SIEMENS
Ingenuity for life

User Experience with KNIME for Large Industrial Data Sets and Applications

Bernhard Lang at KNIME Summit in Berlin, March 2017

Siemens Corporate Technology

Outline – User Experience with KNIME ..

SIEMENS

1. Data analytics applications for selected branches and domains
2. From data analytics applications to solution packages in KNIME
3. Application 1 of analytics framework: commodity price forecasting
4. Application 2 of analytics framework: semantic analytics
5. Key Takeaways

Siemens' installed base of products and solution at customer sites – Tremendous amount of data that can be leveraged for new applications

SIEMENS

The amount of data produced by Siemens products in one day

>10 gigabytes per day

Siemens EnergyIP smart grid platform

>20 gigabytes per day

Siemens gas turbine

>30 gigabytes per day

17.000 Siemens train units

>50 gigabytes per day

Siemens computer tomograph

>80 gigabytes per day

Siemens controllers in particle accelerator CERN

>160 gigabytes per day

Siemens wind turbines

>1 terabytes per day

Siemens traffic management system (one city)

Source: own rough estimations 2016

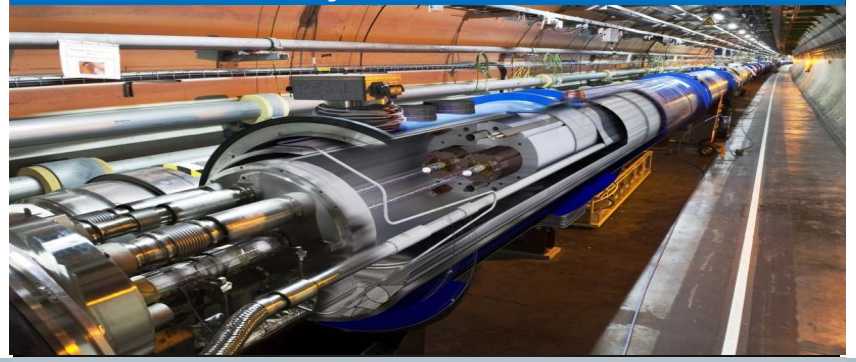
20 years of experience in industrial data analytics applications:
Selected examples for data-driven value generation

SIEMENS

Energy Analytics

Siemens factory in Amberg

Root Cause Analysis & Failure Prediction



CERN Large Hadron Collider

Predictive Maintenance

> 20 high-speed trains at Renfe Spain

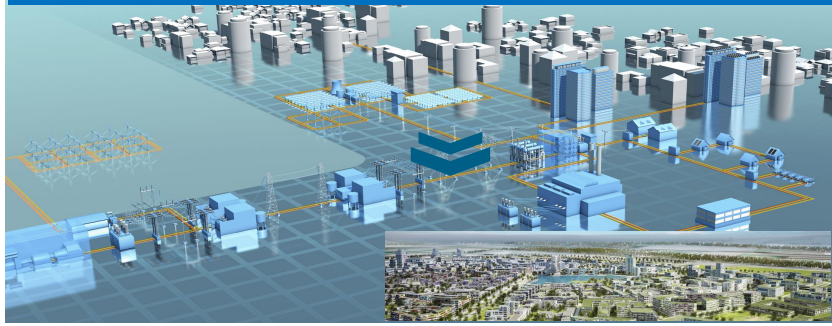
> 20 years online learning neural networks

Deployed in > 30 steel plants

Condition Monitoring & Wear Prediction

Power plants

Asset Performance Management



Smart Grid, Seestadt Aspern, Vienna

Outline – User Experience with KNIME ..

SIEMENS

1 Data analytics applications for selected branches and domains

2 From data analytics applications to solution packages in KNIME

3 Application 1 of analytics framework: commodity price forecasting

4 Application 2 of analytics framework: semantic analytics

5 Key Takeaways

Data Analytics Solution Packages – Let's extract multi-purpose analytics modules (software) out of successful analytics applications in products

SIEMENS

Siemens Data-Driven Services require smart data analytics tools

Solution Packages for Smart Data Analytics

Descriptive Analytics
What happened?
- Visual Analytics
- Service Intelligence
- Enterprise Search

Diagnostic Analytics
Why did it happen?
- Diagnostic Advice
- Condition Monitoring

Predictive Analytics
What will happen?
- Forecasting Services
- Predictive Maintenance

Prescriptive Analytics
What shall I do?
- Autonomous Learning
- Product Configuration
- Operation Planning

Purpose:

Reusable data analytics elements that can be applied for different branches and divisions

Using open source KNIME as analytics integration environment

Benefits:

- Time to market
- Cost efficiency
- Standardization
- Key learnings across Siemens' divisions and branches

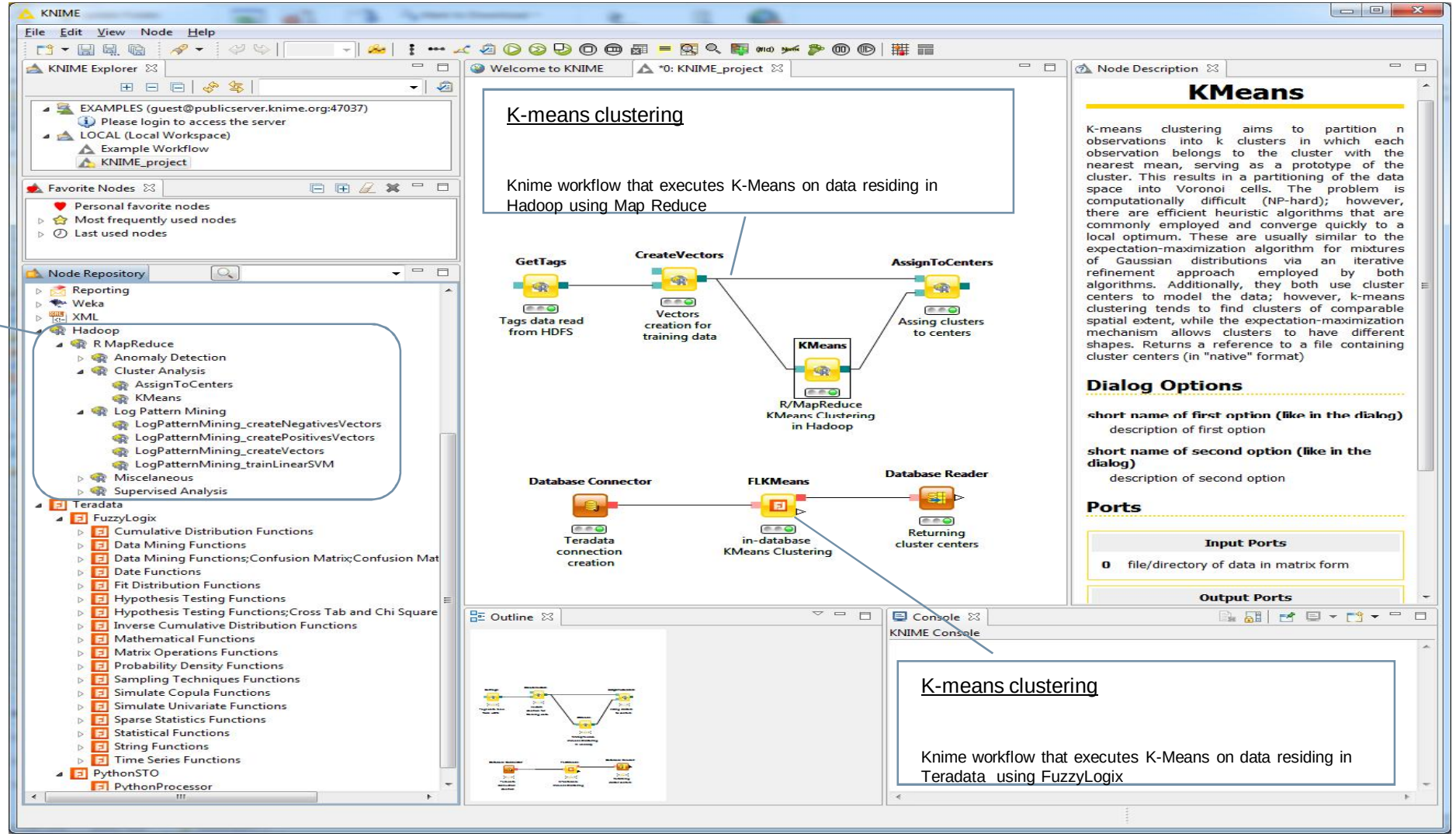
KNIME as a Data Analytics Workflow Editor – Integration of third party analytics tools (also Siemens' solution packages)

SIEMENS

Hadoop Nodes

Nodes that interact with Hadoop. The actual analytic code is written in R and executes as Map Reduce (*approx. 25 nodes already available*)

Python/C++/Java are also supported



More info about the KNIME Analytics Platform:

http://www.knime.org/

Outline – User Experience with KNIME ..

