Ref. Ares(2018)5271371 - 15/10/2018

Boost 4.0

BIG DATA FOR FACTORIES

Big Data Value Spaces for Competitiveness of European Connected Smart Factories 4.0

Horizon 2020 EU Grant Agreement 780732

Title	D3.3 - Big Data Models and Analytics Platform v1
Document Owners	Alexandros Nizamis - CERTH
Contributors	Konstantinos Sipsas - INTRA, Dimosthenis Ioannidis - CERTH, Silvia Castellvi - i2CAT, Damiano Manocchia - SIEMENS, Stefanie Kritzinger - RISC, Cedric Chappuis - ESI, Paulo Figueiras - UNINOVA, Serena Dondana – SAS, Paloma Taboada - TRIMEK, Sebastian von Enzberg – FhG, , Alicia Gonzalez -INNO, Fernando Ubis (VIS), Thanasis Naskos (ATLANTIS)
Dissemination	Public
Date	14/10/2018
Version	V1.0

Version history

09/07/2018	BOOST4.0 D3.3 v0.1 - ToC
31/07/2018	BOOST4.0 D3.3 v0.2 – State-of-the-art analysis
24/08/2018	BOOST4.0 D3.3 v0.3 – Introductory sections
25/09/2018	BOOST4.0 D3.3 v0.4 – Chapter 4 related to Task 3.5
27/09/2018	BOOST4.0 D3.3 v0.5 – Chapter 5 related to Task 3.6
28/09/2018	BOOST4.0 D3.3 v0.6 – Document updates and conclusions
28/09/2018	BOOST4.0 D3.3 v0.7 – Ready for review
11/10/2018	BOOST4.0 D3.3 v0.7 – Added VIS contribution, updated SAS Analytics section
14/10/2018	Final version

Document Fiche

Authors	Alexandros Nizamis – CERTH, Konstantinos Sipsas - INTRA, Dimosthenis Ioannidis - CERTH, Silvia Castellvi - i2CAT, Damiano Manocchia - SIEMENS, Stefanie Kritzinger - RISC, Cedric Chappuis - ESI, Paulo Figueiras - UNINOVA, Serena Dondana – SAS, Paloma Taboada - TRIMEK, Sebastian von Enzberg – FhG, Alicia Gonzalez -INNO, Fernando Ubis (VIS), Thanasis Naskos (ATLANTIS)
Internal Reviewers	INNOVALIA
Workpackage	WP3
Task	T3.5, T3.6
Nature	Report
Dissemination	Public

Project Partners

Participant organisation name	Acronym
Asociación de Empresas Tecnológicas Innovalia	INNO
Volkswagen Autoeuropa, Lda *	VWAE
Visual Components	VIS
Automatismos y Sistemas de Transporte Interno S.A.U.	ASTI
Telefónica Investigación y Desarrollo SA	TID
Volkswagen AG. *	VW
UNINOVA	UNINO
FILL GmbH. *	FILL
TTTECH Computertechnik AG	TTT
RISC Software GmbH	RISC
PHILIPS Consumer Lifestyle B.V. *	PCL
PHILIPS Electronics Nederland	PEN
Interuniversitair Micro-Electronicacentrum VZW	IMEC
Centro Ricerche Fiat S.C.p.A. *	CRF
SIEMENS S.p.A.	SIEMENS
Prima Industries S.p.A	PRIMA
Politecnico di Milano	POLIMI
AUTOTECH ENGINEERING, AIE *	GESTAMP
Fundació Privada I2CAT, Internet I Innovació Digital A Catalunyai2cat	I2CAT
TRIMEK S.A.	TRIMEK
CAPVIDIA N.V,	CAPVIDIA
Volvo Lastvagnar AB *	VOLVO
Chalmers Tekniska Hoegskola AB	CHAL
Whirlpool EMEA SpA *	WHIR
SAS Institute Srl	SAS
Benteler Automotive GmbH *	BAT
It.s OWL Clustermanagement	OWL
Fraunhofer Gesellschaft Zur Foerderung Der Angewandten Forschung E.V.	FhG
Atlantis Engineering	AE
Agie Charmilles New Technologies SA *	+GF+
Ecole Polytechnique Federale De Lausanne	EPFL
Institut Für Angewandte Systemtechnik Bremen GmbH	ATB
Rheinische Friedrich-Wilhelms-Universitat Bonn	UBO
Ethniko Kentro Erevnas Kai Technologikis Anaptyxis (CERTH)	CERTH

The University of Edinburgh	UED
Institute Mines Telecom	IMT
Industrial Data Space E.V.	IDSA
FIWARE Foundation EV	FF
GEIE ERCIM EEIG	ERCIM
IBM ISRAEL – Science and Technology LTD	IBM
ESI Group	ESI
Eneo Tecnología, S.L	ENEO
Software Quality Systems S.A.	SQS
Consultores de Automatización y Robótica S.A.	CARSA
INTRASOFT International	INTRA
United Technologies Research Centre Ireland, Ltd *	UTRC-I
Fratelli Piacenza S.p.A. *	PIA
RiaStone - Vista Alegre Atlantis SA *	RIA
Unparallel Innovation, Lda	UNP
Gottfried Wilhelm Leibniz Universität Hannover	LUH

*LHF 4.0 – Lighthouse Factory 4.0   * RF – Replication Factory 4.0

Executive Summary

The present document is a deliverable of the BOOST 4.0 – Big Data Value Spaces for Competitiveness of European Connected Smart factories 4.0 project, funded by European Union under Horizon 2020 Research and Innovation programme (H2020). The deliverable presents the first version of the Big Data Models and Analytics Platform report that contains the initial results of Task 3.5 Big Data Models for Cognitive Manufacturing and Task 3.6 Big Data Analytics Platform.

BOOST 4.0 aims to establish a set of European big data light-house smart connected factories. In order to achieve this, a set of data-driven cognitive manufacturing services will be developed. The advanced big data analytics and technologies supporting cognitive manufacturing processes will be deployed in the BOOST 4.0 Big Data Analytics Platform. The Platform from Task 3.6 will customize and extend existing digital manufacturing platforms and will support advanced cognitive models and data visualisation techniques developed in T3.5.

This report comes in an early stage of the project (M9) and the two involved tasks start at M4 and M7 respectively, so only some first results will be presented. The complete outcome of the Tasks 3.5 and 3.6 will be available at M24 when these tasks will be completed. The results will be drawn at the second iteration of this document, Big Data Models and Analytics Platform v2.

The report presents the first big data algorithms implemented across smart engineering, planning, operations, production and after sales to leverage high value data-driven operations. The report also presents the vision of smart data apps, services and platforms to leverage on the capabilities offered by the European Industrial Data Space (EIDS). The report presents the various big data analytic platform and discusses on the expected extensions to align big data analytics platform capabilities with the larger volumes and heterogeneous data streams that will be provided from IT, OT, IIoT Engineering and collaborative manufacturing contexts. The document introduces the expected big data capabilities of the boost 4.0 platform federation to implement hybrid twin, real time and collaborative interactive decision workflows for the benefit of factories 4.0.

Keywords: IDS Smart Data Apps, Big Data Analytics, Digital Manufacturing Platforms, Big Data Lakes, Hybrid Twin, early detection, predictive diagnostics, visual analytics, AI, IIoT, 3D point cloud, time series analytics, big data stream analytics, ETL.

Disclaimer

Acknowledgement

This document is a deliverable of Boost 4.0 project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 780732.

Table of contents

Abbreviations and Acronyms

Acronym	Meaning
A.I.	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
CEP	Complex Event Processing
DSS	Decision Support System
ERP	Enterprise Resource Planning system
FDT	Fault Detection Tool
IoT	Internet of Things
KPIs	Key Performance Indicators
ML	Machine Learning
MES	Manufacturing Execution System
PaaS	Platform as a Service
PMT	Predictive Maintenance Tool
RDBMS	Relational Database Management Systems
SIEM	Security Information and Event Management
SSA	Singular Spectrum Analysis
SMS	Smart Manufacturing Systems

List of Figures

List of Tables

1 Introduction

1.1 Scope

The scope of this deliverable is to describe the work that has been done and the research has been conducted in Task 3.5 Big Data Models for Cognitive Manufacturing and Task 3.6 Big Data Analytics Platform as well.

Boost 4.0 project is organised as shown in the Figure below. WP3 is setting up and aligning digital infrastructures, big data service marketplaces and big data platform capabilities to the Boost 4.0 smart data and European Industrial Data Space principles for data sovereignity.

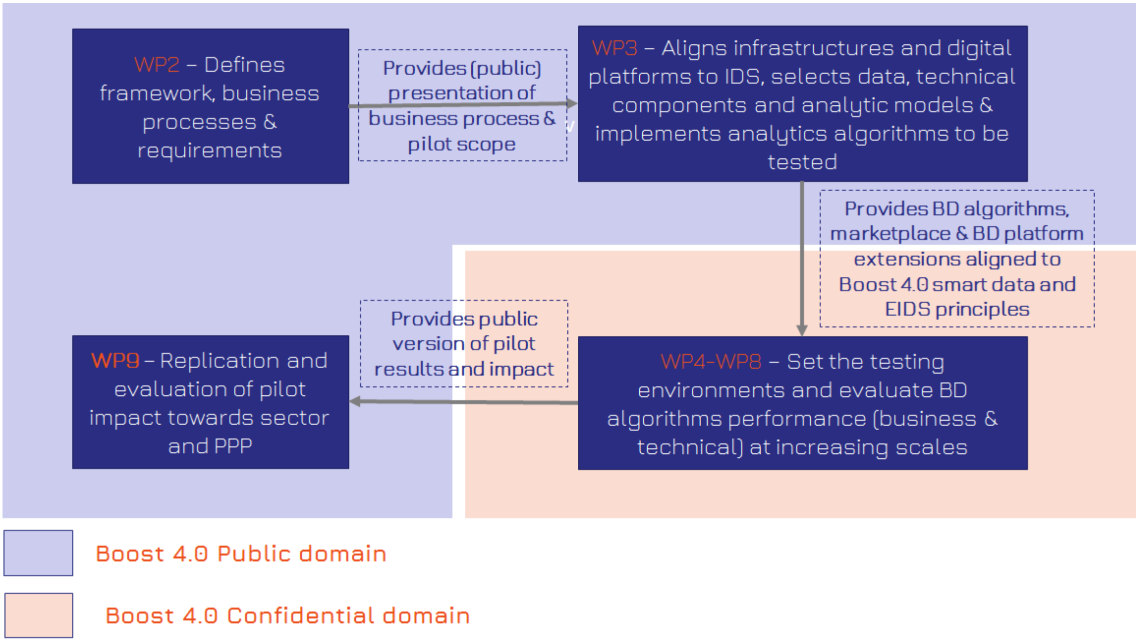

Figure 1 BOOST 4.0 Project Workpackage organisation

Predictive data models, forecasting algorithms, machine learning techniques, visual analytics tools etc. that will be developed in Task 3.5, they will be applied and tested in the pilot cases alongside with the tools, and the supporting analytics platform from Task 3.6..

