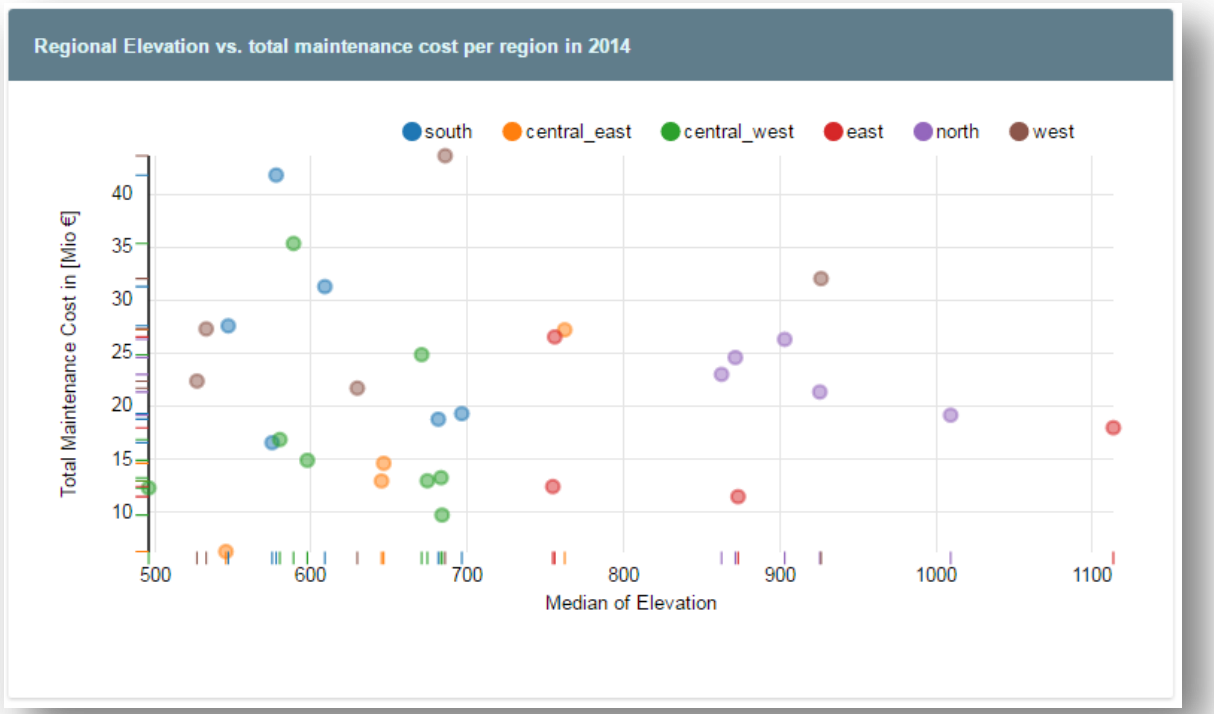
- Easily combining information of different data sources
- Receive descriptive statistics
- Understand nexus between different data, such as maintenance costs and construction year of feeder
- Applying interactive charts for high usability

