TROIA

Efficient electricity distribution and sustainable energy management through Big Data analytics and machine learning

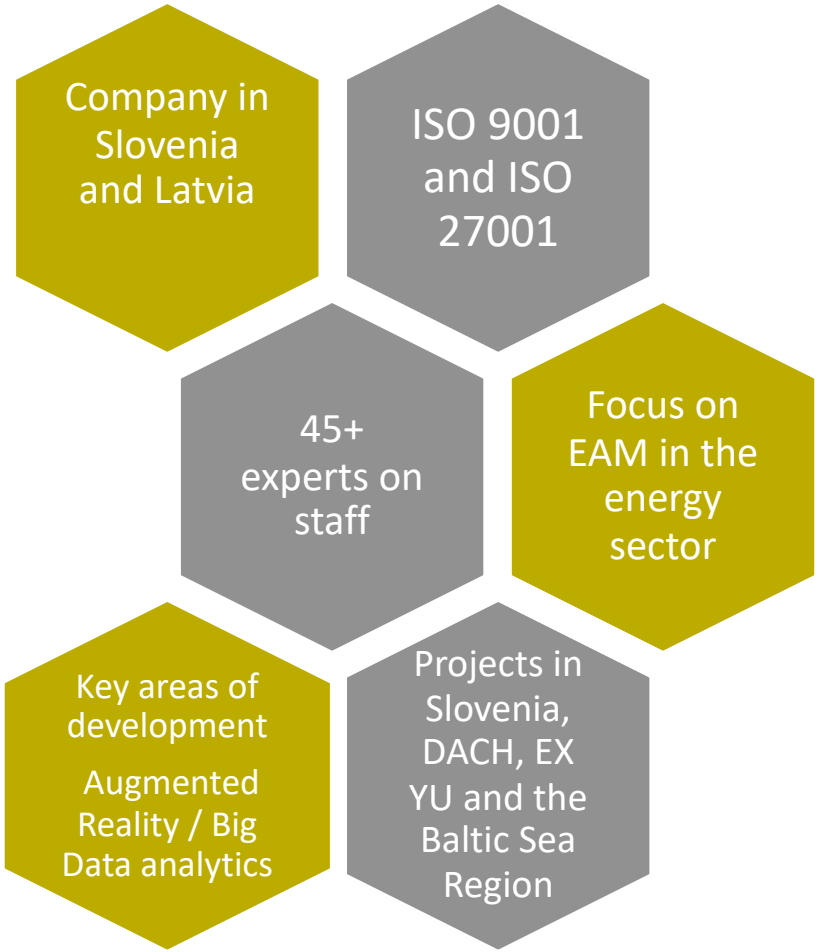
dr. Andrej Somrak

ENERDAY 2023
5. 5. 2023

AGENDA

- INTRODUCTION
- BIG DATA/ANALYTICS PLATFORM
- DATA VISUALIZATION USE CASES
- CHALLENGES IN IMPLEMENTING SOLUTION
- AI/ML
- DISQUSSION & QUESTIONS

- Company founded 2010
- The core team has 20+ years of experience in the energy industry
- Our customers in Energy sector: distribution and transmission companies



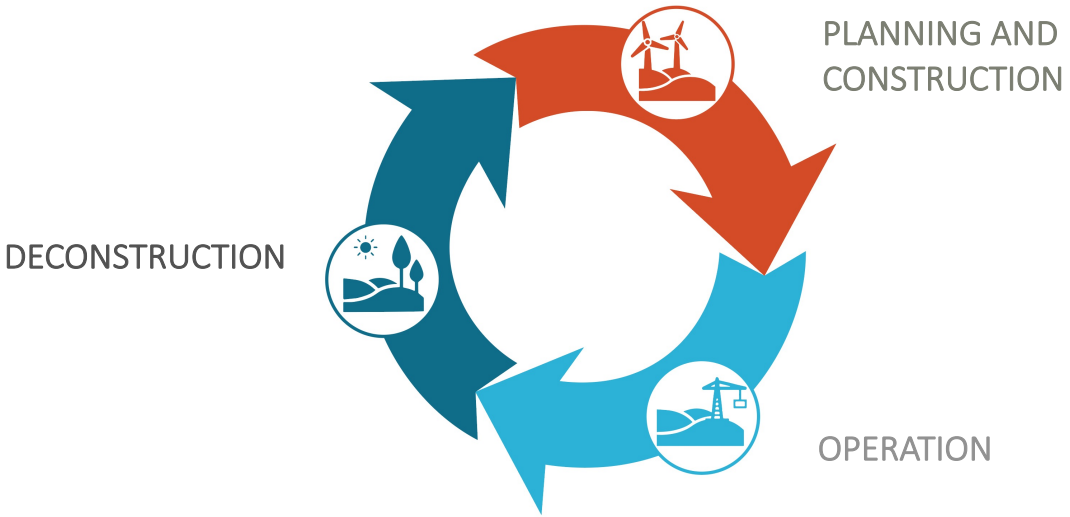
DACH-Partner: RODIAS

For more than 35 years, we have been **supporting utilities in the digitalisation** of their processes. In the **nuclear industry**, RODIAS has been setting the standard for **plant management systems** in the German-speaking market for several decades.