SIEMENS

1 Data analytics applications for selected branches and domains

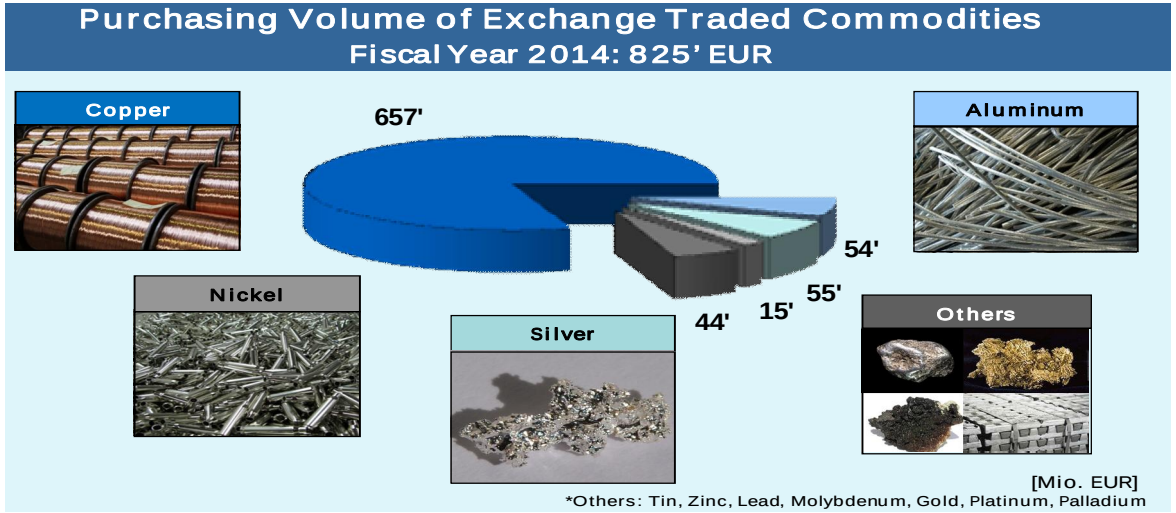
2 From data analytics applications to solution packages in KNIME

3 Application 1 of analytics framework: commodity price forecasting

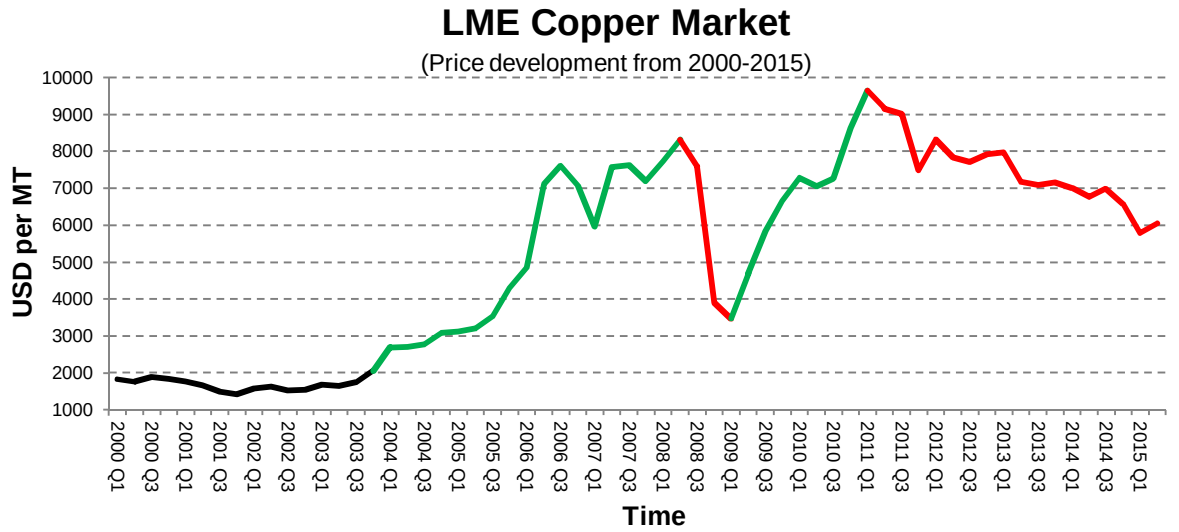
4 Application 2 of analytics framework: semantic analytics

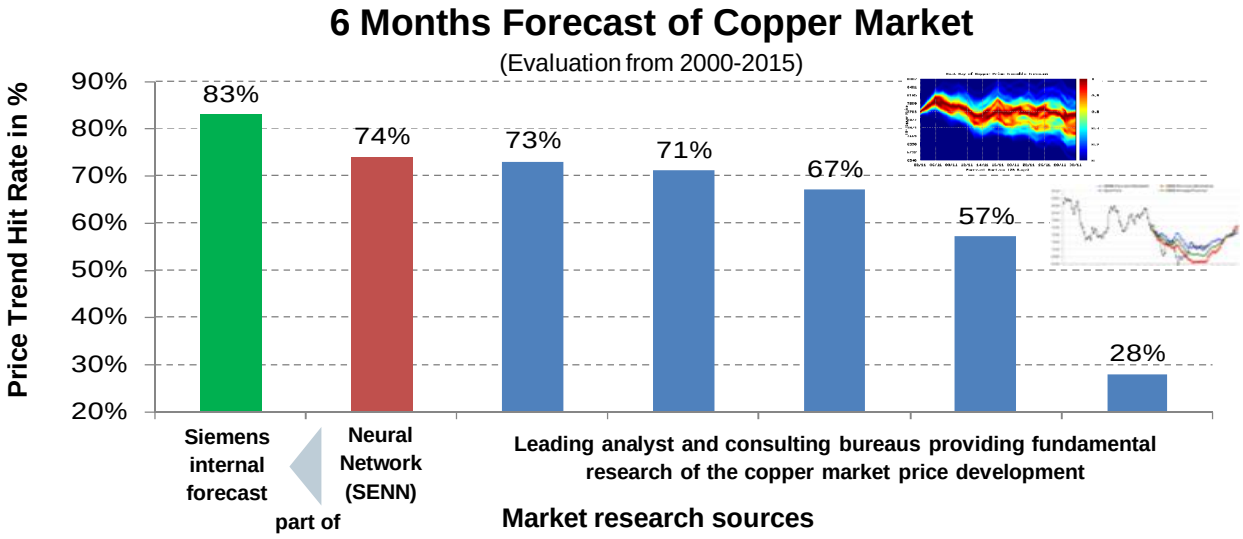
5 Key Takeaways

Solution Package Forecasting: Commodity Prices with Neural Networks

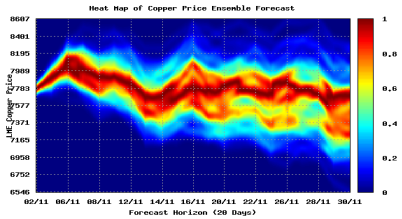


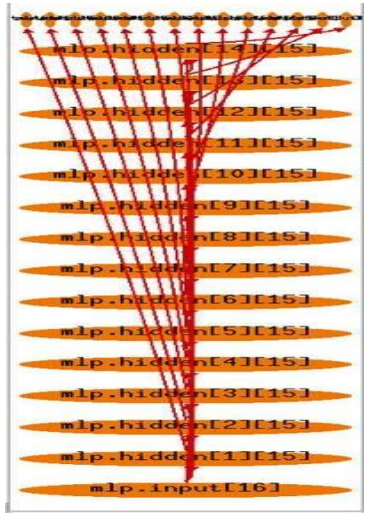
- LME Copper: 79% of total purchasing volume
- How to cope with the increased market volatility?
- Optimal procurement decisions on the basis of accurate market forecasts and research
- Benchmarking of different vendors for market intelligence
- **Siemens internal forecast** is based on **Neural Networks forecast** and selected other sources
- Extension to other metals and energy