Customized Reports

Examples of report-charts



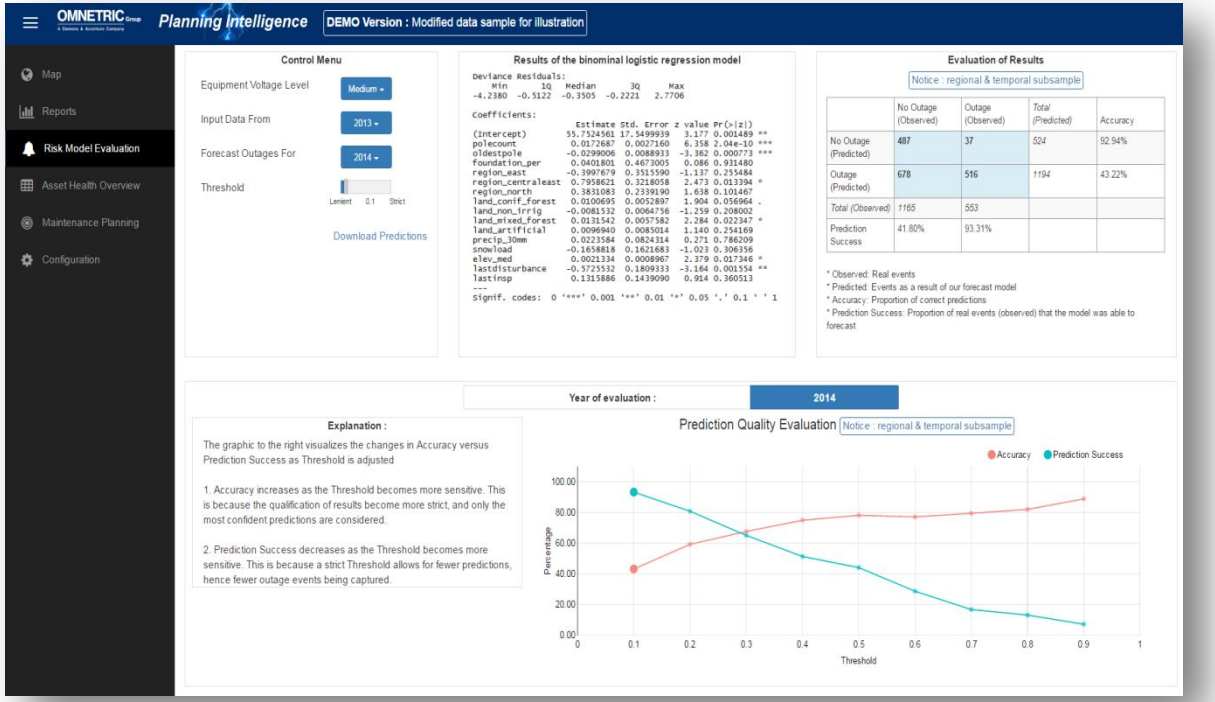
Comparing pole condition (affected by woodpecker, insects, being partly twisted, etc.) and the age-group (anonymized)



Comparing maintenance expenditures of grid segments (in Mio €) and elevation. Colored dots show the geographical location.

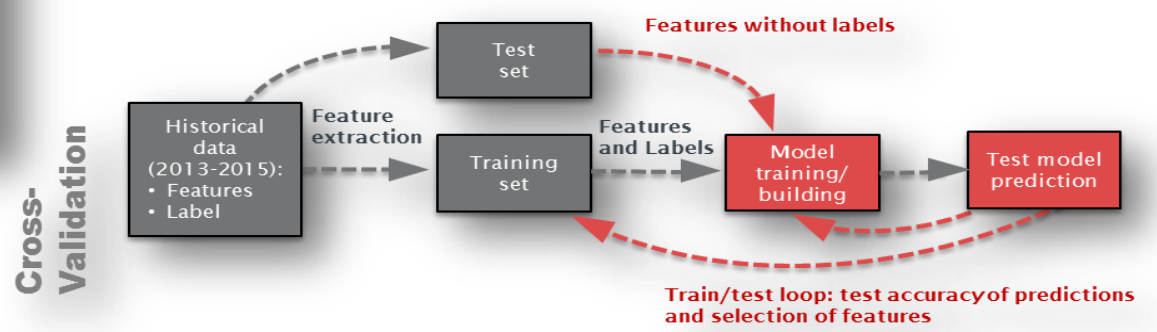
Risk Model Evaluation

The evaluation of the model helps the user to get a feeling on the reliability of the results