The outcome of these two tasks will be strongly correlated with the rest of the developed components of WP3. Many of implemented solutions of Tasks 3.5 and 3.6 will be delivered as IDS apps and services. The use of IDS architecture, Vocabularies, IBM Hyperledger Fabric and FIWARE connectors will enable the creation of data-driven, secure and novel services for the BOOST 4.0 Online Collaborative Analytics Service Marketplace. These will enable apps injection to connectors to add the provided services to the top of the data exchange. The provided services will be related to data processing and analytics and will enable the remote execution of algorithms over the Marketplace. By using the aforementioned technologies and tools of the BOOST 4.0 platform, the developed data analytics services will rely on security, data sovereignty and standardized interoperability.

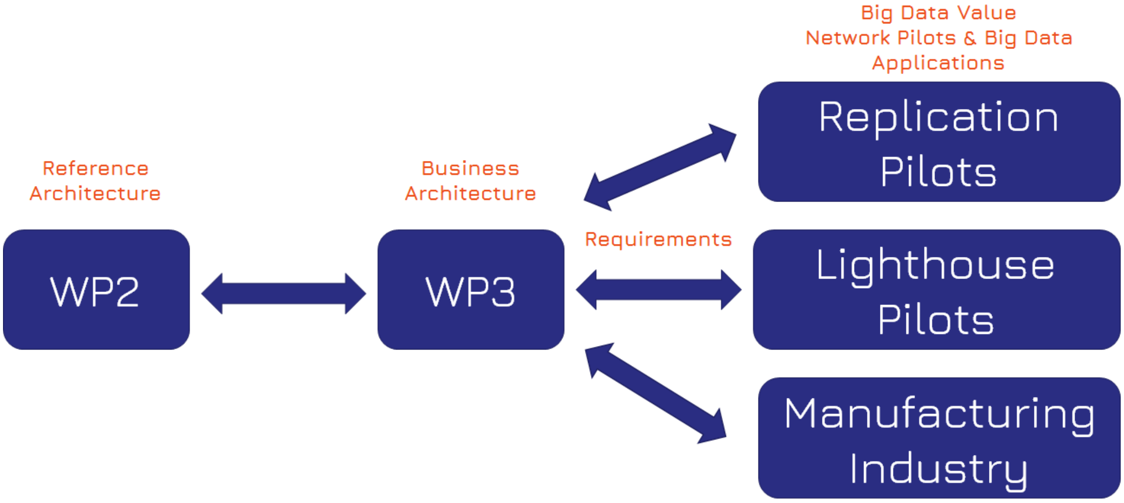

Figure 2 BOOST 4.0 project Internal and external dependencies

1.2 Document Structure

After the introductory sections, this deliverable is structured as follow: Section 2 includes a thorough state-of-the-art analysis in the field of predictive and cognitive data models, fault diagnosis methodologies, and prediction and forecasting techniques in accordance with visual analytics techniques.

Section 3 describes predictive and cognitive data models, fault diagnosis methodologies, and prediction and forecasting techniques that are under development or they are already developed in the BOOST 4.0 project. Moreover, their relevance with the pilot cases in which they are going to be applied and tested will be presented as well.

Section 4 is devoted to the Big Data Analytics Platform and its tools. Details about platform's development or about the extensions of already existing platforms are described in this section. Furthermore, the way that the implemented tools support and ensure data sovereignty is documented in this section as well.

Conclusions and next steps related to the activities of Tasks 3.5 and 3.6 are drawn at the closing section of this document.

2 Big Data Models for Cognitive Manufacturing – State-of-the-Art Analysis

BOOST 4.0 project aims to establish a set of European big data light-house smart connected factories based on the big amount of data coming from these factories. In order to achieve this, a set of data-driven cognitive manufacturing services and tools will be developed and applied in these data. In this first period of the Project a thorough analysis have been conducted in the fields of cognitive manufacturing, predictive modeling, faults and their root cause diagnosis, and visual analysis for big data. It was considered as an urgent need to conduct this analysis before the beginning of any implementation related to the aforementioned fields as the Task 3.5 aims to offer the best available solutions to the pilot partners of the project in the part of data analysis, and to contribute beyond the state-of-the-art in the fields of cognitive manufacturing, AI and machine learning.

2.1 From Cognitive Modeling to Cognitive Manufacturing

The idea of cognitive modeling is founded on the assumption that human intelligent behavior involves computation. Chronologically, the idea of cognitive modelling concurs with the beginning of computers. Cognitive modelling nowadays is an area of computer science that deals with simulating human problem-solving and mental processing in a computerized model. Such a model can be used to simulate or predict human behavior or performance on tasks similar to the ones modeled and improve human-computer interaction. Indicative techniques that support the above role are [1, 2]

The adoption of the Internet of Things (IoT) in industry makes the factories more instrumented and interconnected than ever, resulting in more complex systems with increased requirements in monitoring, reliability, risk management and decision-making. IoT provides the ability to gain valuable data of all the aspects in our factories. Furthermore, the amount of data produced and communicated via IoT must be analyzed through analytics to identify patterns in the data, model behaviors of equipment and predict failures based on a variety of variables that exist in the manufacturing setting. Another key factor towards an automated production, in which cognitive modeling can play a major role, is the capability for a factory to be flexible, adaptive and reliable in the momentary situations and unforeseen conditions, in order to derive an efficient production scheme and aid the decision makers. The reasoning behind the need in using cognition modeling in manufacturing is inspired from the fact that humans are capable of acting competently under uncertainty, reliably handling unpredicted events and situations and quickly adapting to changing tasks, capabilities and environments [3, 4]. Therefore, the facts of life, induced by the rapid evolution and exploitation of technology and information, force us to broaden the limits of knowledge and techniques of cognitive modeling, in order to serve the demands of the automated manufacturing. The research in cognitive manufacturing the past decade has introduced interesting and innovative ideas on how artificial cognition could be incorporated in industry, combining the field of IoT, psychology, machine learning, data mining and artificial intelligence in the realm of the Big Data era. Such ideas are also discussed in [5, 6].

2.2 Early detection – Predictive modeling

The anomaly detection problem, in its most general form is not easy to solve and most of the existing anomaly detection techniques solve a specific formulation of the problem. A key aspect of any anomaly detection technique is the nature of the input data. The form and the nature of attributes of each data instances, can determine the applicability of anomaly detection techniques. In general, most researches have adopted techniques, such as machine learning, data mining, information theory and statistics so as to deal with this crucial problem. In literature, one can find many approaches on anomaly detection such as classification [7, 8,10], clustering approach [7, 8, 9], statistical approach [11, 12], link analysis [14], "Dynamic" event-stream processing approach [16]. Anomaly detection techniques are applied in many areas such as virtualized computer machines [16], communication and social networks [13, 14] and urban data [15].

New techniques, innovating algorithms and modifications of existing ones tailored to the needs and characteristics of the to-be-addressed use case, are utilized either on historical off-line or in real time data. These techniques will continually examine various conditions of the collected data, such as normality and linear trend maintenance through the calculation of linear trend profile of monitored features [17, 18]. Another interesting technique is Complex Event Processing (CEP), because CEP is intended to manage data in motion and therefore is useful for in real time data situations. Complex Event Processing is a technique for tracking, analyzing, and processing data as an event happens. This information is then processed and communicated based on business rules and processes. The idea behind CEP is to be able to establish the correlation between streams of information and match the resulting pattern with defined behaviors such as mitigating a threat or seizing an opportunity. CEP is an advanced approach based on simple event processing that collects and combines data from different relevant sources to discover events and patterns that can result in action. CEP techniques such as [16, 19, 20] can play a major role in detection of situations of interest that are known in prior to be problematic. The main objective of this process is to find possible deviations from normal conditions, detect for malfunctions and failures during industrial production processes and to point for alerts. For this part of detection, prior knowledge of the behavior of the data, will not take into account. As for off-line and real-time predictive modeling, a series of known and popular machine learning techniques, such as Support Vector Machines [21], Decision Trees [22], Random Forest [23], Back-propagation network [24] and their boosting versions, such as Adaboost. SAMME algorithm for ensemble learning on two-class and multi-class classification scenarios [25] are commonly used in order to prevent future failures and abnormalities, to find hidden insights on various operations of the plant and to provide reliable decisions and results. A series of feature extraction techniques, such as Principal Component Analysis, Canonical Correspondence Analysis, Mutual information, are applied as a pre-processing step so as to find those features that contain the most useful information and will improve performance of predictive models. Moreover, in order to understand (and trust) the predictive models, researchers use Local Interpretable Model-Agnostic Explanations (LIME), a technique that explains the predictions of any machine learning classifiers [26, 27]. Another technique that aids in understanding the model behavior via influence functions is discussed in [28]. Artificial neural networks have been widely used in research and for practical applications since the early 80's. The evolutionary artificial neural networks utilize evolutionary algorithms like the most known Genetic Algorithm [29] so as to provide an alternative approach on process optimization. Deep learning techniques suitable for the form of existing data, such as fully-connected and convolutional neural networks [30] are the one-step-forward in machine learning, as an attempt to improve and enhance predictive performance.

2.3 Diagnosis of faults and their root cause

Understanding the root cause of an observed symptom in a complex system has been a major problem for decades. The dramatic evolution in automation of industry and the exponential increase of the amount and velocity of data to support complex human decisions, demands the continuous increase in complexity and expense of industrial systems. This fact sets stricter principles by means of tolerance for performance degradation, productivity decrease, and safety hazards, which makes detection and identification of potential abnormalities and faults as early as possible and real-time fault-tolerant operation a necessity [31, 32]. The diagnosis of fault can be divided into two main categories: model-based or signal-based.

2.3.1 Model-Based Fault Diagnosis Methods

Model-Based Fault diagnosis was firstly introduced by Beard [33] in 1971 in order to substitute the hardware superfluity with analytic techniques. Fault diagnosis algorithms that have as a basis the existence of an analytical model are used in order to monitor the consistency between the outputs of a physical system and the outputs that the model predicts. There are two broad families of models for root cause analysis: Deterministic models and Probabilistic models. In deterministic models, there are no uncertainty in the known facts or the inferences expressed in the model. On the other hand, probabilistic models are able to handle this uncertainty. Deterministic models can be further divided into three subcategories based on the technique is used for the model. The subcategories are the Logic, Classifier and Process Model. Probabilistic models are also further divided into four subcategories. The subcategories are the Logic, Bayesian, Classifier and Process model.

2.3.2 Deterministic and Probabilistic models

This section contains a brief description of the most common approaches of deterministic models: logic, classifier and process models. There are various existing implementations suited for logic models such as Propositional Logic [34, 35, 36], First-order Logic [37, 38] and Fault Tree [39, 40]. Implementations suited for classification models include Decision Trees, SVM [51, 52] and Neural Networks [53, 54]. Processes in industrial environment becoming more complex and less tolerant to the consequences of faults, because the risks and price of faults are higher. In order to reduce these faults, an analytical model for fault diagnosis needs to be interfered in the process. Existing process-model implementations include Automata/FSM [55] and Petri Nets [56].