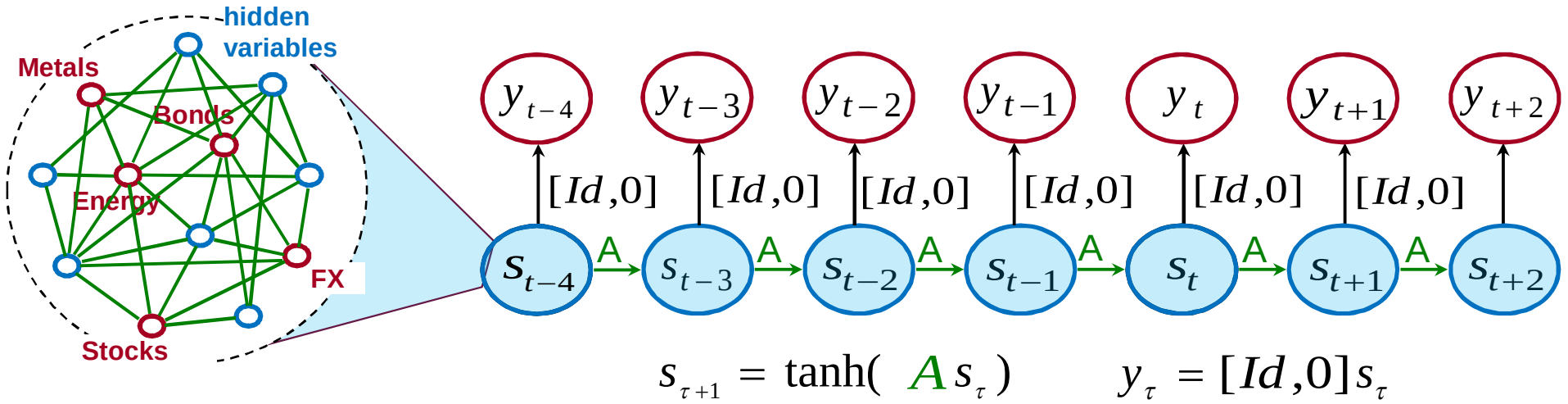




Market Price Predictions with Large Neural Networks







$$s_{\tau+1} = \tanh(A s_\tau) \qquad y_\tau = [Id,0] s_\tau$$

- Our neural networks model coherent markets as interacting dynamical systems
- Given the subset of the observables we reconstruct the hidden variables
- Different market structures and multiple time scale dynamics can be addressed
- From open to closed dynamical systems: The model is dynamical consistent, symmetric in all variables and present time does not play any special role

Zimmermann; Tietz; Grothmann: Forecasting with Recurrent Neural Networks, In: Neural Networks: Tricks of the Trade, **2nd ed.**; Springer, 2012

TopicRadar: A Market Mood Indicator for Commodity Price Forecasting

SIEMENS

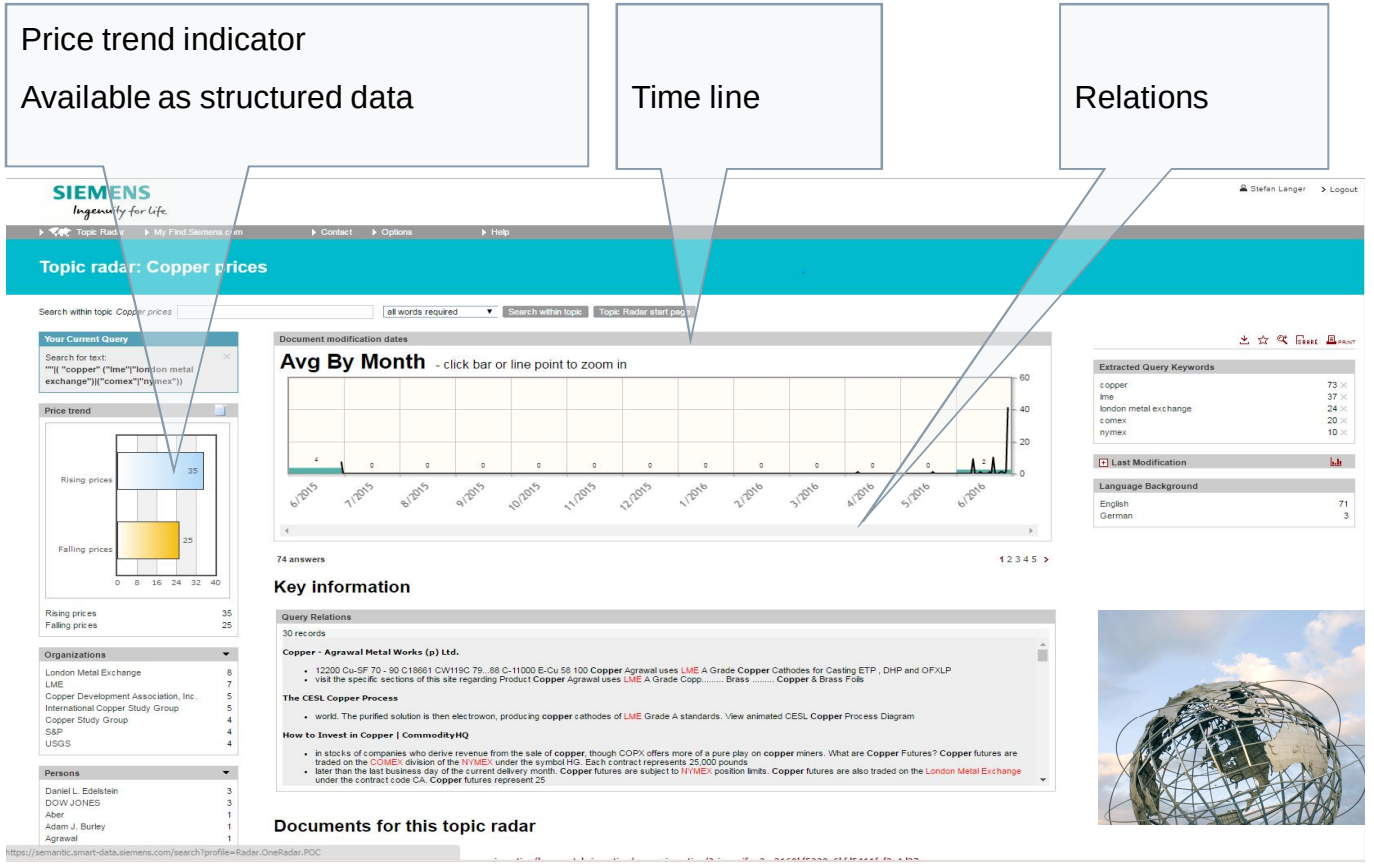
Enrich and analyze internal and external unstructured data

