**SIEMENS**

Siemens Corporate Technology, RTC GTF BAM | December 10, 2013

# Siemens Data Analytics @ CERN openLab

**Dr. Mikhail Roshchin, Michal Skubacz
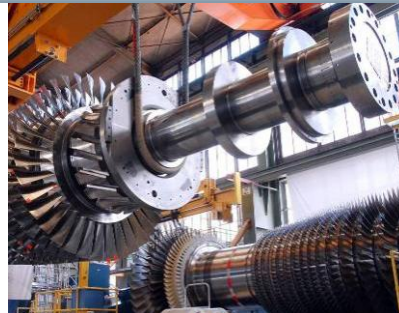Mikhail Kalinkin, Alexander Loginov**

# Siemens – Innovations in four Sectors

Siemens: Facts and Figures

## Energy

**World record in energy efficiency**

- New H-class gas turbine

## Industry

**Efficiency in industrial production**

- Totally Integrated Automation Portal

## Healthcare

**Affordable and personalized healthcare**

- MAGNETOM Spectra

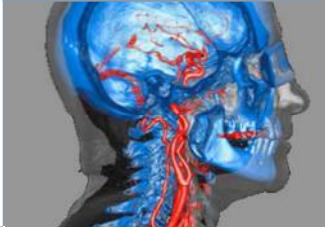## Infrastructure & Cities

**Intelligent infrastructure**

- Driverless subway trains

29,500 R&D engineers | 17,500 SW engineers | 160 R&D locations in 30 countries | 4.2bn EUR R&D spending in 2012 | 57,300 patents | 8.900 invention disclosures in 2012

# Siemens has a strong position in Vertical IT

Software domains and respective tools and systems in the Siemens sectors (examples)

| Domains | Industry | Infrastructure and Cities | Energy | Healthcare |
|---|---|---|---|---|
| **Engineering Software** | • Product&Production Lifecycle Tools<br>• Equipment commissioning | • Engineering Tools for rail, traffic and building automation | • Power Plant Engineering | |
| **Supervisory Software** | • Manufacturing Execution Systems<br>• Distributed Control System<br>• Scada Systems | • Meter Data Mgt.<br>• Traffic Mgt.<br>• Rail Operations & Control<br>• Building Mgt. | • Distributed Control System<br>• Fleet Mgt.<br>• Energy Mgt.<br>• Power Plant Mgt. | • Advanced visualization<br>• Hospital and laboratory information systems |
| **Embedded Software** | • Programmable Logical Controller<br>• Motion Controller Drives, Intelligent field-devices | • Building Automation Equipment<br>• Intersection Control<br>• Train Control | • Power Plant Automation Systems | • Imaging platforms inside the modalities<br>• Diagnostic systems |
| **Intelligent devices** | | | | |