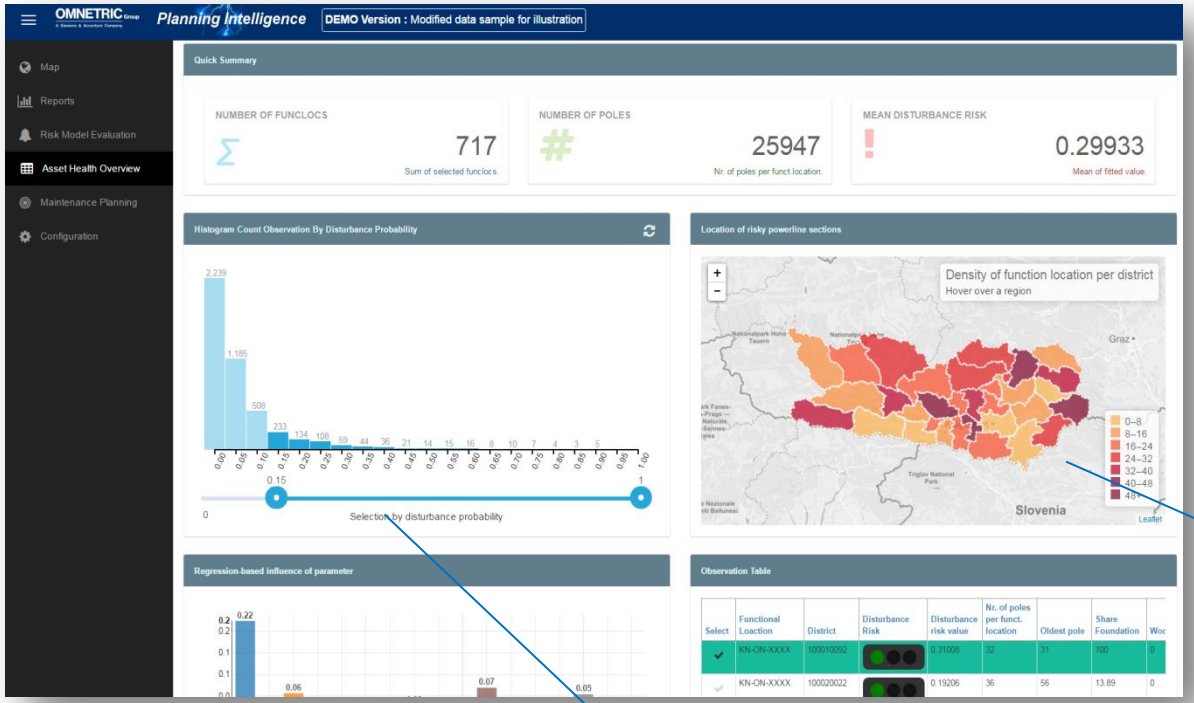
How does the evaluation work?

- Divide data into TEST and TRAINING set
- TRAINING: e.g. data from 2013 → used to create a machine learning model
- TEST: e.g. data from 2014 → used to evaluate the quality of the result from the machine learning model
- Comparing different TEST and TRAINING data give the user a feeling on the reliability of the predictions



Asset Health Overview

Learn about your assets' health status and impact of different parameters



Asset health based on outage probability

- Sub-select feeders depending on the outage risk (blue histogram)
- Quantify assets being at risk (e.g. 717 feeders / 25.947 Poles with a mean outage probability of 0.29933 percent)
- Identify critical regions: consider the map to see which (e.g. political) districts are mainly affected by assets in poor condition