Named entity and event recognition

- Extract named entities or events
  - query entities,
  - company names,
  - price trend indicators

Relation extraction

- Detect, locate & display relation between query entities and events



Modular Framework of Data Analytics Solution Packages applied for Commodity Price Forecasting

SIEMENS

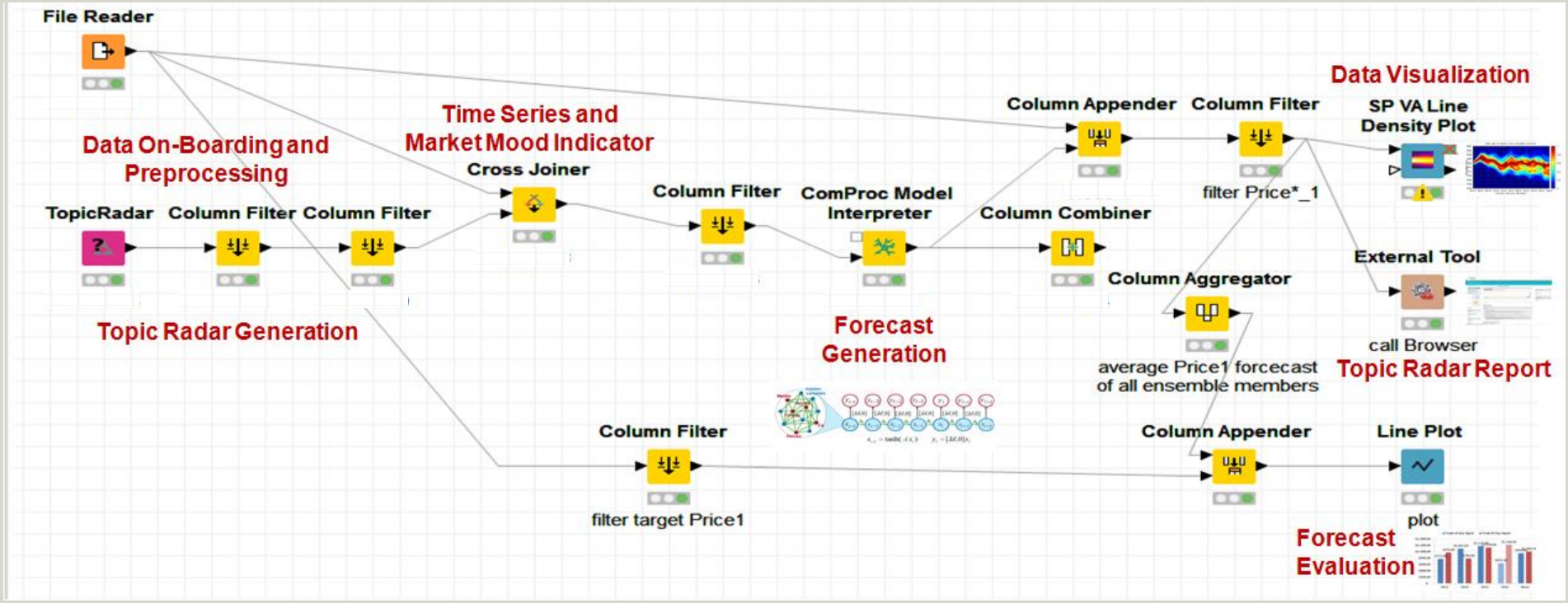
Task: Forecasting of market trends supports hedging strategies and optimal procurement

Forecasting Services:
A forecast solution and a service, which covers the dynamic of markets and their uncertainty

Enterprise Search:
Supports the exploitation of human generated content

Visual Analytics:
Consulting and realization over the complete Visual Analytics process based on reusable building blocks

Autonomous Learning:
Identifies the relevant control variables and generates continuous control policies



Business Impact:
- Identification of potential market scenarios and estimation of related market price risks
- Optimal procurement decisions and decision support for competitive advantages
- Forecasts are provided as a service. Performance based contracting

Outline – User Experience with KNIME ..

SIEMENS

1 Data analytics applications for selected branches and domains

2 From data analytics applications to solution packages in KNIME

3 Application 1 of analytics framework: commodity price forecasting

4 Application 2 of analytics framework: semantic analytics

5 Key Takeaways

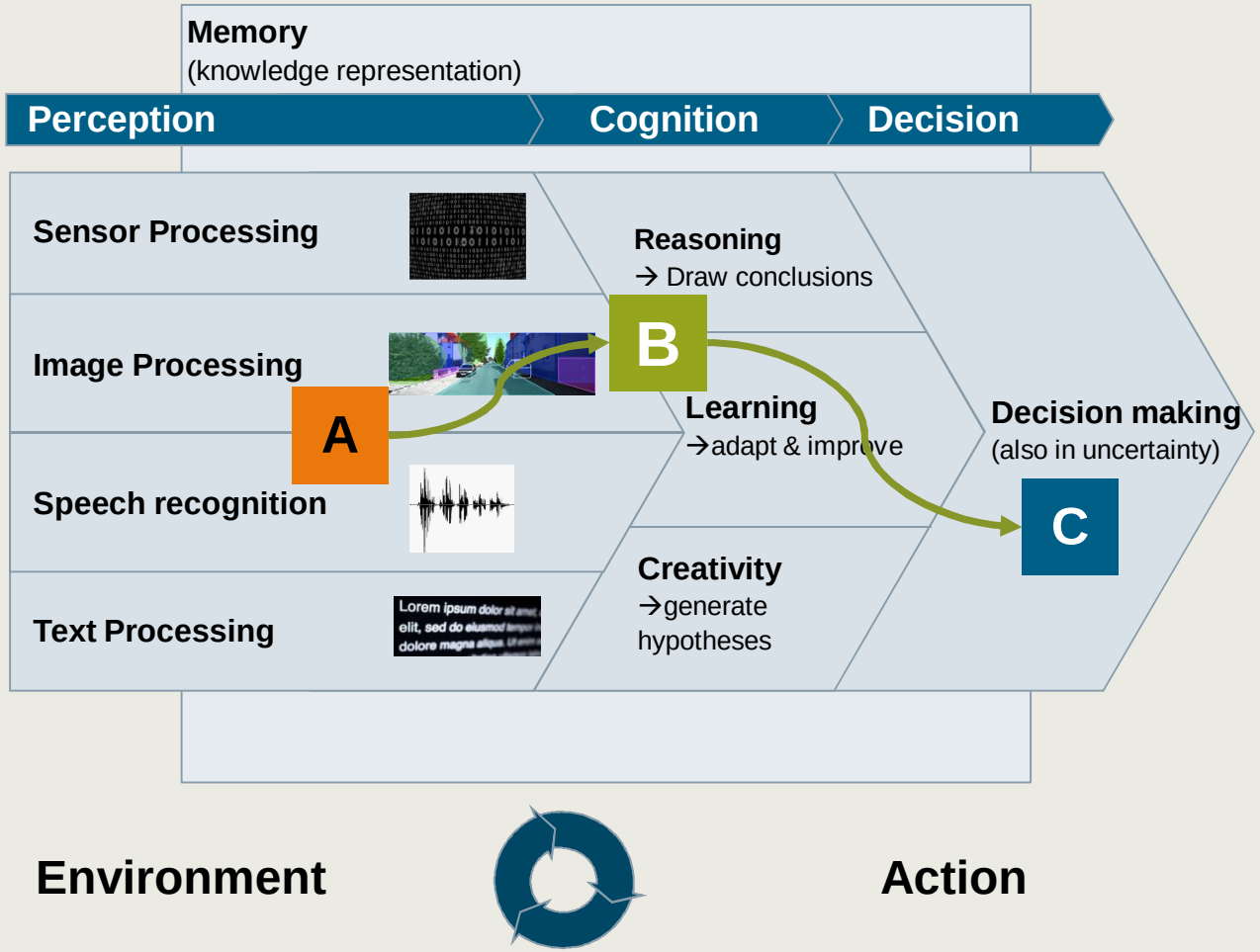
SIEMENS
Ingenuity for life

Semantic Analytics.
The Why and the How.