This section contains a brief description of the most common approaches of probabilistic models: logic, classifier and process models. Logic modelling or qualitative modelling is appropriate in processes whose behavior can be observed merely as a sequence of events. A discrete–event description of the system used and the diagnosis is fulfilled by comparing the observed event sequence with the discrete–event dynamics of the model. There are various implementations suited for logic models such as Fuzzy Logic [57], Dempster-Shafer theory [58], Fuzzy Fault Tree [59] and Possibilistic Logic [60] implementations. There are various implementations suited for classification models such as Bayesian MSVM and Probabilistic Neural Network algorithms. In [61], is presented an implementation for Bayesian MSVM and in [62] is presented an application for fault diagnosis for power circuits using Bayesian MSVM. Bayesian implantation includes Bayesian Networks [63, 64], Probabilistic Relational Models [65], Markov Logic Networks [66, 67], Hidden Markov Models [68], and Relational Sum-product networks [69]. There is a wide range of implementations, addressing the field of probabilistic process-model based fault detection. The State of The Art of probabilistic process modeling in fault detection is represented by implementations of Stochastic Petri Nets [70] and Stochastic DES [71].

2.3.3 Signal-Based Fault Diagnosis Methods

In contrast to model-based methods, signal-based methods are not utilized via models that take an input and predict an output to be compared with the input. Signal-based methods utilize measured signals, which represent the fault and subsequently a diagnostic decision is made based on the symptom analysis. The nature of the signal, either time-domain or frequency-domain, that is about to be processed, points to the signal-based method that fits to the specified signal. The three families of signal-based method for analysis are: time-domain [72, 73], frequency-domain [74, 75], time-frequency domain [76, 77].

2.3.4 Reliability Block Diagrams and System Reliability

Reliability could be addressed as ''the probability of a device performing its purpose adequately for the period of time intended under the operating conditions encountered'' [78]. Reliability is a necessary feature for in mission-critical domains such as aerospace, military, manufacturing and power industries. A reliability block diagram is introduced in order to represent graphically how the components/subsystems of the system are reliability-wise related. The Reliability Block Diagram (RBD) can be used in both the design and operational phase, in order to identify potential areas of poor reliability and where improvements can be made to lower the failure rates for the equipment.

2.4 Big data and Visual analytics

In modern manufacturing organizations, the challenge of collecting sufficient data for the efficient control of processes has shifted to analyzing vast and continuously changing amounts of data for the extraction of useful information in real time [79]. A wide spectrum of sensors has been spread across industrial environments over the last decades causing the explosion of information in such environments in terms of volume, velocity, variety and veracity; categorizing this issue under big data analytics. Unfortunately, most manufacturing companies do not make good use of all the generated and collected data [80]. Smart Manufacturing Systems (SMS) are foreseen as the solution to this through smart technologies implementing novel real-time control & data analytics solutions [81].

Past research in databases and information retrieval has focused on storage, search and retrieval of information, functionalities that cannot cover the current need for the automated extraction of knowledge from big data resources [82]. On the basis of this gap, and in focus of the related issues and challenges, big data analytics has been continuously evolving and targets a variety of applications including manufacturing [83]. The analysis of data of sheer volume and dimensionality remains one of the main challenges in this field of research coupled with the fast data rates of modern measurement and assessment mechanisms.

Visual analytics has historically played a key role in business processes optimization. Existing tools [84-86] can be of great assistance for the visualization of spatiotemporal data in the shop floor, providing e.g.: temporal plots and heat maps indicating specific types of activities; representation of movement data joint with statistical analysis suitable to assist in discovering patterns and correlation; visualization of business processes detailing communication activities and summaries in addition to geospatial distribution and scheduling.

3 Big Data Models and Analytics Services In BOOST 4.0 Project

The BOOST 4.0 project aims to provide services and tools related to descriptive, predictive and prescriptive analytics and deal with the application of big data at different stages of the factory life cycle and the supply chain. BOOST 4.0 plans to contribute on the state-of-the-art in cognitive manufacturing using the latest available techniques and methodologies in big data modeling, predictive and prescriptive analytics. The services related to predictive and prescriptive models, and data and visual analytics that will be available through the BOOST 4.0 Online Collaborative Analytics Service Marketplace to the pilot partners or other vendors from the same domain and even from different domains categorized as follow:

- Services, tools and algorithms related to the **detection** of deterioration rate of production machines and their root cause. Model-based Fault diagnosis methods such as deterministic models and probabilistic models will be designed or adopted. Besides the Model-based Fault diagnosis methods Signal-based Fault diagnosis models will be offered as well.
- Services, tools and algorithms related to **prediction** of production of defected products, based on the modelling of production assets' deterioration rate. Predictive modeling and early detection algorithms will be developed and used in order to add these types of services to the BOOST 4.0 Marketplace. Traditional techniques such as machine learning, data mining, information theory and statistics able to deal with this crucial problem will be offered. Moreover, innovative algorithms and new techniques such as Complex Event Processing (CEP) analysis techniques and trend analysis will be designed and be available to the Marketplace as well.
- Services, tools and algorithms related to the **advanced data visualization** and visual analytics. New ways of interaction and display technologies to support analytical reasoning and collaborative decision making in the big data available to the Big Data Value Space of the BOOST 4.0 will be adopted.

The use of the above-described methodologies and cognitive modeling for a cognitive manufacturing will enable the creation of services that will predict unforeseen conditions during the production process and will boost the real-time decision-making. In this chapter, the aforementioned methodologies are analyzed in accordance with the BOOST 4.0 pilots in which these methodologies will be applied and tested before become available to BOOST 4.0 Online Collaborative Analytics Service Marketplace.

3.1 Industrial Data Space apps

The IDS Connector Architecture as discussed in Boost 4.0 deliverable D3.1, uses Application Container Management technology to ensure an isolated and secure environment for individual data services. To ensure privacy of sensitive data, data processing should take place as close as possible to the data source. Any data preprocessing (e.g., filtering, anonymization, or analysis) should be performed by Internal Connectors. Only data intended for being made available to other participants should be transferred to External Connectors. Data Apps are services encapsulating data processing and/or transformation functionality bundled as container images for simple installation by Application Container Management.

The IDS reference model distinguishes among three types of data apps:

- self-developed Data Apps, which are used by the Data Provider's own Connector (usually requiring no certification from the Certification Body),
- third-party Data Apps, which are retrieved from the App Store (and which may require certification), and
- Data Apps provided by the Connector of the Data Consumer, which allow the Data Provider to use certain functions before data is exchanged (e.g., filtering or aggregation of data) (and which may also require certification).

In addition, data apps can be divided into two more categories:

- **System Adapters** are Data Apps on the Data Provider side, establishing interfaces to external enterprise information systems. The main task of a Data App belonging to this category (in addition to wrapping the enterprise information system and perhaps transforming from an internal data model to a data model recommended or standard for a given application domain) is to add metadata to data.
- **Smart Data Apps (or Data Sink Connectors)** are Data Apps on the Data Consumer side, executing any kind of data processing, transformation, or storage functionality.

Normally, the data provided from, or sent to, a Smart Data App is already annotated with metadata (as described in the Information Layer section). Using an integrated index service, the Broker manages the data sources available in the Industrial Data Space and supports publication and maintenance of associated metadata.

Furthermore, the Broker Index Service supports the search for data sources. Both the App Store and the Broker are based on the Connector Architecture (which is described in detail in the following paragraphs). The Figure below illustrates the internal structure of the Connector.

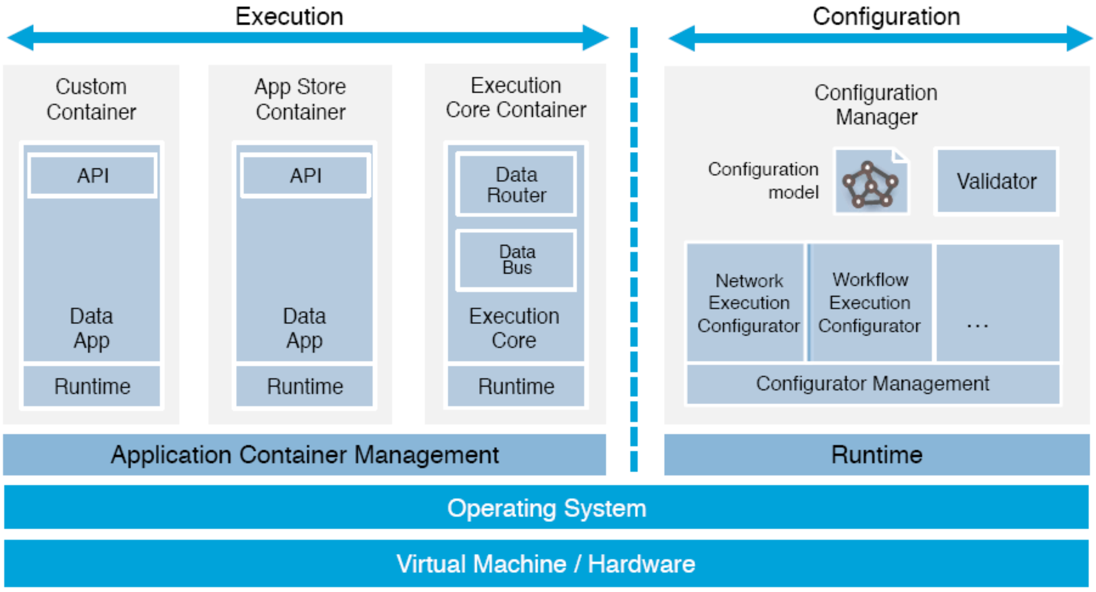

Figure 3 Reference Architecture of an IDS Connector

A concrete installation of a Connector may differ from this structure, as existing components can be modified and optional components added. The components shown in the Figure above can be assigned to two phases: Execution and Configuration.

The **execution phase** of a connector involves the following components:

- **Application Container Management:** In most cases, the deployment of an Execution Core Container and selected Data Services is based on application containers. Data Services are isolated from each other by containers in order to prevent unintended interdependencies. Using Application Container Management, extended control of Data Services and containers can be enforced. During development, and in case of systems with limited resources, Application Container Management can be omitted. Difficulties in container deployment can be handled by special Execution Configurators
- **An Execution Core Container** provides components for interfacing with Data Services and supporting communication (e.g., Data Router or Data Bus to a Connector).
- **A Data Router** helps configure Data Services to be invoked according to predefined configuration parameters. In this respect, it is responsible of how data is sent (and received) to (and from) the Data Bus from (and to) Data Services. Participants have the option to replace the Data Router component by alternative implementations of various vendors. Differences in configuration can be handled by specialized Execution Configurator plug-ins. If a Connector in a limited or embedded platform consists of a single Data Service or a fixed connection

configuration (e.g., on a sensor device), the Data Router can be replaced by a hard-coded software, or the Data Service can be exposed directly.