**17.500 SW engineers**

December 10, 2013       Dr. Mikhail Roshchin

# Smart Data to Business: Leveraging additional business opportunities based on smart domain data

Data sources and processing



December 10, 2013      Dr. Mikhail Roshchin
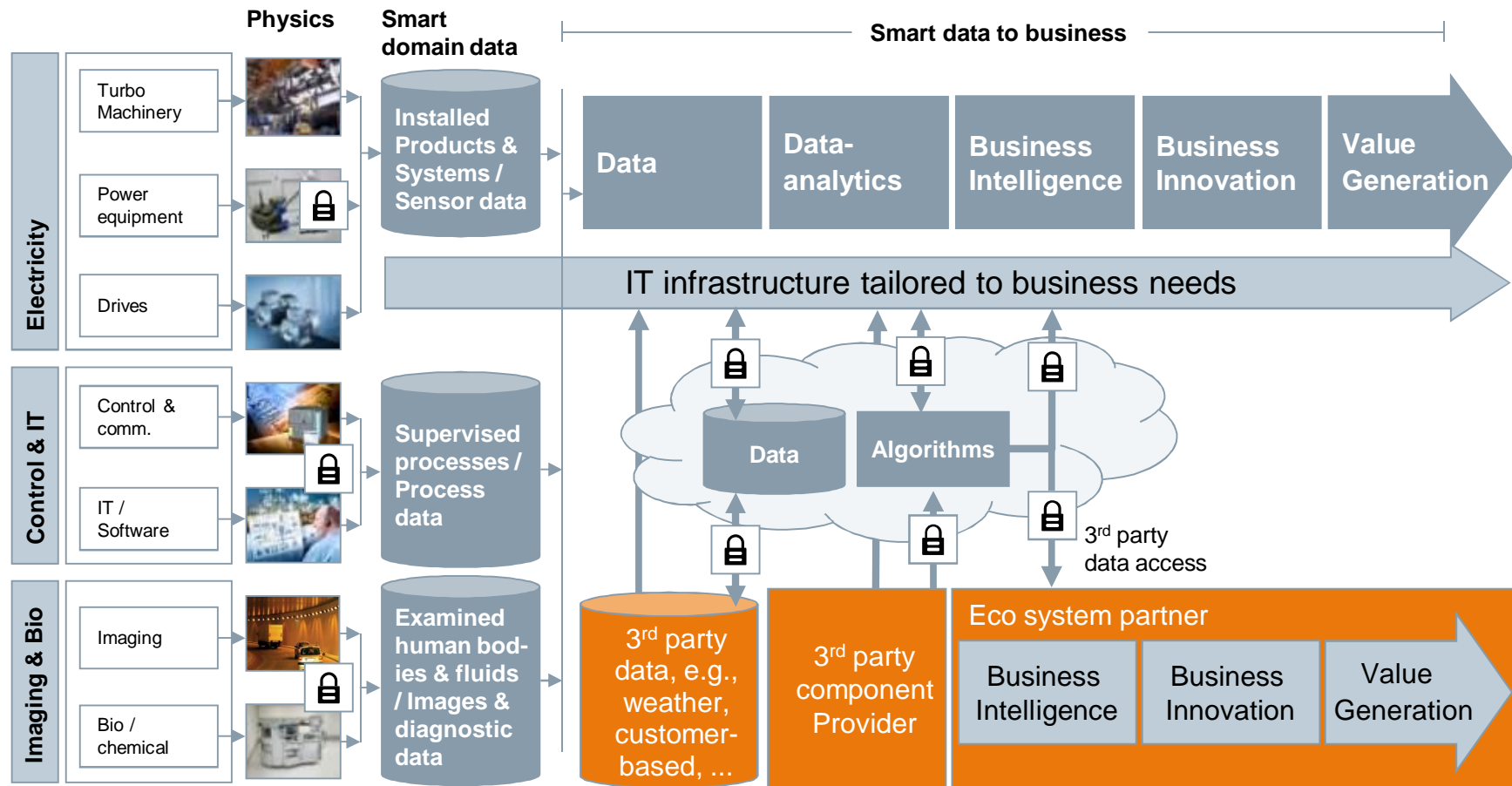
# Core know-how and technology for data analytics applications needed for Siemens offerings

**SIEMENS**

Smart data to business: Building blocks (examples)

| Physics | Data | Data Analytics | | | Business Intelligence | | Business Innovation | Value Generation |
|---------|------|----------------|---|---|----------------------|---|---------------------|------------------|
| **Sensors, process, etc.** | | **Data Integration** | **Model and Analyze** | **Visualize and Advise** | **Applications** | | **Business Models** | **Customer Value** |
| Sensor design | Sensor data integration | Modeling, data structure & integration | Machine learning, reasoning | Dashboards | Online cond. monitoring | Forecasting and control | Enhanced products / solutions | • Availability<br>• Productivity<br>• Additional revenue<br>• End customer satisfaction<br>• … |
| Sensor integration, placement, … | Data pre-processing (e.g. quality, timestamps) | Relational and NoSQL databases | Math. modeling, Optimization | Information retrieval | Resource efficiency (energy, water, waste) | Fleet intelligence | Services | |
| | External data e.g. internet, weather | | Data mining, Language processing | Dialog | Customer analytics | Logistics, production processes | Consulting | |
| | | | | | | Business process perf. mgmt. | Performance contracting | |

**Technology Consulting**

**IT Integration**

# Large, Heterogeneous, Complex

## Large volumes of information

→ data streams (raw operational data, events, problem-specific, spatial, temporal): e.g. **> 10 GB/day**