Throughout the last decade, we **entered** other industries like **mobility&transport**, **pulp&paper**, **facility management** and **automotives**.

For the **future**, we are positioning ourselves in the **hydrogen market**.

Our future-oriented solutions enrich **Enterprise Asset Management systems with big data, analytics and machine learning use cases**.



IMPLEMENTING BIG DATA SOLUTION/PLATFORM

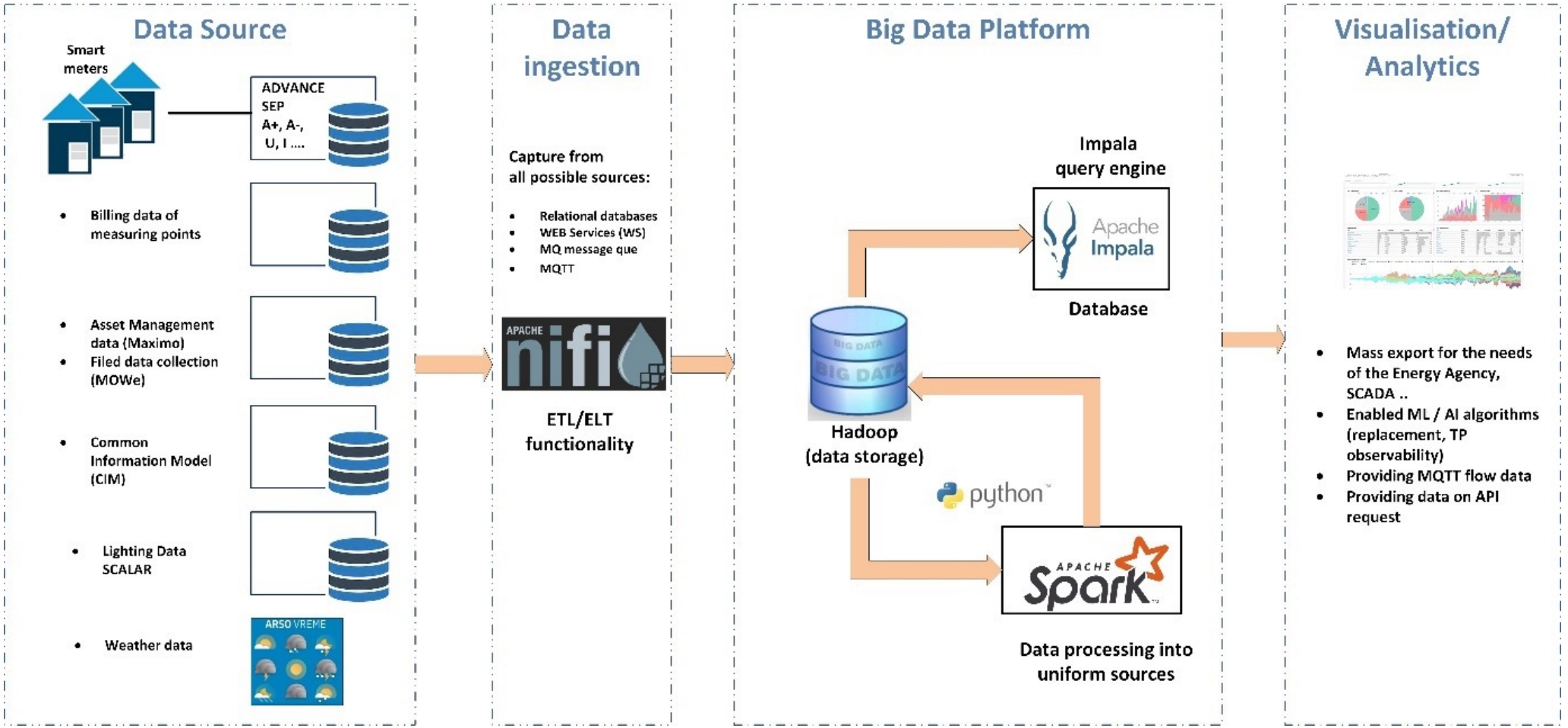
- **Digitalisation**:
  - Data-driven organisations
  - Generating huge amounts of data (IoT devices/smart meters)
  - Siloed data

- **The volume and quantity of data** makes it impossible to manage data in the traditional way (storage/processing)
  - ➢ Implementing big data solutions

- Network management, optimisation and development

Accurate and timely data and information is the cornerstone of optimising energy use and provides many untapped opportunities to reduce energy use, thereby lowering our carbon footprint and making our operations even more sustainable.