- **The Data Bus** exchanges data with Data Services and Data Bus components of other Connectors. It may also store data within a Connector. Usually, the Data Bus provides the method to exchange data between Connectors. Like the Data Router, the Data Bus can be replaced by alternative implementations in order to meet the requirements of the operator. The selection of an appropriate Data Bus may depend on various aspects (e.g., costs, level of support, throughput rate, quality of documentation, or availability of accessories).
- **An App Store Container** is a certified container downloaded from the App Store, providing a specific Data Service to the Connector.
- **A Custom Container** provides a self-developed Data Service. Custom containers usually require no certification.
- **A Data Service** defines a public API, which is invoked from a Data Router. This API is formally specified in a meta-description that is imported into the configuration model. The tasks to be executed by Data Services may vary. Data Services can be implemented in any programming language and target different runtime environments. Existing components can be reused to simplify migration from other integration platforms.
- **The Runtime of a Data Service** depends on the selected technology and programming language. The Runtime together with the Data Service constitutes the main part of a container. Different containers may use different runtimes. What runtimes are available depends only on the base operating system of the host computer. From the runtimes available, a service architect may select the one deemed most suitable.

The **configuration phase** of a connector involves the following components:

- The **Configuration Manager** constitutes the administrative part of a Connector. Its main task is the management and validation of the Configuration Model, followed by deployment of the Connector. Deployment is delegated to a collection of Execution Configurators by the Configurator Management.
- **The Configuration Model** is an extendable domain model for describing the configuration of a Connector. It consists of technology-independent, inter-connected configuration aspects.
- **Configurator Management** loads and manages an exchangeable set of Execution Configurators. When a Connector is deployed, the Configurator Management delegates each task to a special Execution Configurator.
- **Execution Configurators** are exchangeable plug-ins which execute or translate single aspects of the Configuration Model to a specific technology. The procedure

of executing a configuration depends on the technology used. Common examples would be the generation of configuration files or the usage of a configuration API. Using different Execution Configurators, it is possible to adopt new or alternative technologies and integrate them into a Connector.

- **The Validator** checks if the Configuration Model complies with self-defined rules and with general rules specified by the Industrial Data Space, respectively. Violation of rules can be treated as warnings or errors. If such warnings or errors occur, deployment may fail or be rejected.

As the Configuration phase and the Execution phase are separated from each other, it is possible to develop, and later on operate, these components independently of each other. Different Connector implementations may use various kinds of communication and encryption technologies, depending on the requirements given.

Following the principles of the reference IDS connector and availability of (smart) data apps, the Big Data Model and Analytics Platform approach takn by Boost 4.0 will rely on the principles shown in the picture below, where big data platforms rely on smart big data algorithms to deliver advanced functionalities; which are therefore ready to support and operate trusted smart data apps distributions from an IDS App store; e.g. Boost 4.0 marketplace.

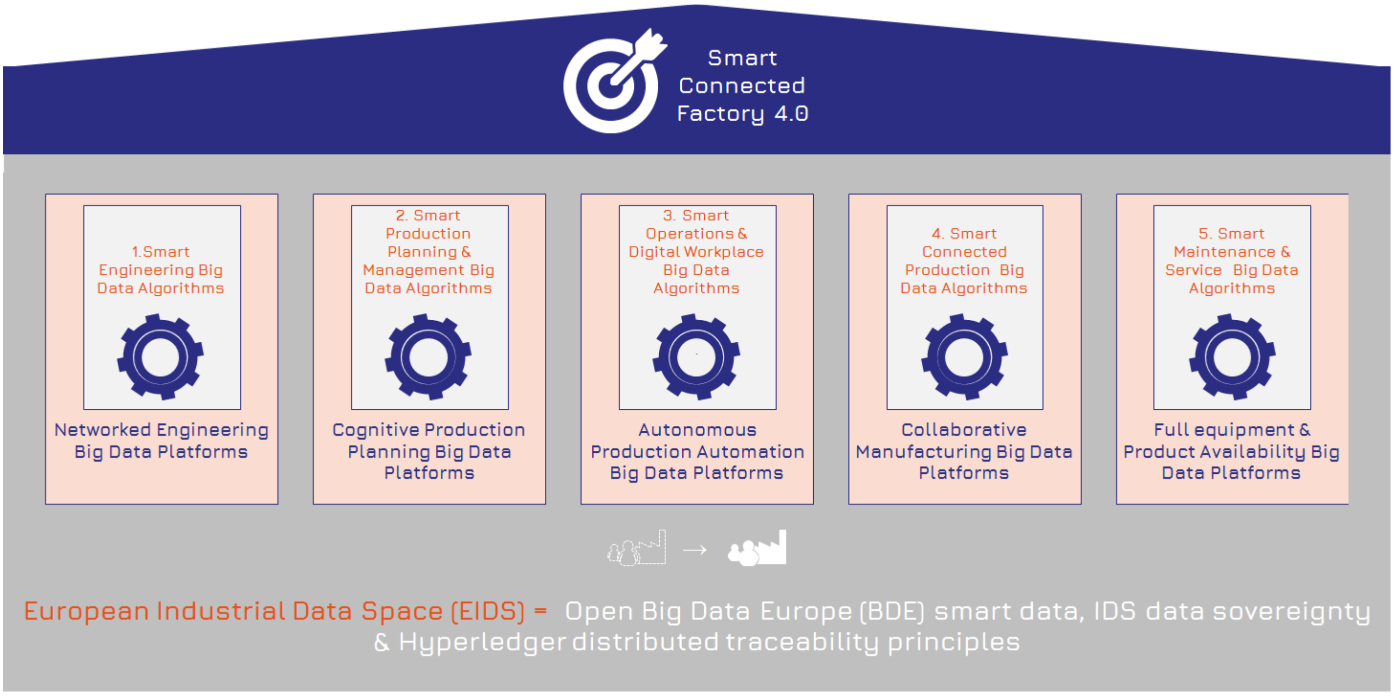

Figure 4 BOOST 4.0 Smart Data App Framework

The following Sections describe the various big data algorithms that will be developed to address the big data value needs from the current Boost 4.0 pilots. Algorithms will cover all phases of the data lifecycle (smart engineering/commissioning, planning, management, operations, production, maintenance and after sales services).

3.2 Big Data Models, Techniques and Methodologies for Smart Digital Engineering Pilots

In order to support the creation of Smart Digital Engineering pilots, the BOOST 4.0 project will use and implement innovative Big Data cognitive manufacturing processes, predictive methodologies and algorithms that enable:

- Predictive maintenance based on closed-loop feedback for zero defect factory tooling commissioning
- General engineering approach for highly flexible machines manufacturing optimization

3.2.1 Predictive Maintenance for Injection Moulding Plant

3.2.1.1 Overview

BOOST 4.0 project aims at developing new Industry 4.0 related approaches and systems for one of the world's leading automobile manufacturers and the largest global carmaker in terms of sold vehicles, the Volkswagen Group. The BOOST 4.0 approaches will be deployed at Volkswagen component tool shop in Braunschweig, internal provider of tools for light metal casting, injection moulding and forming processes for many other factories each with a large number of machines.

Continuous increasing requirements towards product design and quality requirements and therewith narrowing process windows demand feedback to design & manufacturing of tools as well as prevention of tool and/or machine related issues by prediction are key for effective operation of complex processes such as light metal casting, injection moulding, additive manufacturing or complex sheet metal related work.

The component tool shop represents the key department in a circular relationship between the design of cars and the production of related steps (i.e. pre-series production & test, regular production, maintenance, infrastructure management). The component tool shop as kernel department is already operating some 40 milling machines, 20 turning & sanding machines and 5 eroding machines that are used to manufacture diverse tools for additional types of manufacturing processes (i.e. tools for metal casting, injection moulding, assembly and metal press). Those different manufacturing processes are imposing different challenges on how to use data, for generating information and knowledge.

Especially metal casting is a manufacturing process that is operated at high temperatures (i.e. in ranges from 660-1500 °C) where process parameters cannot be easily monitored by usual sensors and measurement approaches. At the same time, there is rather limited knowledge available on the usage of measured process parameters for monitoring and adjusting the process directly. On top of that, diverse sources of untapped data needs to be processed and analysed for being able to find potential implications of design decisions with respect to the automotive part, used alloy, mould and machine. However, the current manufacturing process is somehow a black box handled rather based on past designs and engineers' profound technical competency. The metal casting process itself does not use detailed deterministic real-time process control models. The motivation is to develop related models based on an evolutionary learning from data collected in different life cycle phases of the manufacturing process. Therefore, the use case shall correlate the data that can be collected in the different product and process life cycle phases (e.g. mould design, machine set-up, tool & machine operation, maintenance as well as quality monitoring w.r.t. the automotive part).

Therefore, in BOOST 4.0 VW aims at developing new Industry 4.0 related approaches and systems for gathering large amounts of data for being able to understand, diagnose, predict and finally to prescribe the conditions in the manufacturing process for an optimal overall equipment effectiveness. This requires an integrated approach for modelling information, gathering data, deriving information & knowledge as well as to transfer the approach to diverse departments locally and at diverse locations and possibly even transferring related knowledge to first tier suppliers. Finally, for being able to generate knowledge that will facilitate the achievement of a zero defect objective it is envisaged to develop a modular improvement approach that allows for an evolutionary introduction in heterogeneous manufacturing environments. It shall be possible to use the approach in departments that are operating old machines and tools with rather limited abilities for data provision as well as with machines, moulds and tools that are already able to provide rich data sets in real time.

3.2.1.2 Big Data Models, Techniques and Tools for a Zero Defect Component Tool Shop

The BOOST 4.0 solutions for the VW tool shop aims to provide to objective scenarios as they are depicted in the following figure.

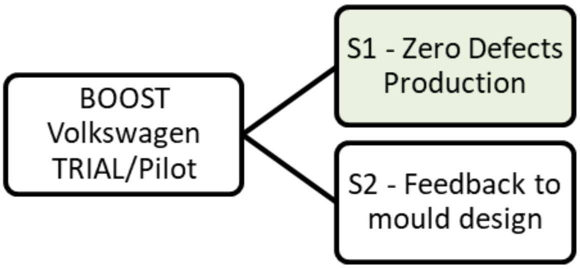

Figure 5 BOOST 4.0 Objectives for VW Tool Shop

For reaching these objectives, two approaches will be followed responding each one to the specific VW need:

The *first method*, for the '*Zero Defects Production*' goal, is dedicated to a better control of the casting process during production (in process).

For this need, a Hybrid Twin approach based on data analytics tools embedded in ESI Simulation platform for casting, and sensor data coming directly from the casting process will be adopted.

The workflow would be:

1. Realize a first set of simulations with some evolving parameters.
2. Classify these simulation results in order to get a full understanding of the casting process. This will lead to the creation of a vector of information characterizing each of these simulations. It will look like a DNA map of simulations (see figure below).
3. Compare sensors data in real-time with the known behaviours identified in the DNA map, and take a decision based on this comparison.