→ integration of databases required: e.g. **>10 data sources, > 300 TB**

→ complex queries and complex algorithms: e.g. **joints & analytics on-demand**

## Lack of standardization

→ heterogeneity of data sources/structures

→ data quality issues, inconsistent data types

→ various interfaces for data access: e.g. **APIs with OO access & SQL**

## Complex data access

→ no direct access to data by engineers possible, additional IT stuff required

→ treatment for missing and incomplete data required

## Challenges in Data Analytics
# High speed, Real-time, Operational decisions

**High speed analytics on historical data**

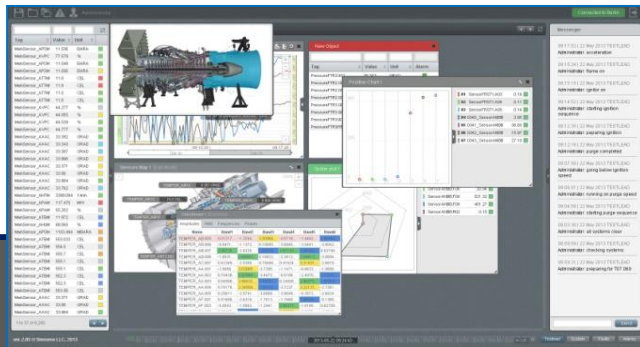→ Explore more "what if" alternatives to make better strategic, tactical and operational decisions

**Real-time analytics on real-time data**

→ continuous monitoring of events to enable faster and better responses to emerging operational threats and opportunities

**Run decision models in real time**

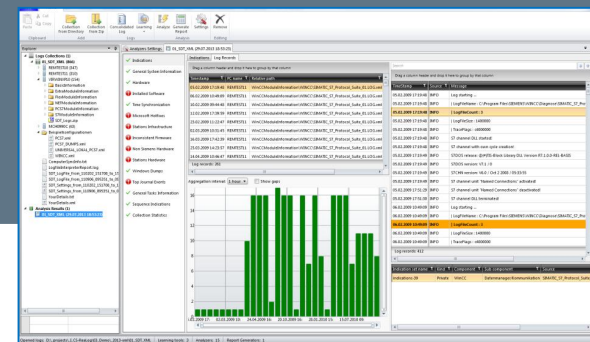→ involve sophisticated decision management into business processes to make smarter operational decisions

# Approach to the System Health Check of the LHC based on ELVis & WatchCAT @Siemens

- Data Analytics Framework
- Complex Event Processing
- Data Mining
- R-Extensible
- Time series data
- Streams, NoSQL
- Sensor Validation
- Web Visualization

**ELVis: Web-based Sensor Data Analytics**

**WatchCAT: Alarm Management & Event Monitoring**



**LHC experiments***



| Detector Control System | Gas Control System | Detector Safety System | Run Control |

Oracle DB (each experiment, infrastructure)

Electricity

**Accelerators***



Cryogenics

Vacuum

Magnet protection

Cooling & Ventilation

Interlocks

Access Control

LHC Logging

**CERN Infrastructure + Experiments**

# ELVis Platform:
# Real Time Data Analytics for CERN
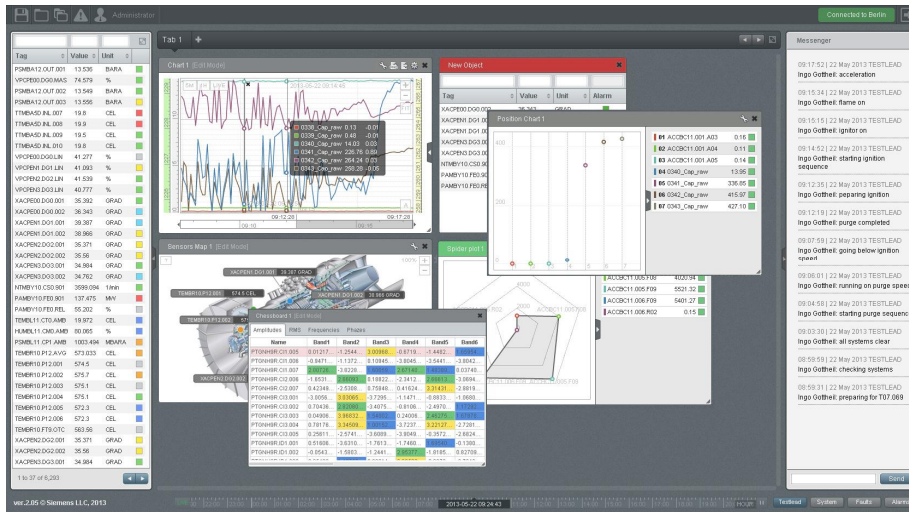


## Benefits

▪ Robust & powerful solution for sensor data analytics with impressive characteristics:

- 10-100 times more efficient than traditional web tools
- Real-time processing of 1000x of sensors per second
- No data loss: no single point of failure design
- Customizable algorithms for sensor data analysis
- Scalable from laptop to world-wide multi-site clusters

▪ Built-in intelligence for sensor data validation done without burden of model creating and learning

▪ Real-time HTML5 visualization of multiple high-speed data in a conventional browser

## Approach

Utilize real-time web technologies recently emerged due to multi-billion R&D investments of Internet giants (*Google, Facebook, Twitter, etc*)

Employ them in the field of industrial monitoring through ELVis - a web-based platform for processing, visualization and archiving multiple streams of time-series data from sensors.