BIG DATA/ANALYTICS PLATFORM – ATHENA ANALYTICS



BIG DATA/ANALYTICS PLATFORM – ATHENA ANALYTICS

- **Open Source** Solutions (Cloudera)
- **Central data platform**
- It provides:
  - Security
  - Robustness
  - High mass data processing capability
  - Scalability:
    - Data volume
    - Scope of data sources

- **Business Analytics** (Data Analytics, Visualisations, Dashboards, Reports)
  - Power BI, Tableau, Grafana, Superset, QGis, Neo4j
  - Ad Hoc Queries (Hue)

- **Advanced Analytics** (AI/ML/Data mining)

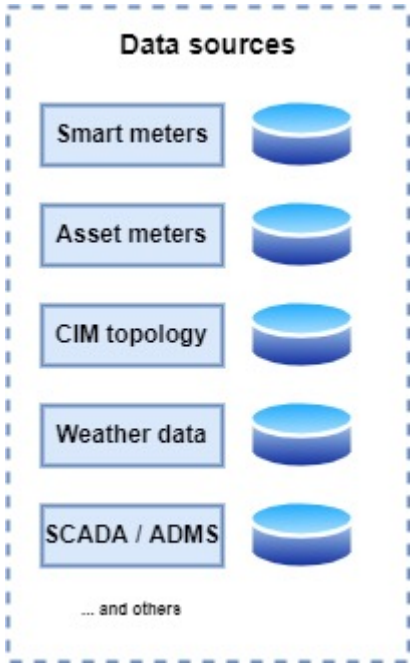
- **Data services**
  - on-demand access to data by external stakeholders (REST API)

Capture data from different sources:

- Flow data (sensors, iOT, smart meters)
- Relational databases
- Online sources

- **Data processing/preparation/ETL/ELT**
  - Cleaning and editing
  - Standardisation, segmentation, aggregation
  - Deduplication, imputation
  - Transformation, substitution, conversion
- **Implementation options**:
  - On premise/private cloud
  - Public Cloud/Hybrid Cloud

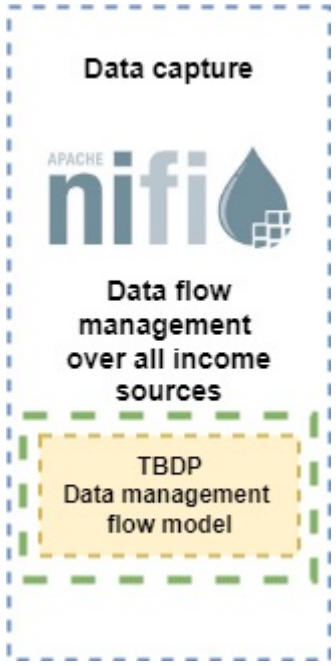
BIG DATA/ANALYTICS PLATFORM – ATHENA ANALYTICS



Support various DATA SOURCE TYPE

Support different DATA PROTOCOLS

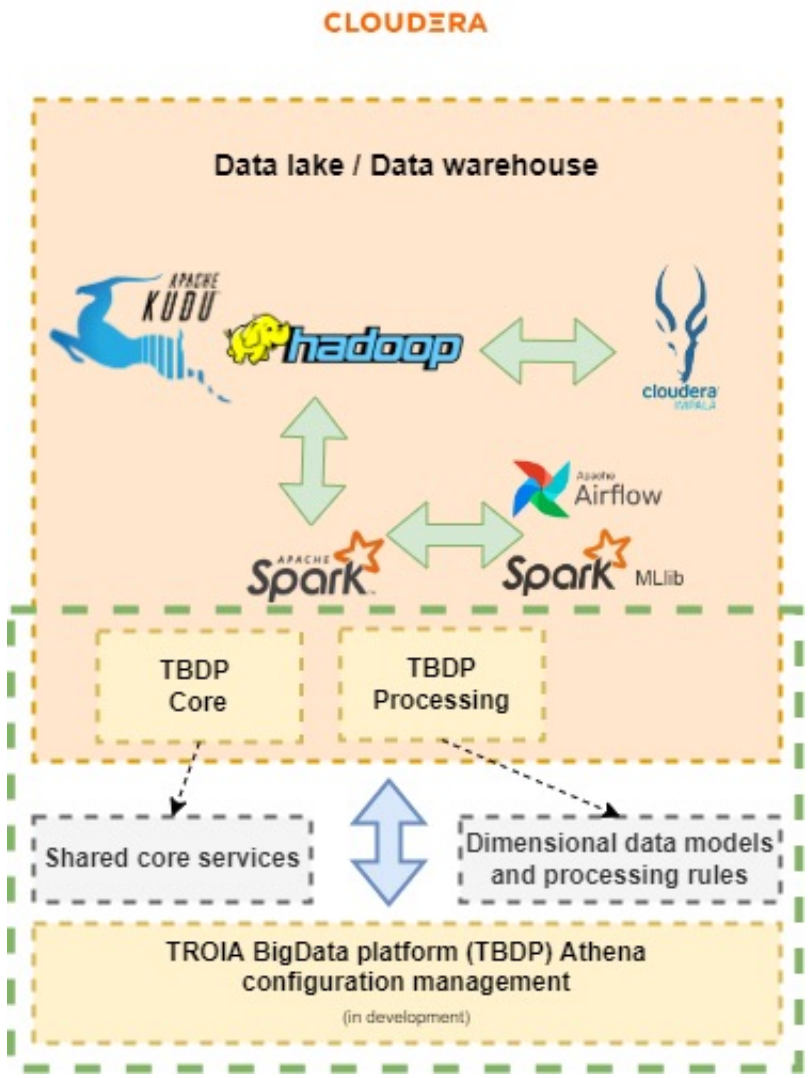
Quick data adoption due to strong domain knowledge