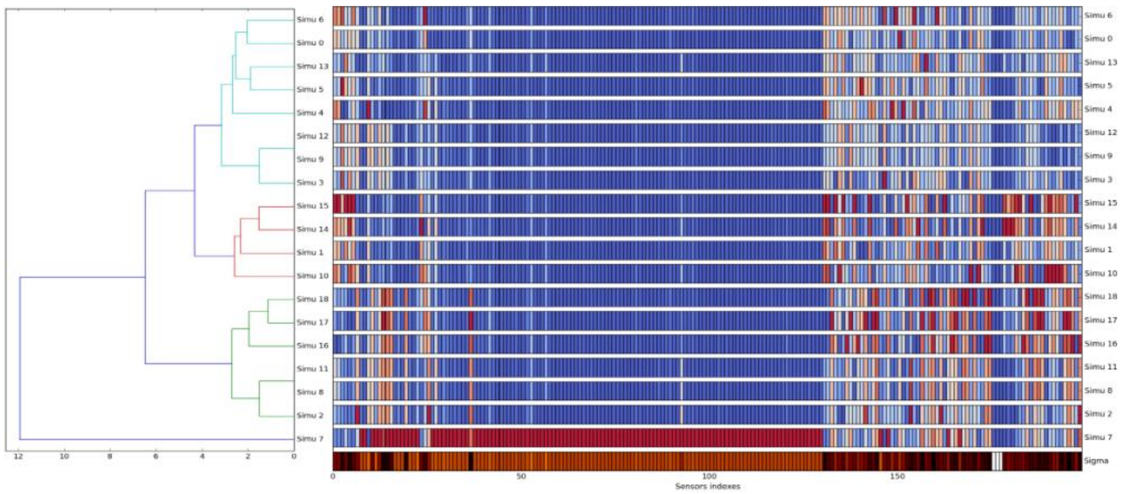

Figure 6 Example of classification (hierarchical clustering) on left, and DNA map of simulations on the right

This work will mainly consist to managing and manipulating big databases of simulation results (temperature / porosity / velocity / …)

The total amount of data available coming from the simulation software (between 50Gb and 100Gb of data for each simulation) may vary according to the number of simulations required to get the most representative DNA map. Also, some advanced numerical techniques will be deployed with the purpose to speedup the casting simulation process.

The *second method*, for the '*Feedback to mould design*', is focused on the root cause failure inspection (out process) based on a big data visualization tool (INENDI Inspector).

This tool will allow to track the production process of any casting part by navigating through data bases collected during the lifecycle of the product. That means for example, if the traceability of a part is well filled during its all lifecycle, we must be able to find deviations between good / bad quality parts. In the case of the suspected deviation appears during the casting process, then it will be possible to confirm by investigating the collected casting sensor datasets of the defected part (sensor data are stored in a historian tool) with the DNA map of simulations (The first method proposed must be deployed too).

The INENDI Inspector software solution allows any user to handle large amounts of structured or semi-structured data and perform a very rich and deep investigation of the data. Thanks to a set of intuitive and very interactive visualizations of this dataset, the user experience is freed from the classical burden of query-based interactions and the latter can use all his/her logical and deductive power to learn valuable insights from the data.

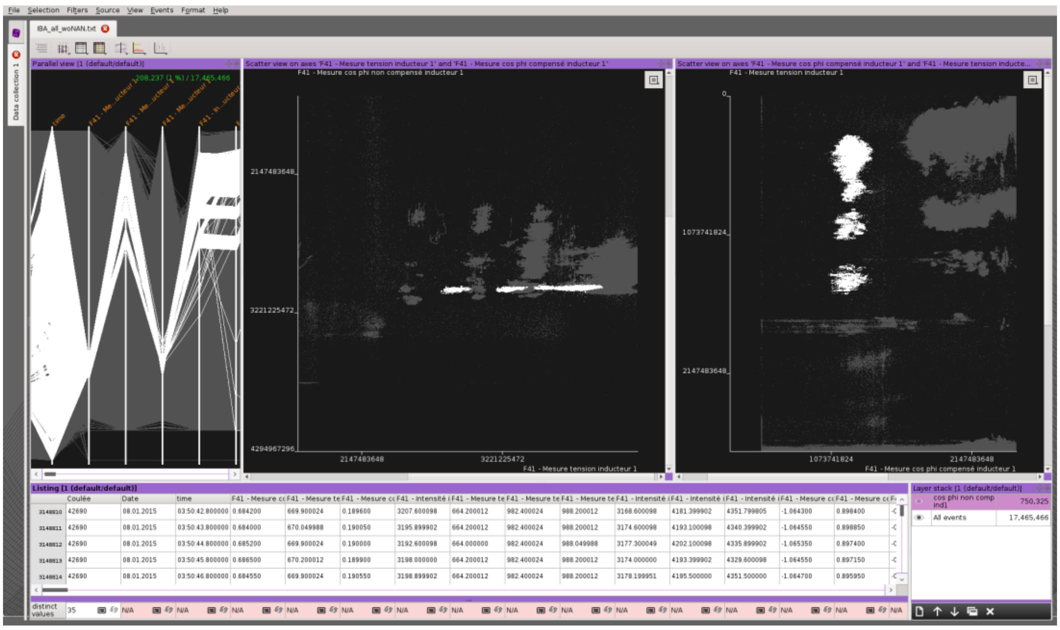

Figure 7 Screenshot of INENDI Inspector with 18 million records

INENDI Inspector has been conceived around a new investigation methodology that outperforms most of the standard data investigation approaches based on statistics or blind applications of Machine Learning. It can be used for different purposes:

- *Initial understanding of large and complex data sets*. INENDI Inspector makes it possible to face a very large amount of data and perform "by hand" a structured and exhaustive decomposition of it, so that one can get a clear and useful understanding of it. Very often, data coming from industrial sensors or large files of event-based records are so rich and complex that it is a tremendous challenge to draw a panoramic understanding of it. INENDI Inspector is especially valuable in the early phases of data discovery and data understanding. Data that were previously considered as too complex or too messy to be used can become valuable assets and can therefore integrate an Industry 4.0 strategy.
- *Discover and describe the detailed structure of datasets*. On Industrial datasets, it is usually very challenging to describe the hierarchy of details that go from the general facts to the smaller-scale details. But this description is very important to have when one wants to use such datasets to build valuable functions such as Fault Detection, Predictive Maintenance, Prognostics, etc. INENDI Inspector provides this very exact type of segmentation of a dataset that is required before using safely Machine Learning.
- *Isolate weak signals very efficiently*. One historical area where INENDI Inspector has been particularly useful is the cybersecurity area. In this domain, it is critical to be able to spot anomalies and weak signals hidden within huge amounts of events logs (logs from network equipment, servers, security appliances, company applications, etc.). Even if the traditional SIEM approach provides an interesting set of functionalities such as data aggregation, data queries and data monitoring, it is usually not suited to detect hidden unknown patterns that are not accessible through basic statistical deviations. In the industrial domain, weak signals detection is also of the utmost importance because they often provide early indications of major failures and catastrophic accidents. Thanks to the specific investigation methods supported by INENDI Inspector, weak signals and small-scale anomalies are revealed in a very natural way. This opens the weak signal detection activity to a larger audience of engineers and junior data scientists. For more expert data analysts, this represents a huge benefit in terms of time savings when in charge of deep data analysis. It really makes a big difference in one's ability to discover new unknown patterns in a data set.
- *Control and improvements of ML algorithms*. When a ML algorithm is being designed to handle a high-dimensional dataset, it is important to control its efficiency and spot parts of the training dataset that generate errors in the ML algorithm. When the training data set is large and complex, it becomes difficult to understand the behaviours of a

ML algorithm and connect these to the original data. In such a situation, INENDI Inspector is used as a ML testbed because it allows to load simultaneously the training dataset and most of the internal variables of ML algorithms. Understanding the mistakes, the false positives, the wrong predictions, etc. becomes much easier and more intuitive, due to the visual nature of the representations provided by INENDI Inspector.

3.2.2 Optimization of Flexible Machines Manufacturing

3.2.2.1 Overview

The FILL pilot related to the BOOST 4.0 project primarily serves the engineering process of the machine builder. It allows for a better understanding of machinery by detecting cause-and-effect relationships due to anomalies and patterns. In addition, maintenance intervals and cycles can be optimized and, as a result, quality improvements of the production and the product can be achieved. Currently there is a lack of digitalization mostly through manual information flow (see production system life cycle). It is planned to resolve this lack within the BOOST 4.0 activities.

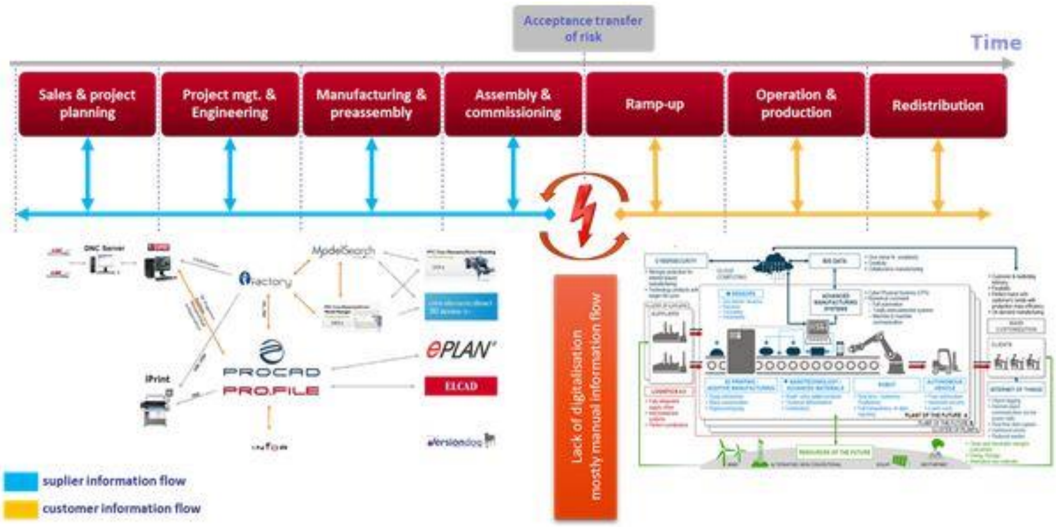

Figure 8 Production System Lifecycle

The technological possibilities for measuring, recording, storing and processing machine and production data are constantly improving. In individual areas, e.g. the CAD/CAM process between engineering and production, systems with feedback loops (closed-loop system) are created. Due to media discontinuities, manual evaluations and activities, however, the existing systems are mostly open-loop systems. This means, especially for globally (distributed) systems, that the response to faults or changes can only be inadequate, delayed and outdated. In order to achieve the vision of a digital, intelligent, agile and self-controlling value added network (engineering and production), the information loop must be closed across all systems involved.

Actually, the information flows between customer and supplier are very diverse. Most of the time these are informal information channels, e.g. the operator (customer) calls the engineer (supplier) directly. Getting the customer information into the engineering management system is therefore a person-dependent process so that the requirements management as well as the tech. change management are insufficiently supported. There are also support and service information used during the engineering. Data analyses to gain knowledge for new developments are currently not carried out. Therefore, the lead-time and quality of new developments depends primarily on the experience and communication skills of the employees in the project team.