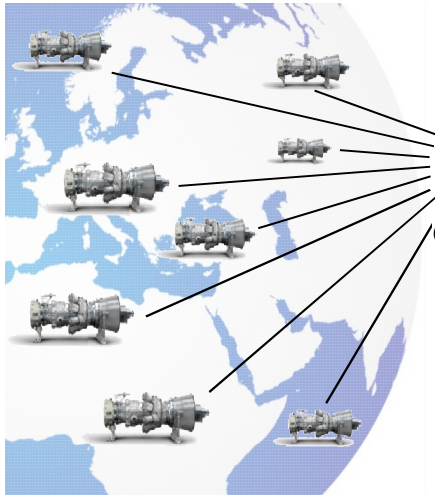
Dr.-Ing. Sebastian Brandt

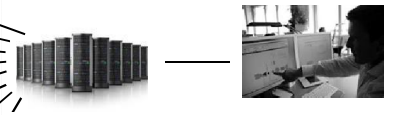
The AI renaissance:
the science and engineering of making intelligent machines

SIEMENS



Example Device Service Automation





Global data centers

Remote diagnostic centers

- Monitoring
- Root cause analysis
- Predictive maintenance
- Reports
- Global service products

A Deep learning for object recognition and labeling in service reports

B Semantic knowledge fusion and reasoning for integrated diagnostics

C Automated planning of maintenance service and activities

B

Connecting industrial knowledge (sources)

SIEMENS

Data Sources

Static aspects

R&D data

Engineering data

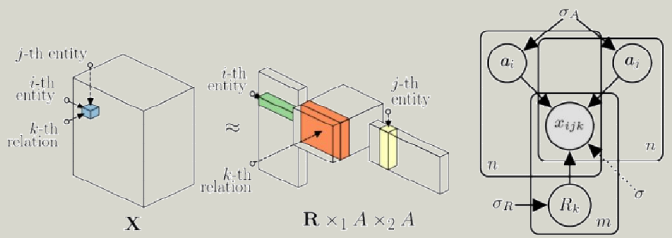
Plant data

Dynamic aspects

Service data

Monitoring data

Relational Learning (e.g. via Tensor Factorization)

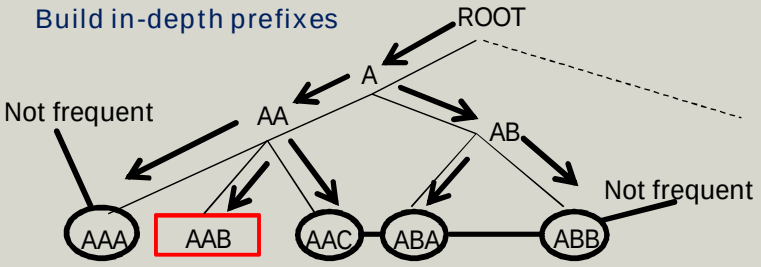


Tresp, Nickel. Tensor Factorization for Multi-Relational Learning, ECML, 2013

Information extraction

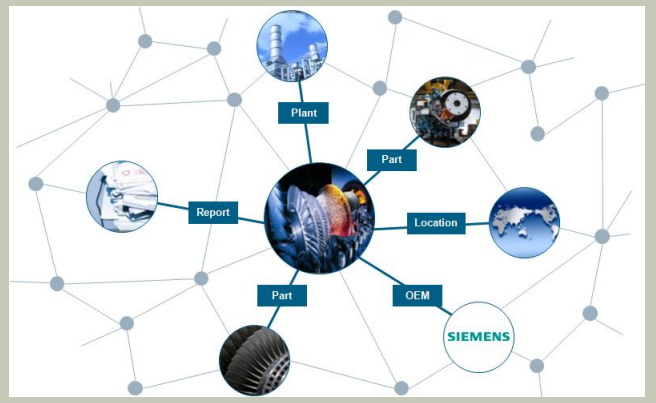
(e.g. Natural Language Processing)

Pattern Sequence Mining (e.g. via PrefixSpan)



Industrial Knowledge Graph

Knowledge fusion into one coherent semantic model



Examples for automated graph construction

B

Enable intuitive end-user access to industrial data

SIEMENS

Data Sources

Static aspects

R&D data

Engineering data

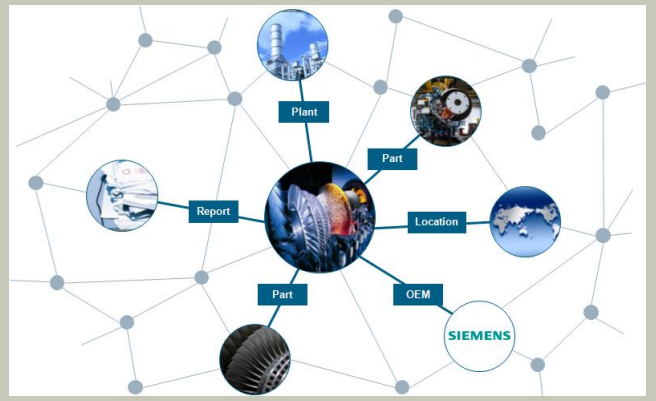
Plant data

Dynamic aspects

Service data

Monitoring data

Industrial Knowledge Graph



Ontology-based Data Access

Normal start?