Graphic controlled data management flow model

Versioning control for data flow model

Support for wide range of data sources and technologies behind



ATHENA ANALYTICS Core

- Shared utility functionalities
- Core data libraries
- Supports many SCD types
- Resilient by design
- Highly configurable

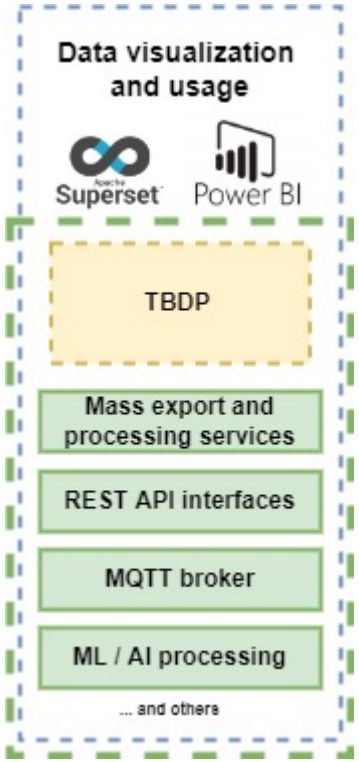
ATHENA ANALYTICS Processing

- Business logic
- Dimensional data models (Kimball DWH approach)
- Processing rules
- Custom environmental variables

ATHENA ANALYTICS configurations (in development)

- Intuitive web interfaces for configuration management
- Graphical view of data model relations
- Other platform configurations
- Deployments

Data visualization / Data export services



Pre-developed and configured client dashboards

Custom developed and configured client dashboards

Quick configuration and adoption

REST API services on demand

Mass export services

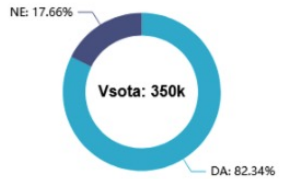
ML model execution

USE CASES

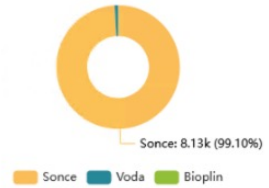
PLC/Smart metering communication	Transformer overload observability	Common area operating picture
Grid losses	Load profile	Overview of maintenance situation
Events at measurement sites	Measuring site load and power consumption	General grid KPI visualisation
Events at measuring devices	Measuring site voltage quality	Digital twins for assets