RISC and FILL are already working on a scientifically close research project called VPA 4.0 – Virtual Production Assistant funded by the Government of Upper Austria. The goal of this project is the conceptual design and development of a virtual production assistant (VPA 4.0) based on data and visual analytics of machine and job data, providing a learning-based knowledge base for the machine. The VPA 4.0 is primarily used to assist local experts. It allows a better understanding of the machine by recognizing cause-effect correlations due to anomalies and patterns. In addition, maintenance intervals and cycles are optimized and as a result, quality improvements of the production and the product can be achieved.

Within the FILL pilot in the BOOST 4.0 project an integrated business process for smart digital engineering using big data extends the V-model. The V-model describes the development of mechatronic systems based on a systematic analysis of the requirements and a distribution of the requirements and loads among the individual disciplines. Furthermore, for the project BOOST 4.0, EIDS connector for the data harvesting process and the model development are planned.

3.2.2.2 Big Data Models, Techniques and Methodologies for the Optimization of Flexible Machines Manufacturing

All the planned analyses for the optimization of flexible machines manufacturing are based on big data coming from machines. On the one hand, sensor data in the meaning of big data streams and on the other hand information of the machine (model-driven). Promising results are expected by the combination of sensors data and model-driven data analytics.

As defined in this first months of the project, the most of FILL data that will become available, will be time series data, and the analytic tools will be designed in order be to find anomalies in such type of data.

Time Series Analysis

The beginning of the iterative process of time series analysis always consists of a visual and a theoretical analysis. The aim is to find characteristic properties of time series that indicate the beginning or the end of a historical trend. These characteristics can range from simple piecewise defined increases to the strength of the expression of previous historical trends. Combinations of different technical indicators using binary relations can also be used experimentally here to find trend signal generating properties. It is important to mention at this point that domain-specific expert knowledge for the food industry, its cycles and peculiarities must be incorporated in order to identify special properties that may only play a role in this area. The many years of industry experience of our team will be of decisive importance here.

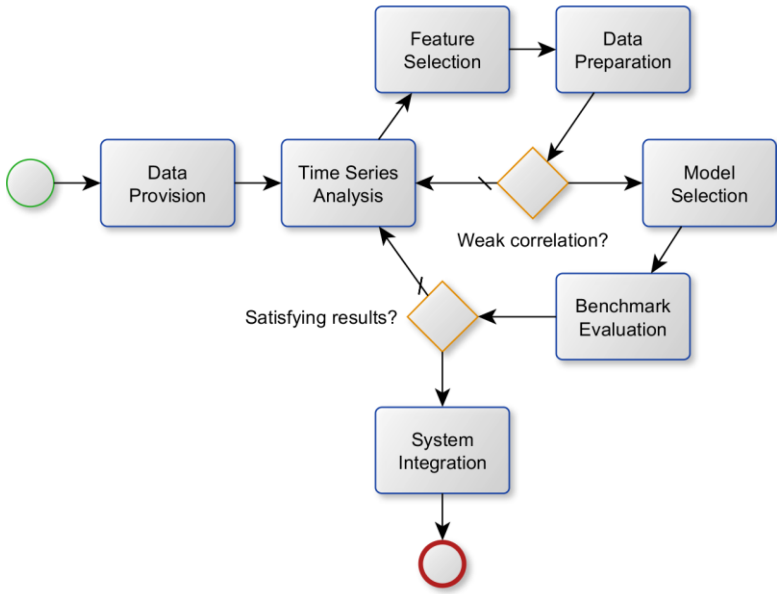

Figure 9 Development processes of time series analysis

Feature Selection

After the time series analysis, the trend related characteristics identified must be described mathematically as characteristics so that they can later be used by the machine learning algorithms. In the literature, for example, the following characteristics are known for trend recognition:

- Trend signals/technical indicators (in combination with binary relations)
- Chart formations (via image recognition processes)
- Historical trends (marked by gradient and duration)

The technical indicators in particular are mostly based on smoothing using moving averages, which always describe the development of the underlying time series with a time lag. On the one hand, this has the disadvantage that trends can only be identified with a parameter-dependent delay. On the other hand, these indicators correlate strongly as characteristics due to the same calculation basis (moving averages). However, the aim of the selection is to describe significant characteristics that describe as many independent characteristics of the underlying time series as possible. Therefore, the properties observed in the analysis and described in the feature selection must also be checked to see how strongly they correlate with each other in order to increase the quality of the later trend recognition and to reduce the run time of the algorithms.

In this context, it is important to note that the exclusive focus on signals for the beginning or end of trends probably ignores one important aspect: that of trendless phases. Particularly in connection with sales figures, there are always phases in which no clear trend is apparent. The limitation to the recognition of such a trend at any time of a time series is therefore unrealistic. Therefore, the recognition and description of features that are symptomatic of such trendless phases will also be part of the planned work.

Data Preparation

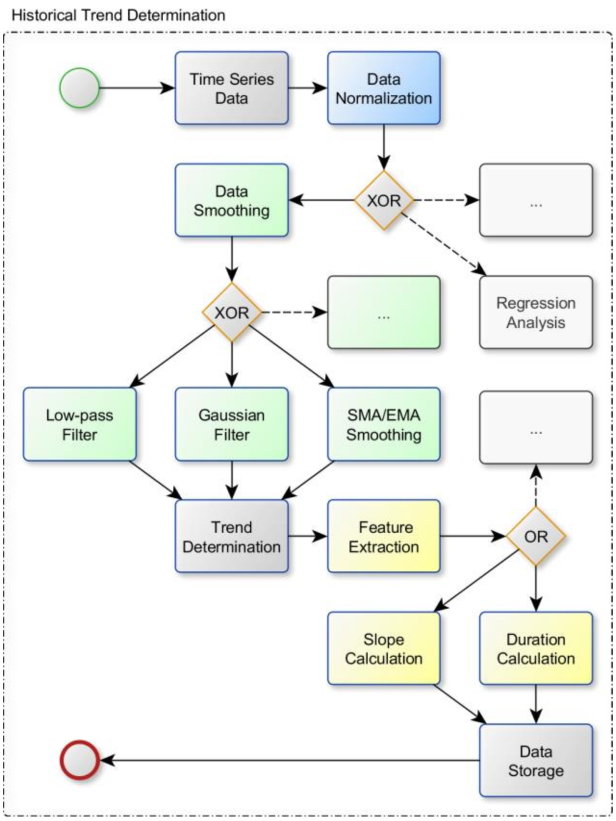

Figure 10 Workflow for the determination of historical trends

This point is of the greatest importance within the workflow, since the data of the underlying time series must be formatted to match the previously selected characteristic in order to represent the characteristic as accurately as possible in the model later. In the case of simple characteristics, the formatting of the data is usually obvious. Features that are more complicated do not only require an automatic parameter selection, but also the corresponding calculation rule is not clearly defined. The best way to illustrate the data preparation step in this case is to use the characteristic of the historical trends:

At the beginning, the given time series must be normalized (blue). This means primarily the removal of seasonality, or periodicity. It is also important to convert time series of absolute sales figures (Absolute Change) into time series that describe the relative change of the sales figures, since the use of absolute values leads to an overfitting of the model to the values of the training data.

To reduce the volatility and the associated noise of the time series, there are various possibilities. One of them is the smoothing of the time series (green), which in turn can be performed with different approaches.

Using the normalized and smoothed time series, the historical trends must be determined, for example from one local extremum to the next. The mathematical description of these trends (yellow) results in the actual features for the machine learning algorithms: Later, correlations between the duration and slope of successive trends can be found if they exist. Whether these two features are sufficient to describe trend significance is the subject of current research.

Finally, the newly calculated features are subjected to a correlation analysis with features already calculated in previous iterations. If it turns out that there is a strong correlation to already used features, the new ones are discarded and the next iteration of the process begins. Otherwise, the data prepared in this way - the calculated features as input and the classification of the historical trends as output are saved for further use (training) of the machine learning algorithms.

At this point it should be noted that historical trends are only one possible characteristic of one of the time series discussed here and that their data preparation cannot be transferred to other characteristics. The effort for the development of a process for the preparation of the time series for other characteristics, as well as its variety, should be comparable.

Model Selection

When selecting models in general and creating prototypes of suitable machine learning algorithms in particular, we will rely on the support of software tools such as Sagemaker (Amazon)[1], Machine Learning Engine (Google) [2]or Deep Cognition[3]. Since it is possible at this point in the development process that a specially prepared data set is available for each

[1] https://aws.amazon.com/sagemaker/
[2] https://cloud.google.com/ml-engine/
[3] https://deepcognition.ai/

feature extracted in the preceding steps and iterations, a suitable and adjusted machine learning algorithm should be searched for for each feature. These software tools provide specially optimized algorithms (Sagemaker), large machine learning frameworks such as TrensorFlow[4], scikit-learn [5]and Keras[6] (Machine Learning Engine), as well as a comprehensive construction kit for the fast creation and setting of neuronal new benefits (Deep Cognition). The in-house development or detailed improvement of individual machine learning algorithms is planned for a later point in time, if we should reach the limits of the already very mature available algorithms in this project. The focus of our work is therefore clearly on the steps in the development process of trend recognition that precede the model selection.

Benchmark Evaluation

The last step of an iteration in the development process of trend recognition is the evaluation of the results obtained. In this context, a benchmark that is adapted to the database and allows the testing of the models selected in each iteration together with the previously prepared data will be developed. These tests should allow a qualitative comparison of the individual iterations with each other, with the aim of further using suitable models and data and discarding unsuitable ones.

Based on a cross-validation procedure, we will test the three data sets selected during data provision with each new feature-model combination and then assess their influence on the quality of trend recognition. The subsequent extension of the benchmark by a scenario analysis represents a possibility to check the robustness of the trend detection for small fluctuations in the underlying data. It can also be used to evaluate to what extent the iteratively developed trend detection is suitable for trend prediction.

3.3 Big Data Models, Techniques and Methodologies for Smart Production Planning & Management Pilots

In order to support the creation of Smart Production Planning pilots, the BOOST 4.0 project will use and implement innovative Big Data cognitive manufacturing processes, predictive methodologies and algorithms that enable:

- The creation of a real-time self-learning virtual factory 4.0
- Digital-cognitive value chain for machine tool manufacturing

[4] https://www.tensorflow.org/
[5] http://scikit-learn.org/stable/
[6] https://keras.io/

3.3.1 Self-learning Factory 4.0 for Inventory Optimisation and Dynamic Production Planning

3.3.1.1 Overview

A self-learning factory 4.0 is planned for modern automotive production facilities. Volkswagen Autoeuropa plant was designed using advanced technology and continuously incorporates the latest developments in automation and computerized production control, in order to meet the high standards required for manufacturing a quality product.