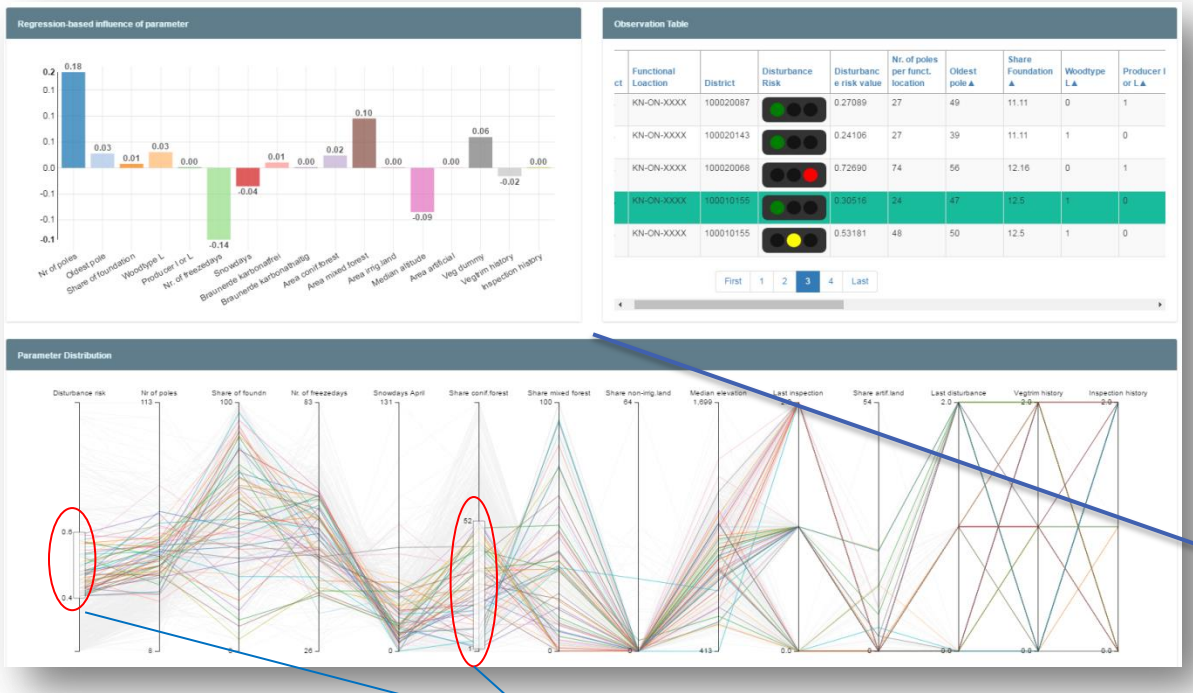
Map, summary of selected feeders or charts below are responsive and adapt automatically to the selection

Use the slider in order to sub-select feeders depending on the outage probability