STATE OF THE SYSTEM

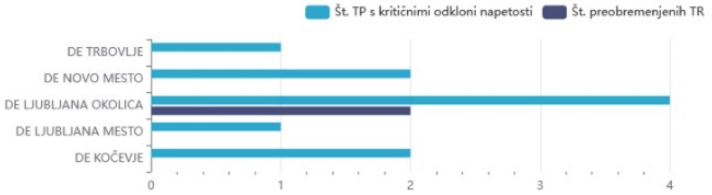
Daljinsko odčitavanje števcev



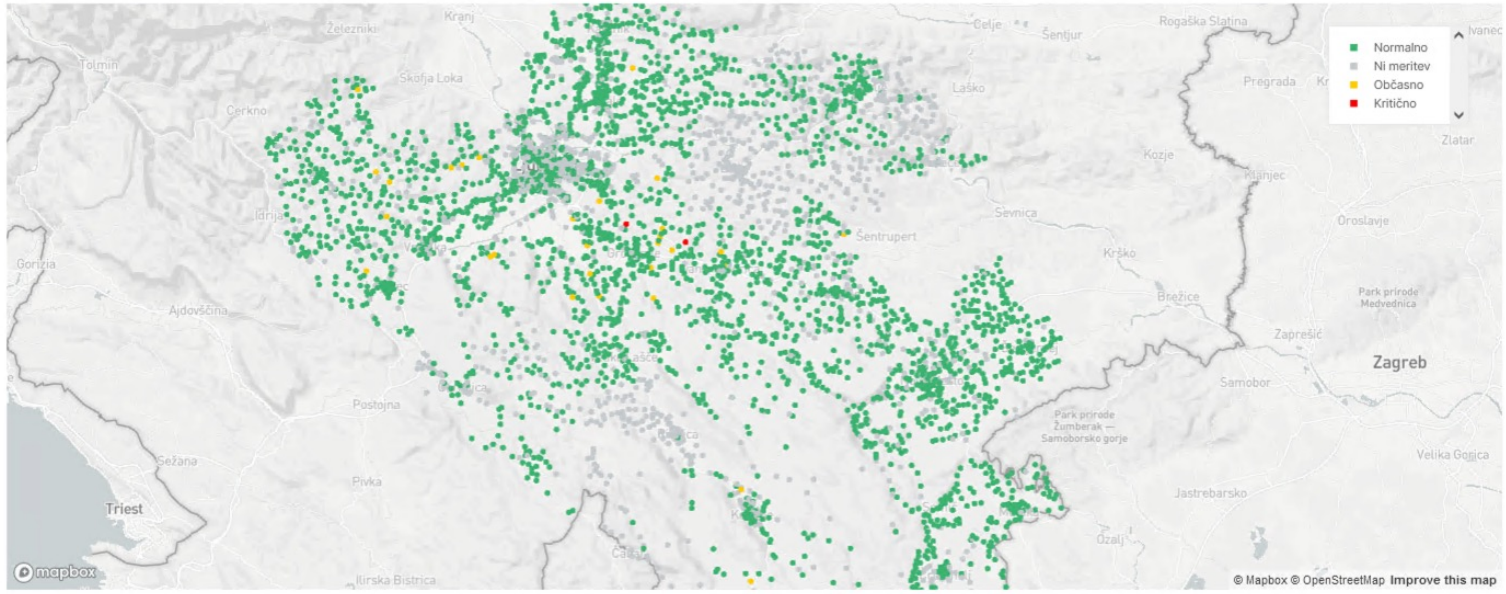
Vir energije proizvodnih naprav



KPI TP



Transformatorske postaje



TRANSFORMER STATION LOAD

Specifikacije TP in hierarhija priklopa

TP	Naziv TP	Tip TP	S_projekt. (kVA)	Leto izgradnje	SN-izvod	RTP/RP	Nadzorništvo	Seznam MM
2000608	VRH PRI LESKOVCU 20/0.4 G-389	JAMBORSKA ŽELEZNA	50	1984	J28 DV 20KV POLICA	RTP 110/20 KV GROSUPLJE	GROSUPLJE 1	Odpri

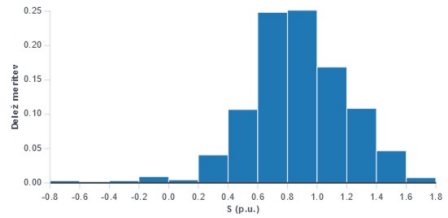
Specifikacije transformatorjev (podatki o nameščenem sredstvu iz MX)

Šifra tr.	Tip tr.	U_1/U_2 (kV)	S_n (kVA)	I_{2n} (A)	u_k (%)	Izg. krat. stika (W)	Izg. praz. teka (W)	Hladilni medij	Leto izdelave	Obratovalno MM	ID števca
3784148	TR. 4HT 50/20-0.4	20/0,4	50	80.83	4.19	1075	188	OLJE-S	1984	8038950	70630275

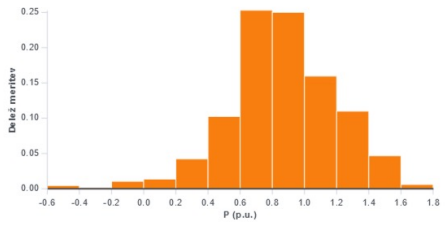
Obremenitev TP / transformatorja

Graf obremenitve (izberite šifro TP / transformatorja / obr. mer. mesta)

