User Experience with KNIME for Large Industrial Data Sets and Applications
Bernhard Lang at KNIME Summit in Berlin, March 2017

Siemens Corporate Technology
<table>
<thead>
<tr>
<th></th>
<th>Outline – User Experience with KNIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data analytics applications for selected branches and domains</td>
</tr>
<tr>
<td>2</td>
<td>From data analytics applications to solution packages in KNIME</td>
</tr>
<tr>
<td>3</td>
<td>Application 1 of analytics framework: commodity price forecasting</td>
</tr>
<tr>
<td>4</td>
<td>Application 2 of analytics framework: semantic analytics</td>
</tr>
<tr>
<td>5</td>
<td>Key Takeaways</td>
</tr>
</tbody>
</table>
Siemens’ installed base of products and solution at customer sites – Tremendous amount of data that can be leveraged for new applications

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

- Energy Analytics
  - Siemens factory in Amberg

- Root Cause Analysis & Failure Prediction
  - CERN Large Hadron Collider

- Predictive Maintenance
  - > 20 high-speed trains at Renfe Spain

- Condition Monitoring & Wear Prediction
  - Deployed in > 30 steel plants
  - Power plants

- Asset Performance Management
  - Smart Grid, Seestadt Aspern, Vienna

> 20 years online learning neural networks

Deployed in > 30 steel plants

Power plants
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Data analytics applications for selected branches and domains</strong></td>
</tr>
<tr>
<td>2</td>
<td><strong>From data analytics applications to solution packages in KNIME</strong></td>
</tr>
<tr>
<td>3</td>
<td><strong>Application 1 of analytics framework: commodity price forecasting</strong></td>
</tr>
<tr>
<td>4</td>
<td><strong>Application 2 of analytics framework: semantic analytics</strong></td>
</tr>
<tr>
<td>5</td>
<td><strong>Key Takeaways</strong></td>
</tr>
</tbody>
</table>
Data Analytics Solution Packages – Let’s extract multi-purpose analytics modules (software) out of successful analytics applications in products

Siemens Data-Driven Services require smart data analytics tools

Solution Packages for Smart Data Analytics

<table>
<thead>
<tr>
<th>Descriptive Analytics</th>
<th>Diagnostic Analytics</th>
<th>Predictive Analytics</th>
<th>Prescriptive Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>What happened?</td>
<td>Why did it happen?</td>
<td>What will happen?</td>
<td>What shall I do?</td>
</tr>
<tr>
<td>Visual Analytics</td>
<td>Diagnostic Advice</td>
<td>Forecasting Services</td>
<td>Autonomous Learning</td>
</tr>
<tr>
<td>Service Intelligence</td>
<td>Condition Monitoring</td>
<td>Predictive Maintenance</td>
<td>Product Configuration</td>
</tr>
<tr>
<td>Enterprise Search</td>
<td></td>
<td></td>
<td>Operation Planning</td>
</tr>
</tbody>
</table>

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)

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

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
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
TopicRadar: A Market Mood Indicator for Commodity Price Forecasting

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

Price trend indicator
Available as structured data

Time line

Relations

Documents for this topic radar
Modular Framework of Data Analytics Solution Packages applied for Commodity Price Forecasting

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

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

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
Semantic Analytics.
The Why and the How.

Dr.-Ing. Sebastian Brandt
The AI renaissance: 
the science and engineering of making intelligent machines

Memory (knowledge representation)

Sensor Processing

Image Processing

Speech recognition

Text Processing

Reasoning 
\[ \text{Draw conclusions} \]

Learning 
\[ \text{adapt & improve} \]

Creativity 
\[ \text{generate hypotheses} \]

Perception 

Cognition 

Decision

Decision making 
(\text{also in uncertainty})

Environment

Action

Remote diagnostic centers

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

Deep learning for object recognition and labeling in service reports

Semantic knowledge fusion and reasoning for integrated diagnostics

Automated planning of maintenance service and activities

Example Device Service Automation

Example

Device Service Automation
Connecting industrial knowledge (sources)