Query

Analytics

SAP HANA

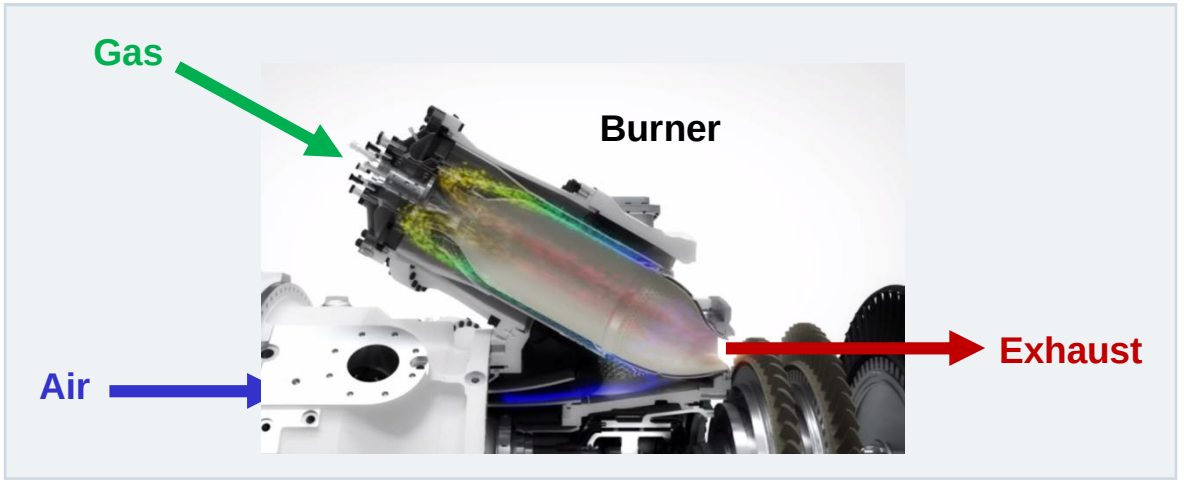
TERADATA ASTER

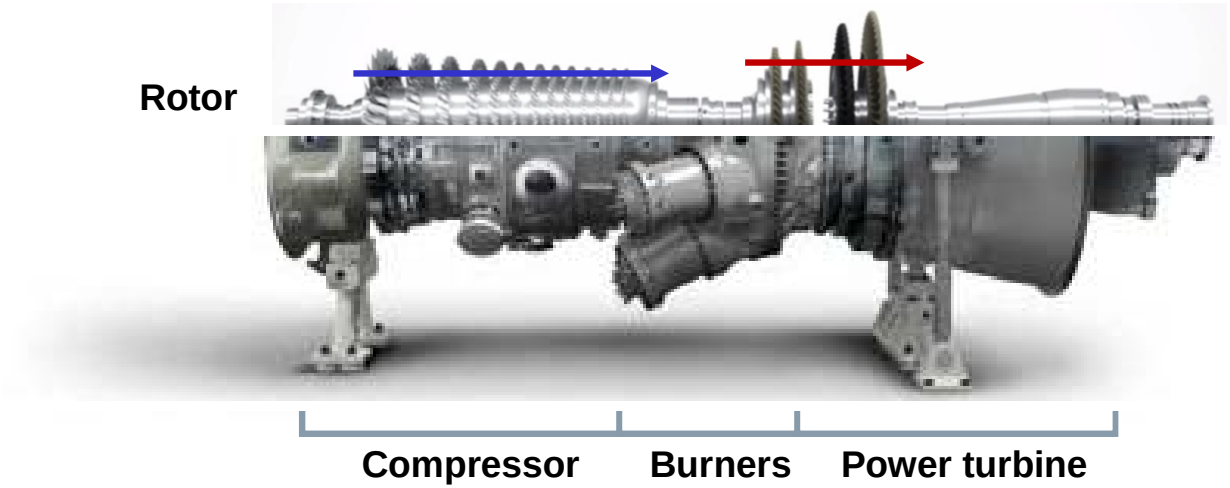
TIBCO

Java

KNIME

Gas turbine crash course





38 MW

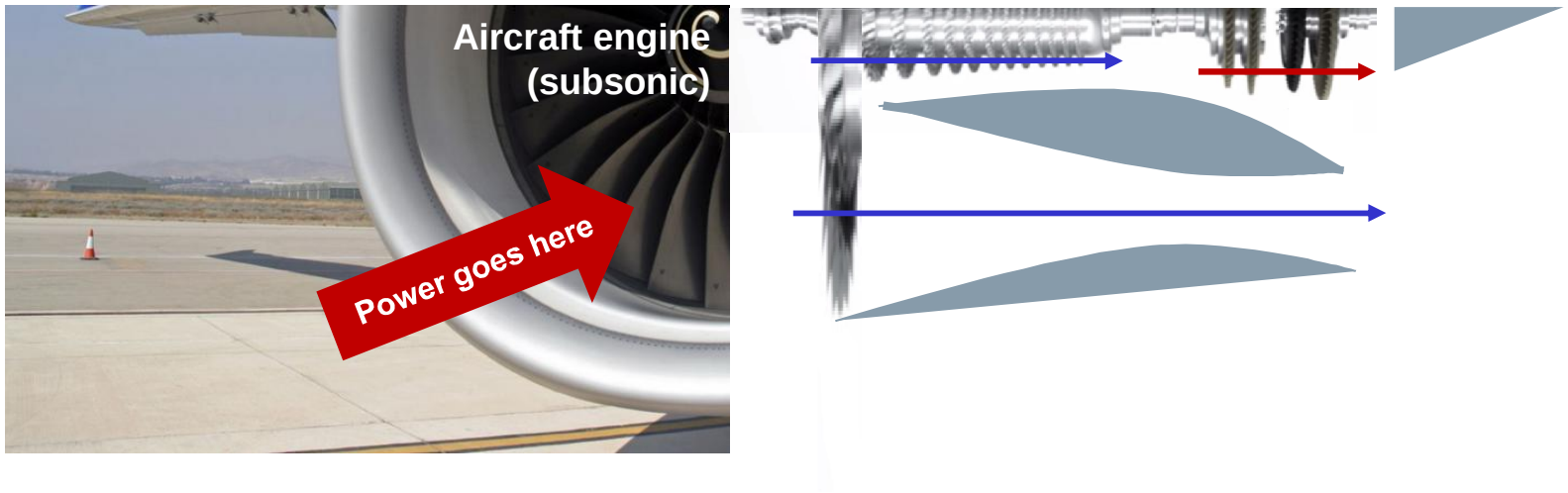
1 000 000 light bulbs

500 cars :-)

Aircraft engine crash course

SIEMENS

Gas turbine



Turbine data: time series + events

Optique

SIEMENS

Signal data

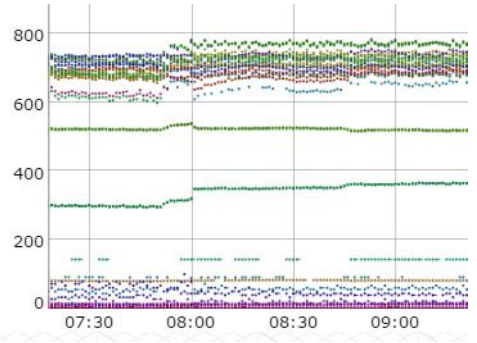
- **Tag**
- Timestamp
- Number

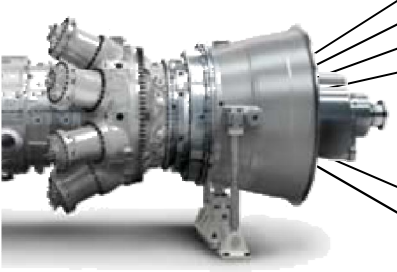
BSX-TC3562-XE01, 12:34:56,21.09.2015; 560

BSX-TMP12A-XE01, 12:34:57,21.09.2015; 0

BSX-TICCFB1-XE01, 12:34:55,21.09.2015; 1

BSX-TIFB2-XE02, 12:34:56,21.09.2015; 9564





up to 3000 sensors per turbine

Event messages

- Category
- Timestamp
- Text

Status, 12:34:56,21.09.2015, “Ignitor on”

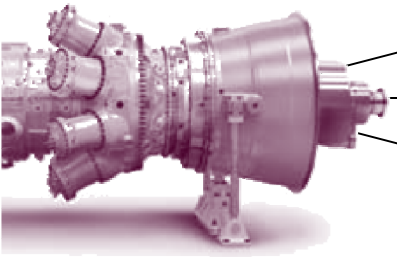
Warning, 12:38:56,21.09.2015, “Overspeed”

event	Time
WATER WASH ALGORITHM CALCULATIONS SUSPENDED	Dec 01, 2009 00:30:14
Normal Stop	Dec 01, 2009 00:38:25
NORMAL STOP IN PROGESS	Dec 01, 2009 00:38:25
Shutdown Spin Will Be Carried Out To Cool Engine	Dec 01, 2009 00:38:25
WATER WASH ALGORITHM CALCULATIONS SUSPENDED	Dec 01, 2009 00:38:25
STOP INITIATED VIA UNIT CONTROLLER OPERATOR PANEL PUSH BUTTON	Dec 01, 2009 00:38:25
TURBINE RUNNING ON LIQUID FUEL	Dec 01, 2009 00:38:25
LIQUID BURNER REVERSE PURGE IN PROGRESS	Dec 01, 2009 00:38:25
STOP INITIATED VIA UNIT CONTROLLER OPERATOR PANEL PUSH BUTTON	Dec 01, 2009 00:38:28
LIQUID BURNER FORWARD AIR PURGE IN PROGRESS	Dec 01, 2009 00:38:33
INSTRUMENT AIR ASSIST PURGE IN PROGRESS	Dec 01, 2009 00:38:35

Optique

SIEMENS

Semantic knowledge fusion and reasoning for integrated diagnostics

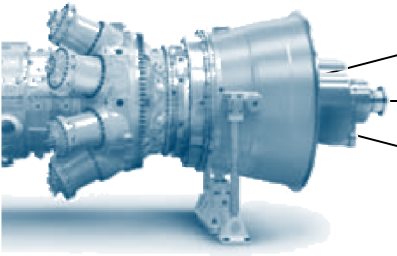


BSX-TC3562-XE01
BSX-TMP12A-XE01
BSX-TICCFB1-XE01
hypothetical identifiers

DC1, DB X1, TY2

Query1

Normal start ?