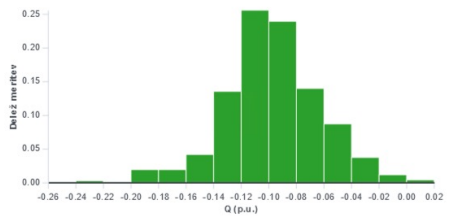
Histogram obremenitev navidezne moči



Histogram obremenitev delovne moči



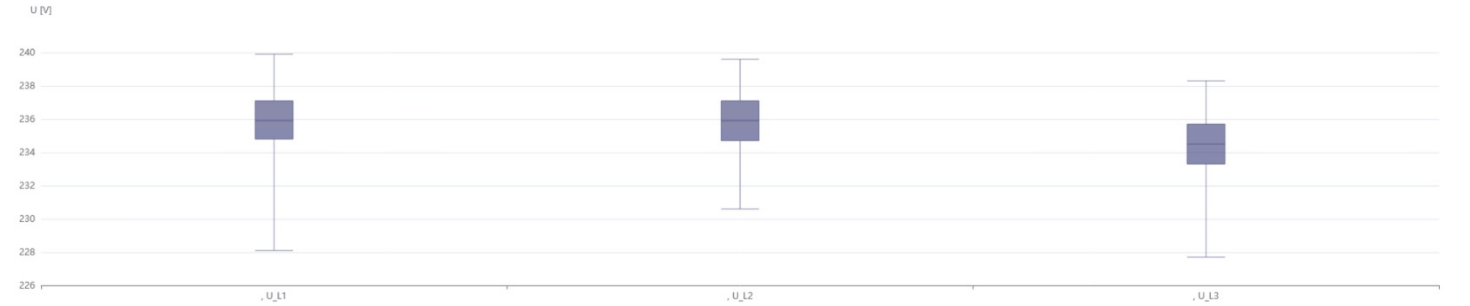
Histogram obremenitev jalove moči



MEASURING SITE VOLTAGE TIME SERIES

Napetostni profil

Graf kvantilov napetostnega profila



Statistični podatki za izbrano časovno obdobje

U_L1 max	U_L2 max	U_L3 max	U_L1 min	U_L2 min	U_L3 min	U_L1 povp.	U_L2 povp.	U_L3 povp.
239.9	239.6	238.3	228.1	230.6	227.7	235.9	235.8	234.5

MEASURING SITE LOAD AND POWER CONSUMPTION

Osnovne informacije

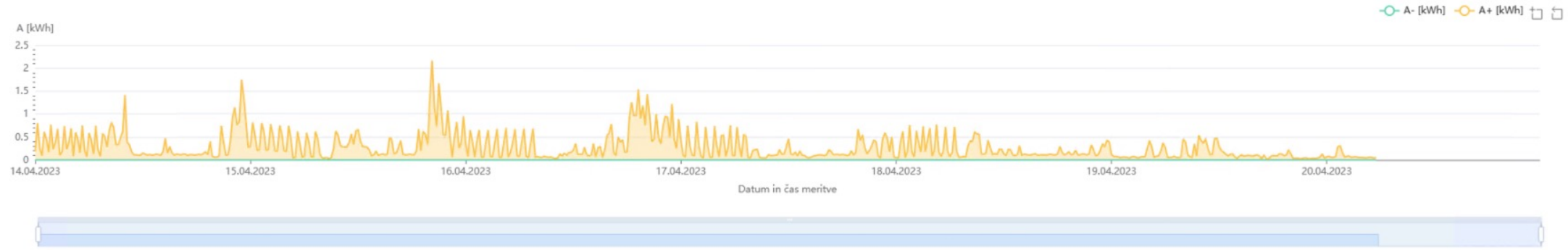
TP	MM	Vrsta MM	Napetostni nivo	Vrsta odjema	Vezalna shema	Tovarniška št.	Proizvajalec	Tip	Daljinsko odčitavanje
TP0454-ŠOLA JARŠE	3102790	OBRAČUNSKO	Nizka napetost	Gospodinjstvo	PS.1A	77260743	ISKRAEMECO	AM550-TD1 PLC	DA