**Data Sources**

- Static aspects
  - R&D data
  - Engineering data
  - Plant data

- Dynamic aspects
  - Service data
  - Monitoring data

**Relational Learning** (e.g. via Tensor Factorization)

**Pattern Sequence Mining** (e.g. via PrefixSpan)

**Industrial Knowledge Graph**

Knowledge fusion into one coherent semantic model

Examples for automated graph construction

Build in-depth prefixes

Not frequent

ROOT

AAA AAB AAC ABA ABB

Not frequent

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

Information extraction
(e.g. Natural Language Processing)

Examples for automated graph construction

Unrestricted © Siemens AG 2017

March 2017

Bernhard Lang - Corporate Technology
Enable intuitive end-user access to industrial data

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
Gas turbine crash course

38 MW
1,000,000 light bulbs
500 cars :-)

Unrestricted © Siemens AG 2017
Page 18 March 2017
Bernhard Lang - Corporate Technology
Aircraft engine crash course

Gas turbine

Power goes here

Aircraft engine (subsonic)

Power goes here
**Signal data**
- **Tag**
- **Timestamp**
- **Number**

**Event messages**
- **Category**
- **Timestamp**
- **Text**

<table>
<thead>
<tr>
<th>Tag</th>
<th>Timestamp</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSX-TC3562-XE01, 12:34:56, 21.09.2015</td>
<td>560</td>
<td></td>
</tr>
<tr>
<td>BSX-TMP12A-XE01, 12:34:57, 21.09.2015</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BSX-TICCBB1-XE01, 12:34:55, 21.09.2015</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>BSX-TIFB2-XE02, 12:34:56, 21.09.2015</td>
<td>9564</td>
<td></td>
</tr>
</tbody>
</table>

up to 3000 sensors per turbine
Semantic knowledge fusion and reasoning for integrated diagnostics

- **Query1**: BSX-TC3562-XE01, BSX-TMP12A-XE01, BSX-TICCFB1-XE01 -> DC1, DB X1, TY2
- **Query2**: MS-XC255-X12, BSX-TC3562-XE01, BSX-TC3562-XE01 -> DC2, DB X2, TY2
- **Query3**: MRR-T8901-8462, CRR-M8393-9272, "Ignitor on" -> DC2, DB X2, TY2, DC2, DB X2, TY4

Hypothetical identifiers:
- 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"
Abstraction enables uniform solutions

Sensor types, turbine structure, site configurations, measurable quantities, Processes

Normal start?

Query

Domain ontology

Semantic mapping

Analytics

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Query**

Sensor types, turbine structure, site configurations, measurable quantities, Processes

**Analytics**

Domain ontology

Semantic mapping

Query

Normal start?
Abstraction enables uniform solutions

Remote monitoring example:

Abstraction enables uniform solutions (Optique)

Sensor types, turbine structure, site configurations, measurable quantities, Processes

Domain ontology

Semantic mapping

Normal start?

Query

Analytics

* http://optique-project.eu/*
Abstraction enables uniform solutions

Sensor types, turbine structure, site configurations, measurable quantities, Processes

Domain ontology

Semantic mapping

Query

Normal start?

Analytics

NLP & Deep Learning

Indexing

Digipen

Observation sheet (hypothetical)

Inspections

Images + Text

Free-text Docs

OCR

NLP

Indexing

Abstraction enables uniform solutions

Remote monitoring example:

Semantics in the wider ecosystem

http://optique-project.eu/

Optique
Data Access & (Virtual) Integration

**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**
- R&D data
- Engineering data
- Plant data
- Service data
- Monitoring data

**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

Monitoring, predictive maintenance, ...

Model-based access as KNIME building block

Generic KNIME analytics workflows

Automated reporting
Outline – User Experience with KNIME

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