However the process for inventory optimisation is heavily reliable on manual processes and in addition to that, the operation is performed inside the factory, where space is limited. On the receiving area trucks are traditionally unloaded by a manually forklift operation, and then the unit loads are transported to the warehouse where they will be stored either in shelve or block storage concept. System wise there is one database to control the parts coming from each truck and then a separate database, which registers the unloading, transportation and storing of the material in the warehouse.

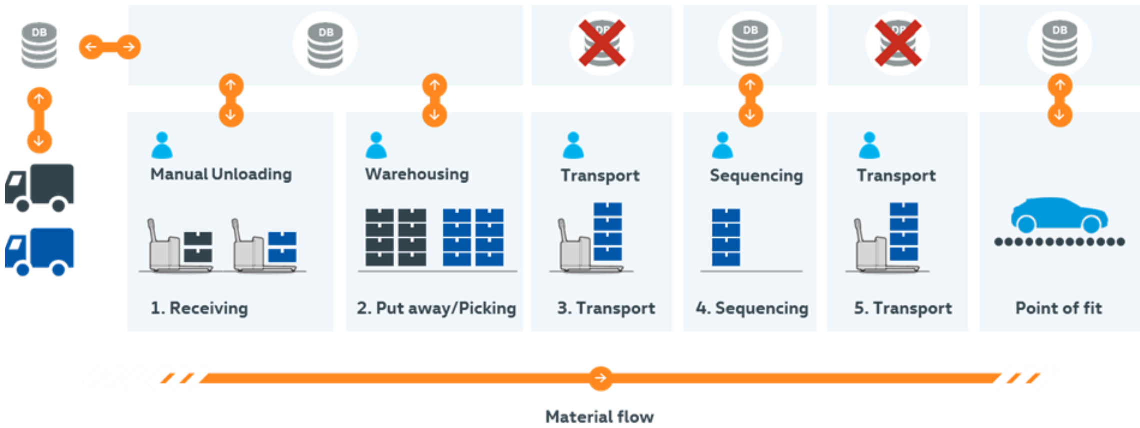

Figure 11 System flow of current process

An automatic line feeding system based on real vehicle demands generates parts call offs after interacting with real time stock data to replenish the points of use at commissioning areas called SUMA's, or directly at the assembly line using a pull methodology/concept. Deliveries are system supported until the point of fit. The next step will be the picking process for the correct sequencing in the SUMA. Here, the operator follows system electronic picking of parts according to the vehicle sequence on the production line. These operations are executed under the principles of the lean production system.

Conventionally the racks with the parts are moved from the picking/sequencing area to the point of fit through trolleys attached to tow tugs. This operation is exclusively performed by the operator and is achieved without resorting to any database, therefore any control or analysis of the process is not possible for the moment. All SUMA line-feeding processes consist in establishing a circuit of deliveries of full unit loads and the collection of the empty trolleys and boxes used in the replenishment process for the assembly line. Finally, the parts are manually delivery at the point of fit by the line-feeding operator.

The envisaged future scenario aims at achieving a full integration of the material flow, from receiving up to the point of fit. Figure 9 shows the system flow integration as it is foreseen in VWAE.

The main objective is to eliminate human intervention or at least reduce to a minimum at all phases from receiving up to the point of fit. The medium to long term vision is that activities will be fully, or at least to a large extent, automatized, namely:

- Automatic scan of each unit load (label)
- Automatic cargo checking and inspection
- AGVs with geolocation and communication capabilities will be used at all transport phases
- Picking Mobile Robots will be used for warehousing activities

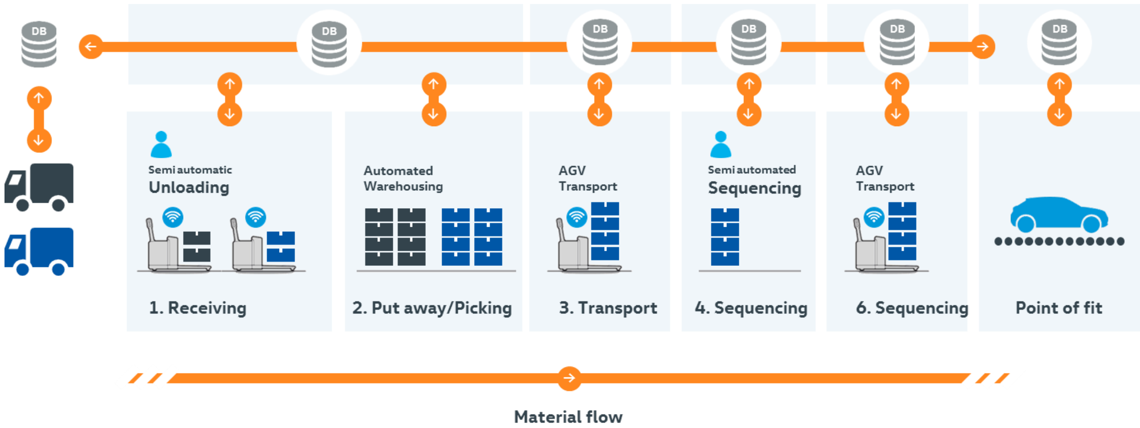

Figure 12 System flow of integration for Industry 4.0

The picking mobile robot will have the flexibility to adapt to different parts which will be linked and orchestrated by the existing sequencing system. The main objective is to lay down the pieces onto the rack in the correct order. The rack will be then picked up by a system composed by a set of AGV's which will handle the automated transport of these racks to the point of fit. Another aspect to consider is the monitoring functionality of the entire system available to the supervisor of the concept. This monitoring system will enable the control and management of all modular components as well as the regulation of key process indicators. Figure 10 shows the envisaged future scenario.

As the trial is quite ambitious from the point of view of focus processes and resources that would need to be available (robots, AGVs, etc.), we will break the trial into several step and will progress accordingly to the results obtained at each.

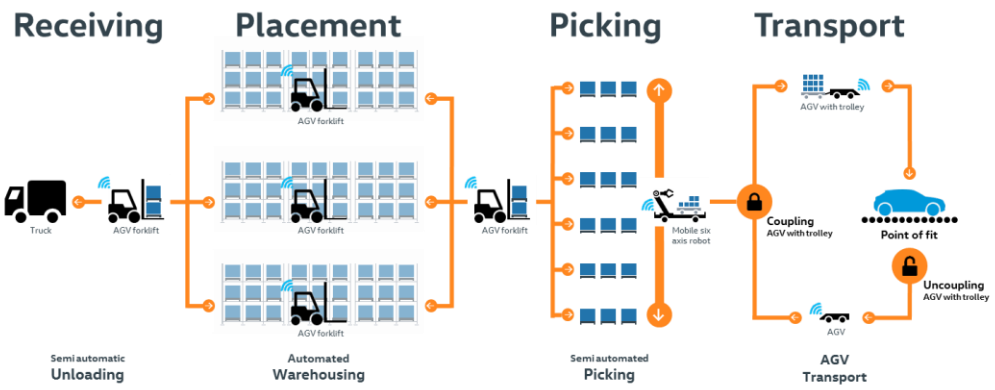

Figure 13 Envisaged future scenario in detail

Note that, for demonstration purposes, we will focus on a specific part, the batteries, trying to model their complete flow. The goal of the internal logistics supply chain is to reliably deliver the part to the point of fit; however, there are parts where this can be achieved relatively easily and others where it is quite complex. Some parts due to physical complexity are quite cumbersome to handle and this is reflected in its process, therefore, parts with this characteristic may not be suitable to choose because they raise undesirable obstacles on the innovation forefront towards industry 4.0 concepts. The batteries then pose as a solid option to tackle since its shape is quite uniform making them easy to pack, store, handle and transport. Also, batteries cover the full scope of Volkswagen Autoeuropa production since they are installed in every car. For those reasons, it is believed that these parts should be considered as vessel for the dissemination of innovative processes since they are quite permeable to experimentation.

3.3.1.2 Big Data Models, Techniques and Methodologies for the Inventory Optimisation and Dynamic Production Planning

Building upon existing Big Data technologies, and supported by BOOST 4.0 Big Data Platform, several models, methodologies and techniques are planned to be used in the VWAE scenario. The first step will be to model the Receiving phase in order to obtain a sort of digital twin that will enable its simulation. Once the model is built, it will be capable of being simulated in order to: (i) test accuracy, (ii) test different scenarios (e.g. using different solutions for label scanning) and (iii) select the most appropriate approaches (e.g. the one that enables a decrease on the time spent on scanning). The same approach will be applied sequentially to other process phases. To do so, a set of Big Data Analytics and Mining techniques will be used.

Since the scenario will build upon a virtual digital twin, designed using *VisualComponents*' (BOOST4.0 partner) simulation software, the techniques for Big Data will be based on batch processing techniques, and will use existing toolkits for offline, batch Big Data Analytics and Mining processing. Some examples are Apache Spark MLLib[7] or Jubatus[8], which run on known Big Data processing frameworks, such as Apache Hadoop[9] and Apache Spark[10]. Next table depicts some of the techniques and algorithms that may be used in the pilot. The final selected algorithms will be presented in details in the second version of this document when concrete results will be available.

Machine Learning Tasks	*Models and Methodologies*	*Alternative Implementations*
Basic Statistics	*Correlation*	
	Hypothesis Testing	Streaming Significance Testing
	Summarization	
	Stratified Sampling	
	Random data generation	
	Kernel density estimation	
	Mathematical Description	Gradient descent, Stochastic gradient descent (SGD), Limited-memory BFGS (L-BFGS),
Feature Extraction Selection and Transformation	*Feature Extractors*	TF-IDF, Word2Vec, CountVecotrizer, FeatureHasher
	Feature Transformers	Tokenizers, WordsRemover, n-grams, PCA, PolynomialExpansion, DiscreteCosineTransform, StringIndexer, HotEncoderEstimator, etc.
	Feature Selectors	VectorSlicer, RFormula, ChiSqSelector
	Locality Sensitive Hashing	Approximate Similarity Join, Approximate Nearest Neighbor Search, Bucketed Random

[7] https://spark.apache.org/mllib/
[8] http://jubat.us/en/index.html
[9] https://hadoop.apache.org/
[10] http://spark.apache.org/

Machine Learning Tasks	*Models and Methodologies*	*Alternative Implementations*
		Projection for Euclidean Distance, MinHash for Jaccard Distance
Classification	*Logistic Regression*	Binomial LR, Multinomial LR
	Decision tree classifier	Gradient-boosted tree classifier
	Random forest classifier	
	Multilayer perceptron classifier	
	Linear Support Vector Machine	
	One-vs-Rest classifier (a.k.a. One-vs-All)	
	Naive Bayes	
Regression/Prediction	*Linear regression*	Generalized linear regression, Available families
	Decision tree regression	Gradient-boosted tree regression
	Random forest regression	
	Survival regression	
	Isotonic regression	
	Tree Ensembles	
	Gradient-Boosted Trees (GBTs)	
	Linear models	SVMs, logistic regression, linear regression
Recommender Systems	*Collaborative filtering*	Explicit vs. implicit feedback, Scaling of the regularization parameter, Cold-start strategy
Clustering	*K-Means*	Bisecting k-means, Streaming k-means
	Latent Dirichlet allocation (LDA)	
	Gaussian Mixture Model (GMM)	
	Power iteration clustering (PIC)	
Frequent Patterns	*FP-Growth*	

Machine Learning Tasks	*Models and Methodologies*	*Alternative Implementations*
	Association Rule Learning	
	Prefix Span	
Models	*KMeans Model*	
	Regression Models	LinearRegressionModel, RidgeRegressionModel, LassoModel, SVMModel, Binary LogisticRegressionModel

Table 1 Big Data Analytics and Mining Methods for Inventory Optimisation and Dynamic Production Planning

In addition to these, Apache Mahout[11] enables developers to build alternative algorithms, through a Scala-based Domain Specific Language that allows new algorithms to be written as mathematical expressions.