Statistični podatki za izbrano časovno obdobje

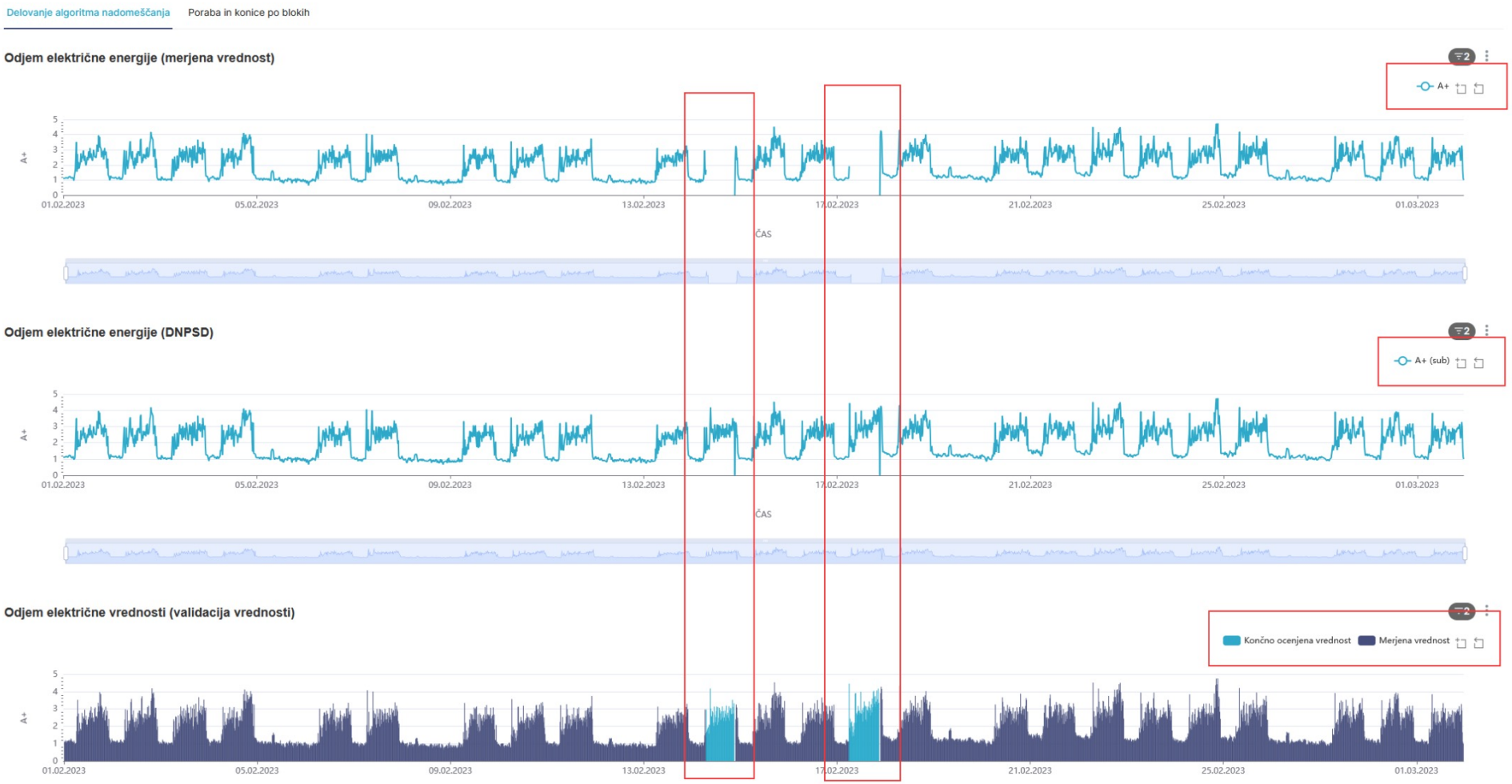
P+_max [kW]	P+_povp. [kW]	A+ [kWh]	P-_max [kW]	P-_povp. [kW]	A- [kWh]
8.7	1.1	165.6	0	0.0	0.0

Dnevna stanja naprave | Prejeta in oddana električna energija | Analiza delovne in jalove moči | Dinamičen tarifni sistem

Prejeta in oddana električna energija



IMPUTATION OF MISSING METER DATA



Dynamic tariff system – change of electrial billing

Osnovne informacije

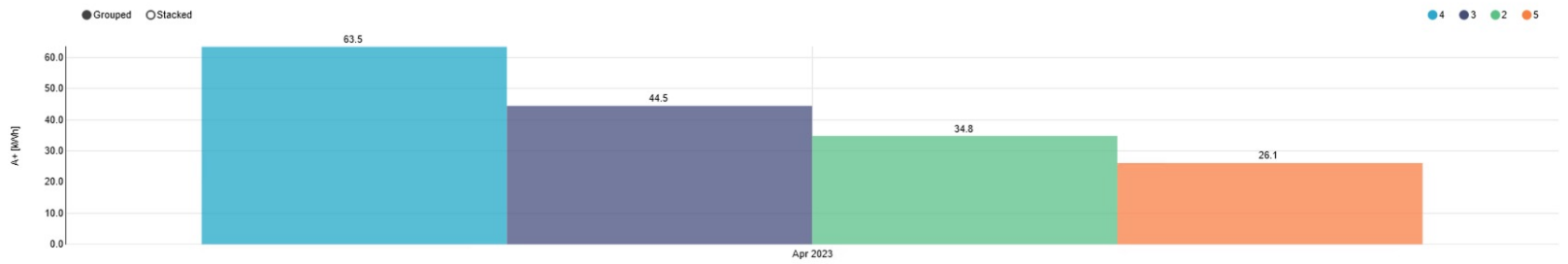
TP	MM	Vrsta MM	Napetostni nivo	Vrsta odjema	Vezalna shema	Tovarniška št.	Proizvajalec	Tip	Daljinsko odčitavanje
TP0454-ŠOLA JARŠE	3102790	OBRAČUNSKO	Nizka napetost	Gospodinjstvo	PS.1A	77260743	ISKRAEMECO	AM550-TD1 PLC	DA