Asset Health Overview

Planning Intelligence

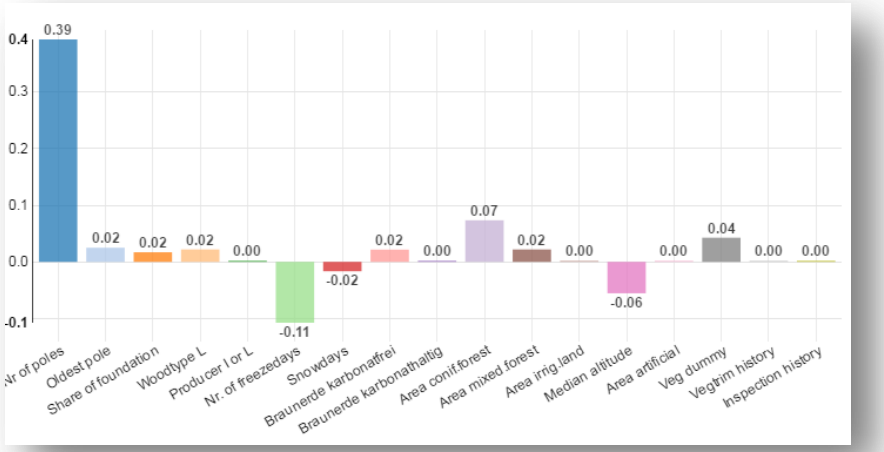
Interactively dive into the data



Select feeders depending on the parameters

Get an insight, why each feeder is exposed to risk

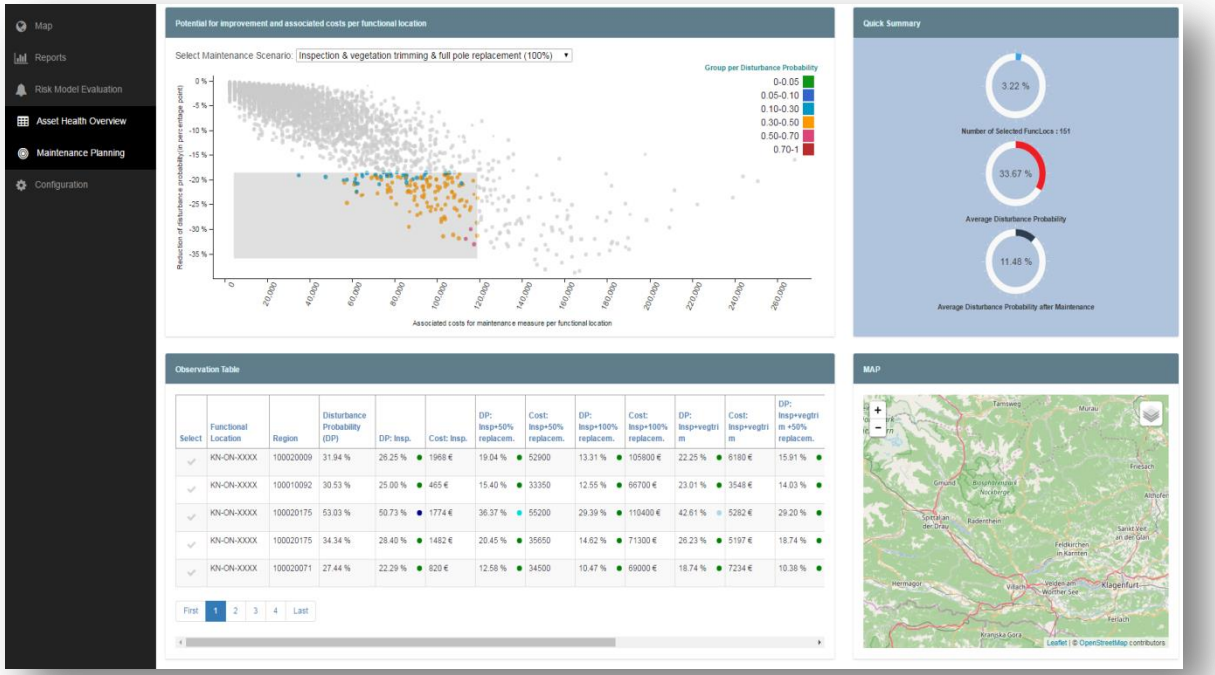
- Consider table with “traffic lights” with detailed information of each feeder and if they are in critical status
- Discover in the bar chart:
  - Is there a parameter influential which can be changed, e.g maintenance?
  - Are environmental parameter (e.g. elevation) responsible?
- Explore the parameter distribution: each line in the chart on the bottom represents a feeder – filter and learn about interactions



Maintenance Planning

Planning Intelligence

Chose most cost-effective maintenance measurements for each feeder



For decision making:

- Compare improvement potential and associated costs for each individual feeder
- Compare different scenarios and get an understanding on the impact of different maintenance measurements
- Choose between different maintenance measurements with feeder-specific costs and improvement potential
- Consider the geographical location of feeders in your decisions