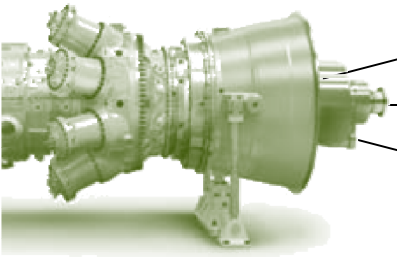


MS-XC255-X12
BSX-TC3562-XE01
BSX-TC3562-XE01
hypothetical identifiers

DC2, DB X2, TY2

Query2

Normal start ?



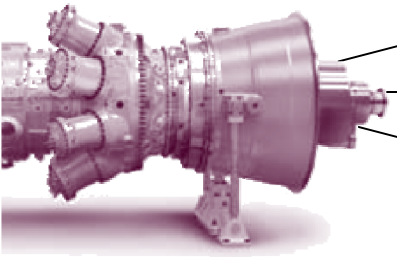
MRR-T8901-8462
CRR-M8393-9272
“Ignitor on”
hypothetical identifiers

DC2, DB X2, TY2
DC2, DB X2, TY4

Query3

Normal start ?

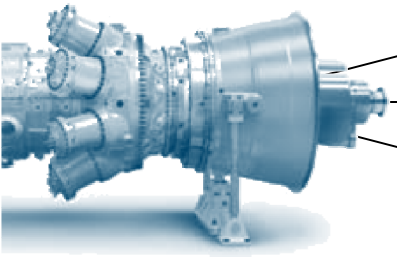
Abstraction enables uniform solutions



BSX-TC3562-XE01

BSX-TMP12A-XE01

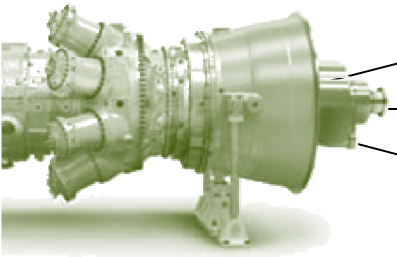
BSX-TICCFB1-XE01



MS-XC255-X12

BSX-TC3562-XE01

BSX-TC3562-XE01



MRR-T8901-8462

CRR-M8393-9272

“Ignitor on”

Semantic mapping

Sensor types, turbine structure, site con-figurations, measurable quantities, Processes

Domain ontology

Normal start?

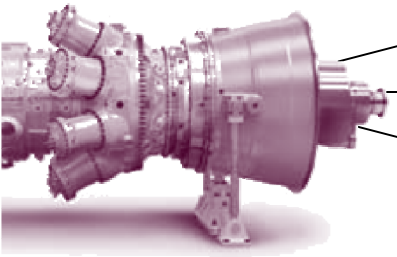
Query

Analytics

Abstraction enables uniform solutions

Optique

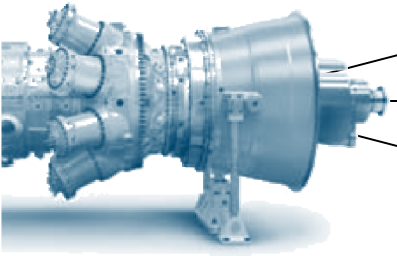
SIEMENS



BSX-TC3562-XE01

BSX-TMP12A-XE01

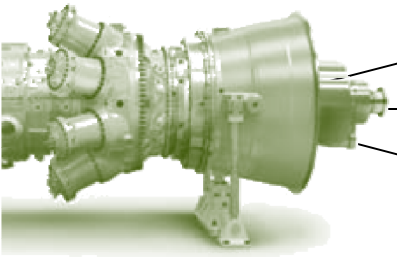
BSX-TICCFB1-XE01



MS-XC255-X12

BSX-TC3562-XE01

BSX-TC3562-XE01



MRR-T8901-8462

CRR-M8393-9272

"Ignitor on"

Semantic mapping

Sensor types,

turbine structure,

site con-figurations,

measurable quantities,

Processes

Domain ontology

Normal start?

Query

Analytics

SAP HANA

TERADATA ASTER

TIBCO

Java

KNIME

Abstraction enables uniform solutions

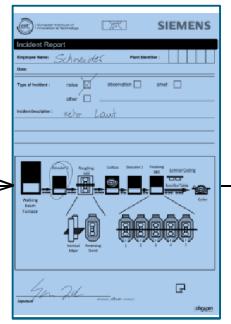
Optique

SIEMENS

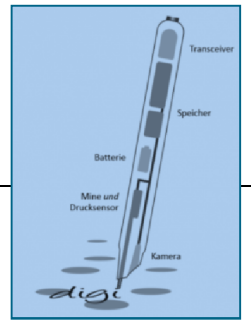
Free-text Docs

OCR → NLP → Indexing

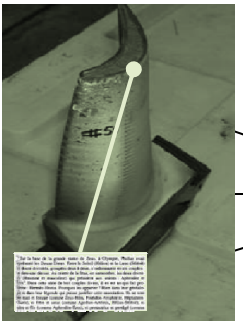
Inspections



Observation sheet
(hypothetical)



Digipen



Images + Text

NLP & Deep Learning → Indexing

Semantic mapping

Domain ontology

Sensor types,
turbine structure,
site con-figurations,
measurable quantities,
Processes

Normal start?

Query

Analytics

* http://optique-project.eu/

Data Access & (Virtual) Integration

SIEMENS

Dashboards

- lightweight: Semantic Media Wikis
- proprietary: via Web Services, ETL, ….

Query

- SPARQL
- based on **Knowledge Graph**
- no joins, columns, tables, etc.

Analytics services

- cross-fleet, cross-platform analytics simplified
- ensure soundness of combined services

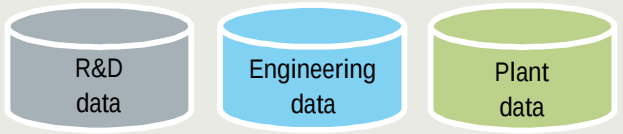
Semantic mappings (R2RML)

- connect knowledge graph to conventional (Big-)data sources
- utilises Knowledge Graph vocabulary
- supports relational database
- supports Big-Data infrastructures
- supports web-services, APIs, document repositories
- supports federation
- supports data streams
- botstrapping

```
?{plantName} : sie:Plant and
sie:hasCustomer ?{customerName}

select Plants.plantName, Customers.customerName
from Plants, Customers, Plant2CustMap
where Plants.plantId=Plant2CustMap.plantId
and Customers.customerId=Plant2CustMap.customerId
```