Statistični podatki za izbrano časovno obdobje

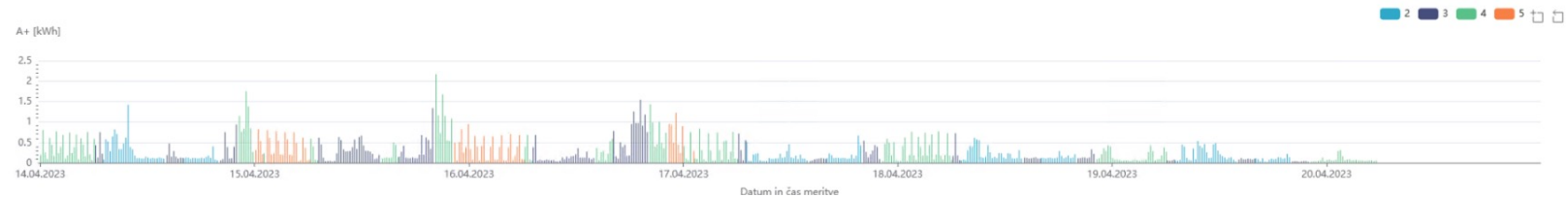
P+_max [kW]	P+_povp. [kW]	A+ [kWh]	P-_max [kW]	P-_povp. [kW]	A- [kWh]
8.7	1.1	165.6	0	0.0	0.0

Dnevna stanja naprave | Prejeta in oddana električna energija | Analiza delovne in jalove moči | Dinamičen tarifni sistem

Agregat profila odjema glede na dinamičen tarifni sistem



Profil odjema energije glede na dinamičen tarifni sistem



Ura dneva	Višja sezona delovni dan	Nižja sezona delovni dan	Višja sezona dela prosti dan	Nižja sezona dela prosti dan
0	3	4	4	5
1	3	4	4	5
2	3	4	4	5
3	3	4	4	5
4	3	4	4	5
5	3	4	4	5
6	2	3	3	4
7	1	2	2	3
8	1	2	2	3
9	1	2	2	3
10	1	2	2	3
11	1	2	2	3
12	1	2	2	3
13	1	2	2	3
14	2	3	3	4
15	2	3	3	4
16	1	2	2	3
17	1	2	2	3
18	1	2	2	3
19	1	2	2	3
20	2	3	3	4
21	2	3	3	4
22	3	4	4	5
23	3	4	4	5

CHALLENGES

- **Data quality**
- **Data governance/security**
  - **SSO (AD/Kerberos integration)**
  - **Apache Ranger** (security, data access)
  - **Apache Atlas** (Data catalogue, lineage)
  - **IBM QRADAR** (SIEM, Auditing)
- **Advanced Analytics Department**
- **In Future:**
  - Data processing in (near) real-time (nRT)
  - Streaming analytics and real-time intelligence
  - Kafka Streaming, Spark Streaming, Apache Flink

Data quality management

- Data needs to be properly verified before being used for BI/data-driven decision-making
- Ensuring data quality has proved to be a major challenge
- Unreliable PLC communication
- How do we technically ensure data quality:
    - Manual verification (domain knowledge required)
    - Finding outliers
    - Replacing missing data – imputation
    - Cleaning duplicate data – deduplication
    - Validation of data type
    - Some data quality issues have been identified in data visualisation
- In progress: Automatic Data Quality monitoring/Data Quality KPIs

Data quality and trust in data are the foundation of a data-driven organisation!

AI/Machine learning

- Spark Mllib
  - Java, Scala, Python, R
- Dedicated compute cluster
- ML Algorithms: classification, regression, clustering, collaborative filtering …
- Featurization: feature extraction, transformation, dimensionality reduction, and selection
- Pipelines: tools for constructing, evaluating, and tuning ML Pipelines
- Persistence: saving and load algorithms, models, and Pipelines
- **MLOps**
  - integrating machine learning models into the software delivery lifecycle
  - end-to-end development, deployment, and operation of machine learning models

AI/Machine learning use cases

Predictive maintenance, condition-based maintenance	Fraud Detection	Reducing consumption and cutting losses
Load Forecasting	Identification of distribution grid	Enables Smart Grid
Generation forecasting	Selecting the optimum location charging stations for electric vehicles	
Fault detection and outage management	RES locations and capacity (power flows, load capacity)	

andrej.somrak@troia.si

TROIA

WE ARE

TROIA

OPTIMIZING THE FUTURE

FOR