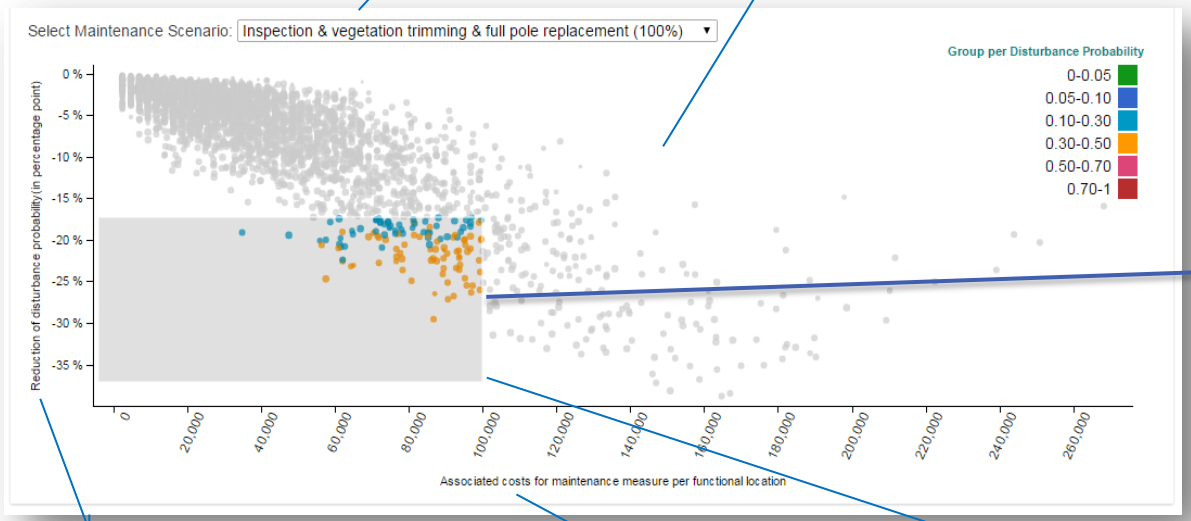
Maintenance Planning

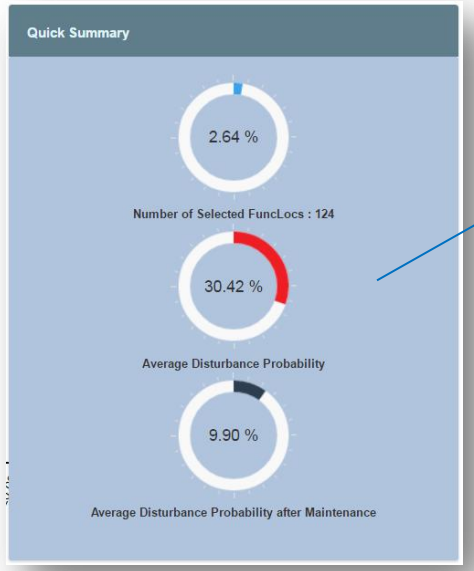
Planning Intelligence

Select depending on costs and improvement potential

Choose a scenario

Each dot is a feeder
(color = outage probability)





Understand how you can improve the selected feeders: Reduce outage probability from 30.42% to 9.90% with the selected scenario

Reduction of outage probability in % points

Associated costs for each feeder

Interactively select interesting feeder, such as:
→ improvement potential >12%
→ points and cost > 100.000

Maintenance Planning

Planning Intelligence

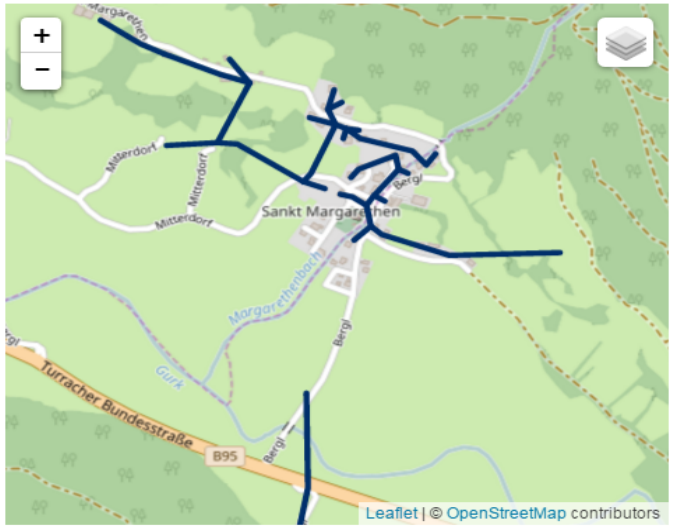
Go through the list of feeder and select a subgroup for maintenance