Static aspects

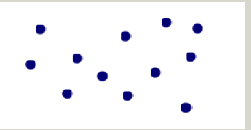


Dynamic aspects

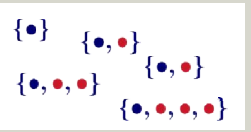
Web data

Data streams

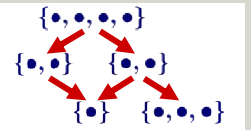
Behind the knowledge graph…



Controlled vocabulary

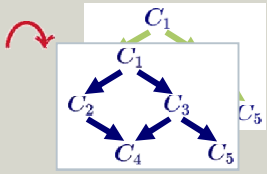


Thesaurus



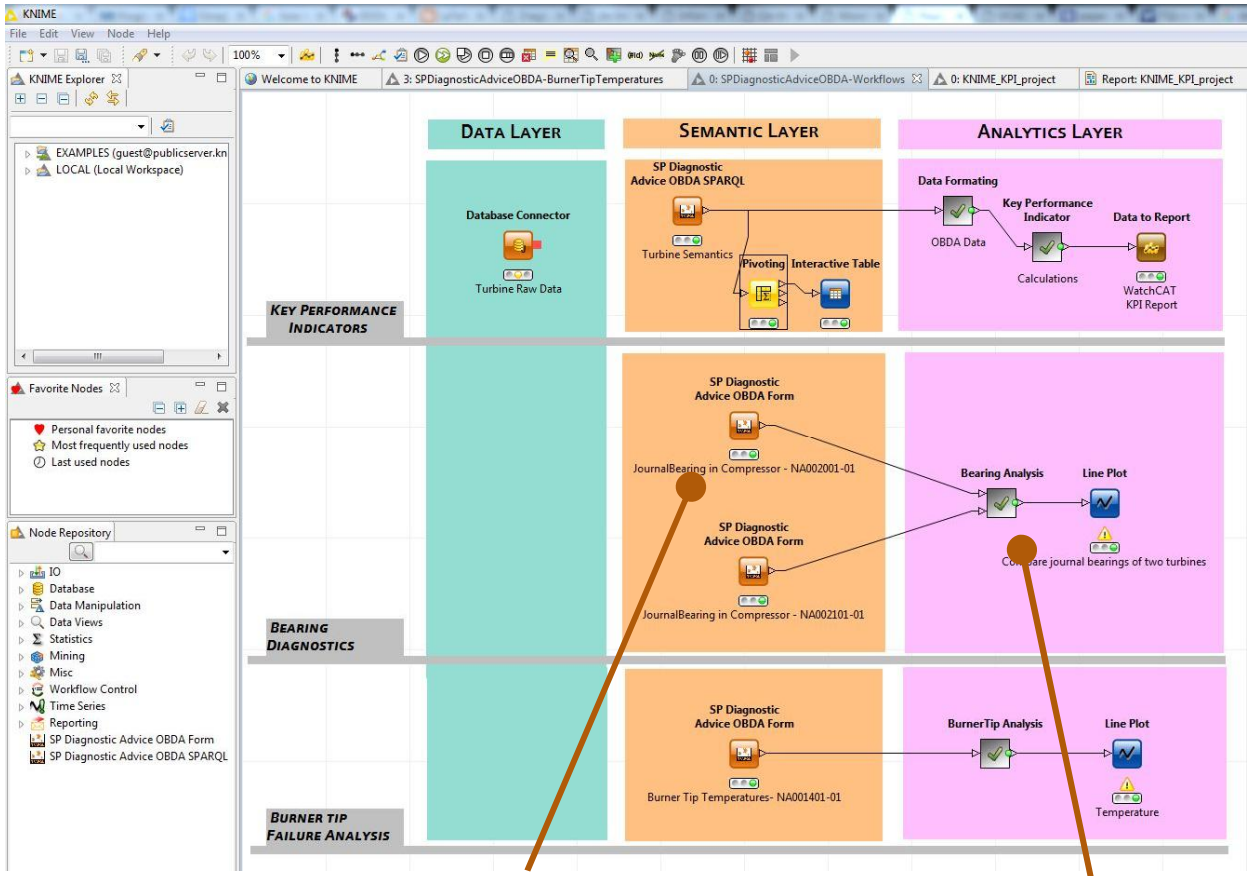
Terminology / Taxonomy

Ontology



Remote monitoring example:
Combining semantics and analytics in KNIME

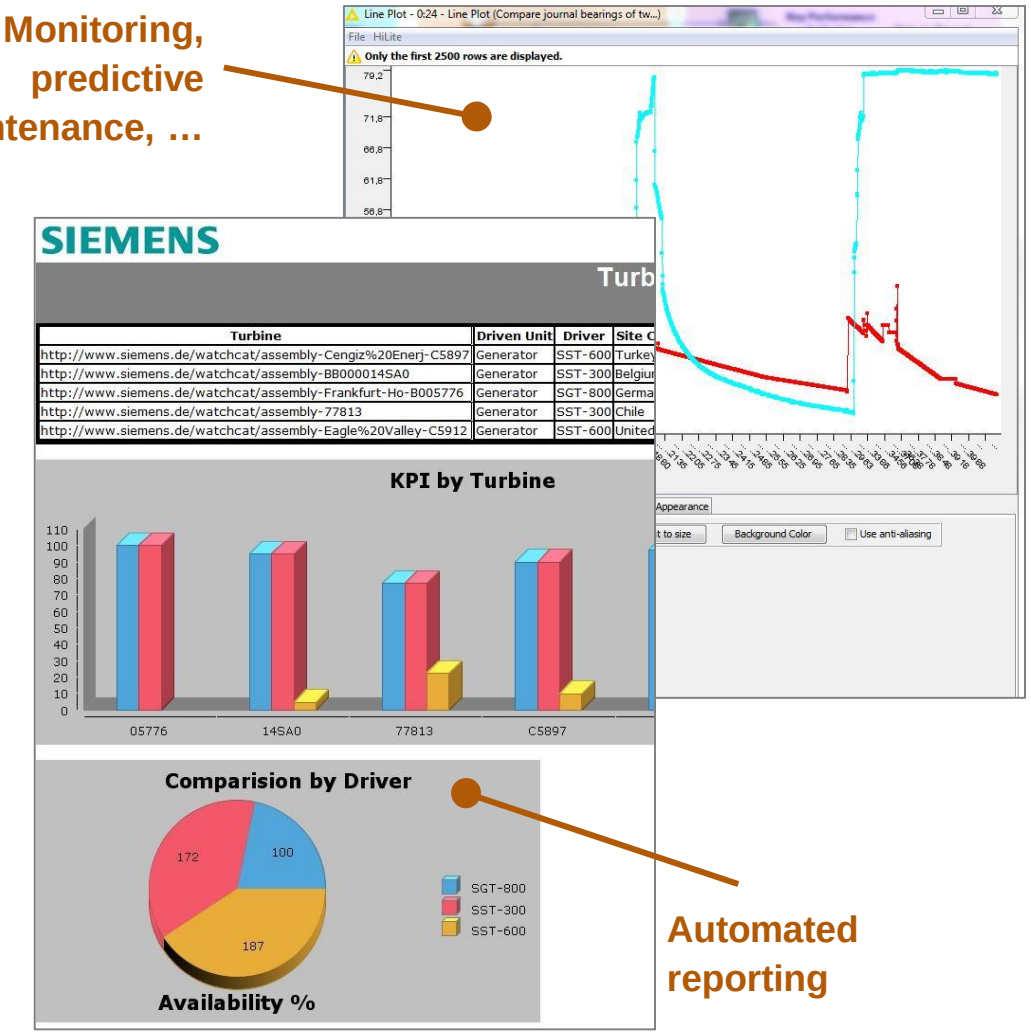
SIEMENS



Model-based access as KNIME building block

Generic KNIME analytics workflows

Monitoring, predictive maintenance, …



Automated reporting

Outline – User Experience with KNIME ..

SIEMENS

1 Data analytics applications for selected branches and domains

2 From data analytics applications to solution packages in KNIME

3 Application 1 of analytics framework: commodity price forecasting

4 Application 2 of analytics framework: semantic analytics

5 Key Takeaways

Key Takeaways

Challenge of big data is not so much the size or the right tool but the inconsisteny of data.

Big data analytics without domain know-how and product/context know-how often fails.

For the last years KNIME is developing fast towards the needs of productive use ☺.

Contact Information

Bernhard Lang
Siemens AG
Corporate Technology

Project Manager
CTI Smart Data to Business

bernhard.lang@siemens.com

+49 (173) 7337 894

Youtube ”Siemens Smart Data”:
https://www.youtube.com/watch?v=ZxoO-DvHQRw

Example Semantic Analytics:

Dr.- Ing. Sebastian Brandt