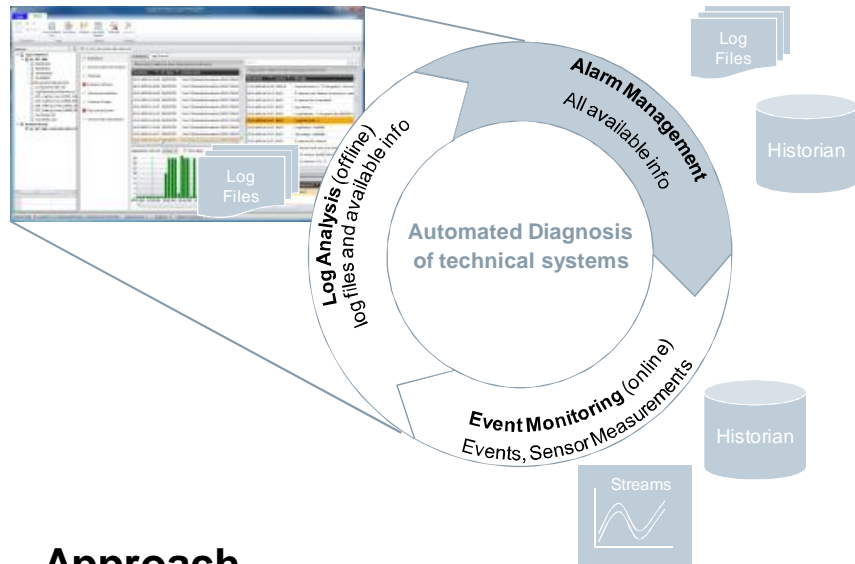Rich data processing and flexible API with support of many programming languages.

## Scope



Online monitoring
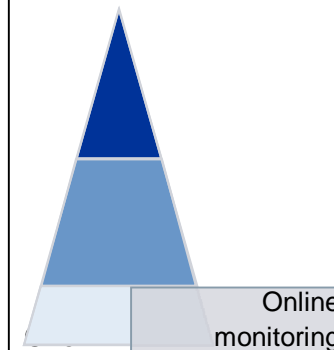
## Key Technologies

**Real-time dashboards**

**Data stream processing**

**Sensor validation**

# WatchCAT Data Analytics Framework:
# Alarm & Event based Diagnostics for CERN



Log Files

Historian

Alarm Management
All available info

Log Analysis (offline)
log files and available info

Automated Diagnosis
of technical systems

Event Monitoring (online)
Events, Sensor Measurements

Historian

Streams

## Benefits

- IEC 62682-compliant alarm management system for CERN using data from Siemens systems including WinCC OA

- Significant speedup of diagnostic process

- Effective alarm management based on both log and sensor information

- Relevant patterns are mined from data, no prior knowledge needed but the generated knowledge can be effectively re-used

## Approach

Integrated analysis of events and alarms coming from heterogeneous sources. Combined offline and online analysis (in work) based on a combination of :

- Data fusion of textual logs, events and sensor data

- Data interpretation using logical reasoning

- Automated sequence mining

- Event monitoring over sensor stream

- Complex event processing , R support

## Scope

Monitoring & Diagnosis

Status: prototype
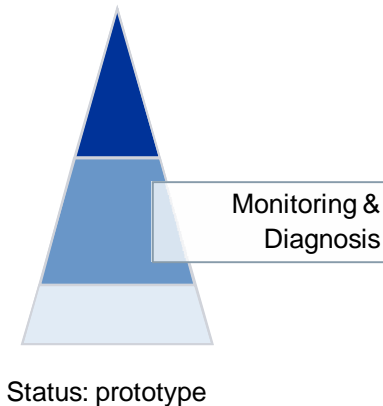
## Key Technologies

Knowledge modeling

Reasoning

Data stream processing

Data mining

Machine learning