Observation Table

Select	Functional Location	Region	Disturbance Probability (DP)	DP: Insp.	Cost: Insp.	DP: Insp+50% replacem.	Cost: Insp+50% replacem.	DP: Insp+100% replacem.	Cost: Insp+100% replacem.	DP: Insp+vegtrim	Cost: Insp+vm
✓	KN-ON-XXXX	100020087	36.16 %	30.04 %	1004 €	18.60 %	42550	16.05 %	85100 €	23.65 %	3969 €
✓	KN-ON-XXXX	100020087	26.06 %	21.09 %	1203 €	15.24 %	31050	11.37 %	62100 €	16.16 %	8806 €
✓	KN-ON-XXXX	100020028	32.93 %	27.13 %	987 €	18.20 %	48300	11.94 %	96600 €	27.13 %	0 €
✓	KN-ON-XXXX	100010114	28.01 %	22.78 %	1064 €	17.39 %	43700	10.50 %	87400 €	20.92 %	7240 €
✓	KN-ON-XXXX	100010114	26.85 %	21.78 %	865 €	11.37 %	28750	8.62 %	57500 €	16.72 %	3635 €

First | 7 | 8 | 9 | 10 | Last

MAP



Select specific feeder which are then illustrated on the map

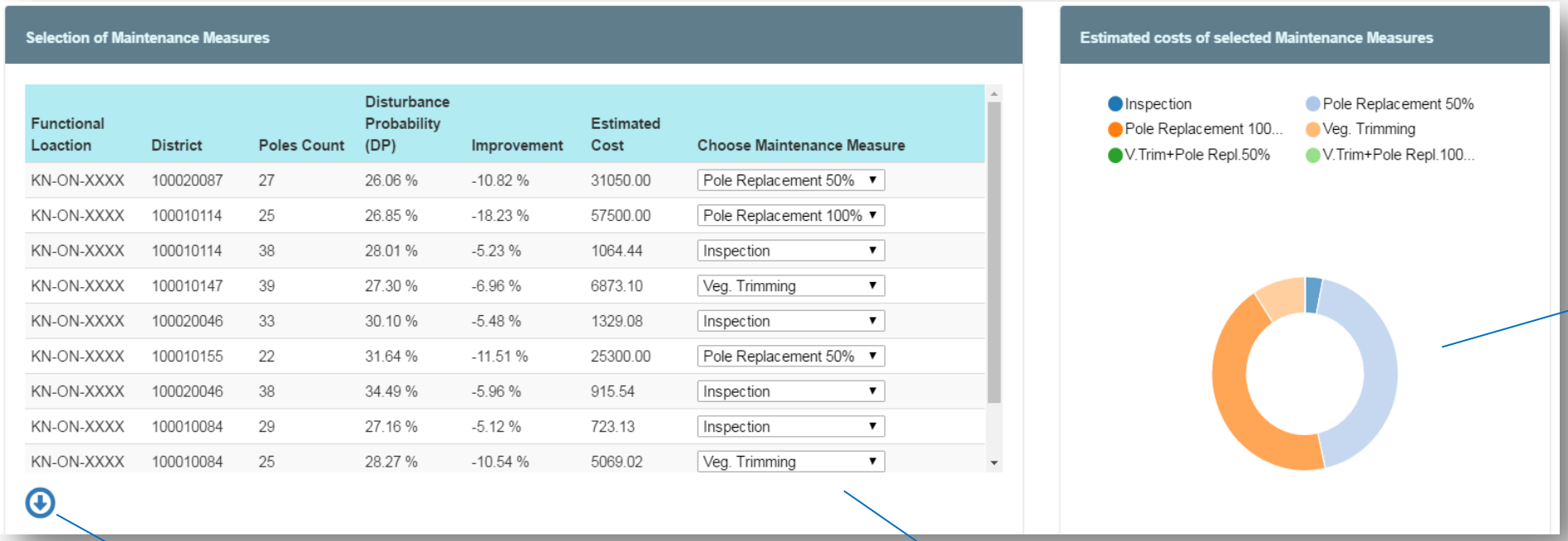
Compare different maintenance measurement + associated cost:

"Insp+50% replacem."
→ Inspection
→ replacement of 50% of the poles of the feeder

Maintenance Planning

Planning Intelligence

Define maintenance measurements and see total costs



Get a summary of all costs with respect to different maintenance measurements

Download the selection

Select the most appropriate maintenance measurement for each feeder depending on:
→ Disturbance (Outage) probability
→ Improvement potential
→ Estimated costs