A Big Data Analytics Infrastructure that has all of these tools, models and algorithms is available from UNNINOVA and is totally compatible with the BOOST4.0 proposed architecture. This software infrastructure has all the necessary components to work in standalone mode, but it also has the ability to use external Big Data tools (Data ingestion, storage, processing, etc.) to run the necessary processes.

3.3.2 Learning Processes and Machine Variables Monitoring for Cognitive Value Chain

3.3.2.1 Overview

A critical issue for manufacturing in Europe is the resulting machine cost with respect to main competition in Asia; differences are of the order of 30%, which can be partially compensated by the product quality and offer in terms of precision and productivity, but puts a high pressure on development, supply chain and operations. In particular, it is necessary to improve the assembly and test of different machine functional groups (mechanical, electrical, electronic) and processes, but also be able to provide to the market innovative features in shorter times than the usual 2-3 years cycle for the upgrade or development of a new device.

Currently, following a planning stage, supplies are scheduled and delivered to the spindle manufacturing plant, where critical elements are machined using GF milling machines. These critical parts are then controlled using measurement devices before going into the

[11] https://mahout.apache.org/

assembly stage. Another quality assurance stage takes place after the assembly before the new spindle integrates the manufacturing flow of a new machine.

This flow does not allow a real time update of the overall planning as the communication of data from the line is mostly paper-based. Additionally there are several manual operations with no real time follow up, and key information about the quality and speed of the process is lost. To summarize, there is no easy access of data from every stage for a reliable control of the production and to reduce the number of defected parts, which are only detected at the quality control points.

Improving the efficiency of the assembly line involves a leap in data management and communication from the engineering stages towards the production stages during the pre-series phase. A faster feedback should be implemented in order to adjust the product quality and enable the series deployment. At the same time, it is crucial to keep the high level of precision and quality, which are the backbone of the product differentiation. Traceability and monitoring tools are partially in place but they remain to be deployed across the lifecycle of the product. Ultimately, GF looks at the reduction of the total cost of ownership in order to provide competitive offers to the segments served, in particular aerospace, automotive, ICT and medical industries.

The ambition here is to develop an adaptive machining system in automated cells, integrated into an open cloud and data analytics infrastructure, taking in account actual information of machine condition and performance in accuracy, surface quality, productivity and sustainability. The eco-systems will be able to diagnose in real-time the state and performance of the machine, and correct deviations through updated simulation models and planning systems.

+GF+ machine tool optimum production factory 4.0 aims to create, digitize and standardize manufacturing-related data across its value chain and beyond, to customers. It will extract and process the relevant information through artificial intelligence, in a common data space, so to make it available to a smart production planning system, targeting optimized, zero-defect manufacturing at a new +GF+ factory in Switzerland.

3.3.2.2 Big Data Models, Methodologies and Tools for a Cognitive Value Chain

The main approach will be to rely on the exploitation of advanced enabling technologies for monitoring complex equipment. It requires the capacity to integrate heterogeneous data sources, and its innovation lies on the incorporation of various domains of knowledge. In this regard, ontologies facilitate a multidisciplinary approach through representation of domain knowledge with its concrete definition and provides semantic interoperability. Implementing an ontology, as a main reference model, will provide a common glossary to integrate various data sources, to facilitate the covering of the life cycles of the factory and its products. In addition, it will work as a meta-model to standardize the integration of further data sources.

Data analytics for predictive maintenance will also be provided in order to facilitate the detection of outlier pots, the prediction of defected outputs, the detection and prediction of pot incidents, the improvement of anomaly root cause diagnosis, and trigger JIT or predictive production activities. In addition, analytic techniques will be developed to decode the data intensive information of the machine production planning and management processes. Pattern recognition and analytical methods will be coupled with graph and multi-level visualizations for enabling plant-wide and local control operators in order to compare the performance of similar nodes, to early predict abnormalities as well as to formulate and validate various hypotheses regarding the performance of the production process.

For the both previous approaches presented, big data visualization tool will be provided by ESI. In a first stage, we will use this tool for doing an audit of the available data sets (sensors, quality control, metrology, customer requests…) in order to check the volume / validity / veracity / traceability of the data. The combination of all data sets generated during the lifecycle of the product must give the possibility to trace back to the root cause failure of the concerned product. The tool provided, INENDI Inspector, must help to travel over these data sets.

Once done, we will browse over the different data sets coming from the milling machine sensors (represented as time series) for identifying patterns. These possible behaviors will then be selected with the interactive tool and exported towards the data analytics functionalities for predictive maintenance proposed by EPFL that will be supported by the ESI Scilab software.

Below, two short descriptions of both aforementioned tools:

The INENDI Inspector solution is a software solution that allows any user to handle large amounts of structured or semi-structured data and perform a very rich and deep investigation of the data. Thanks to a set of intuitive and very interactive visualizations of this dataset, the user experience is freed from the classical burden of query-based interactions and the latter can use all his/her logical and deductive power to learn valuable insights from the data.

INENDI Inspector has been conceived around a new investigation methodology that outperforms most of the standard data investigation approaches based on statistics or blind applications of Machine Learning. It can be used for different purposes:

1. *Initial understanding of large and complex data sets.*

INENDI Inspector makes it possible to face a very large amount of data and perform "by hand" a structured and exhaustive decomposition of it, so that one can get a clear and useful understanding of it. Very often, data coming from industrial sensors or large files of event-based records are so rich and complex that it is a tremendous challenge to draw a panoramic understanding of it.
INENDI Inspector is especially valuable in the early phases of data discovery and data understanding. Data that were previously considered as too complex or too messy to be used can become valuable assets and can therefore integrate an Industry 4.0 strategy.

2. *Discover and describe the detailed structure of datasets.*

On Industrial datasets, it is usually very challenging to describe the hierarchy of details that go from the general facts to the smaller-scale details. But this description is very important to have when one wants to use such datasets to build valuable functions such as Fault Detection, Predictive Maintenance, Prognostics, etc.
INENDI Inspector provides this very exact type of segmentation of a dataset that is required before using safely Machine Learning.

3. *Isolate weak signals very efficiently.*

One historical area where INENDI Inspector has been particularly useful is the cybersecurity area. In this domain, it is critical to be able to spot anomalies and weak signals hidden within huge amounts of events logs (logs from network equipment, servers, security appliances, company applications, etc.). Even if the traditional SIEM approach provides an interesting set of functionalities such as data aggregation, data queries and data monitoring, it is usually not suited to detect hidden unknown patterns that are not accessible through basic statistical deviations.
In the industrial domain, weak signals detection is also of the utmost importance because they often provide early indications of major failures and catastrophic accidents.
Thanks to the specific investigation methods supported by INENDI Inspector, weak signals and small-scale anomalies are revealed in a very natural way. This opens the weak signal detection activity to a larger audience of engineers and junior data scientists. For more expert data analysts, this represents a huge benefit in terms of time savings when in charge of deep data analysis.
It really makes a big difference in one's ability to discover new unknown patterns in a data set.

4. *Control and improvements of ML algorithms*

When a ML algorithm is being designed to handle a high-dimensional dataset, it is important to control its efficiency and spot parts of the training dataset that generate errors in the ML algorithm. When the training data set is large and complex, it becomes difficult to understand the behaviours of a ML algorithm and connect these to the original data.

In such a situation, INENDI Inspector is used as a ML testbed because it allows to load simultaneously the training dataset and most of the internal variables of ML algorithms. Understanding the mistakes, the false positives, the wrong predictions, etc. becomes much easier and more intuitive, due to the visual nature of the representations provided by INENDI Inspector.

Beyond the already implemented scientific capabilities, the **SCILAB Team** is above all providing an engineering environment. This environment based on a flexible platform and a scientific language allows algorithms development & management, scientific applications deployment and overall system control.

1 Calculation engine & platform

SCILAB is an extensible exchange & work platform. It's runnable in command line, can be hidden behind front-end processes and communicate via ASCII, CSV, XML, XLS, or binaries data format. SCILAB also provides a toolbox management system for external contribution management and functionalities could be extended by calls to Java, Python, TCL, Fortran or C/C++ libraries.

2 Cloud – Industrial IoT

SCILAB Cloud is a quite new initiative that aims at providing engineers with a shared platform for data and applications centralization. This Cloud solution is all about data storage space and SCILAB calculation engine embedded within Virtual Machine running on the Cloud that allows the user to deploy his SCILAB applications across his company, all over the world. Applications and data can then be leveraged to make on-line plant monitoring and optimization.

3 XCOS – System modelling & control

One of the major functionalities of SCILAB distribution is the XCOS module. XCOS is a graphical editor to design hybrid dynamical systems models. Models can be designed, loaded, saved, compiled and simulated. Ergonomic and efficient solution for industrial and academics needs, XCOS provides functionalities for modelling of mechanical systems, hydraulic circuits, control systems, etc.

XCOS can finally be used for software-in-the-loop applications such as soft real-time control thanks to the SCILAB C code generator that allows to push an XCOS model on an embedded target.

3.4 Big Data Models, Techniques and Methodologies for Smart Operations & Digital Workplace Pilots

In order to support the creation of Smart Operations & Digital Workplace pilots, the BOOST 4.0 project will use and implement innovative Big Data cognitive manufacturing processes, predictive methodologies and algorithms that enable:

- The creation of autonomous assembly lines based on mobile robots
- Zero defect and zero breakdown injection moulding production process

3.4.1 Autonomous Assembly Lines

3.4.1.1Overview

The creation of an autonomous assembly line, which assembles components for a middle size commercial truck, with high complexity and high variability, is planned for the FCA plant. In this case, automated guided vehicles and autonomous robots replace the traditional linear process of the assembly line in a welding cell. However, there are no planning, monitoring and maintenance precesse for these robots and vehicles. Furthermore, currently there no specific approach to store and analyse data related to the missions of the vehicles, their availability, taking into account the lead time for delivery and the uncertainty related to the interaction with the presence of human operators

The planned cognitive factory 4.0 for FCA will be equipped with an Autonomous Assembly Planning Tool that will contains forecasting mechanisms and a Monitoring and Maintenance System based on the data coming from the mobile robot fleet, and will allow the proper planning of maintenance activities.

The Mindsphere platform from Siemens will be used. The platform is equipped with a wide variety of analytic services able to fulfil most of the autonomous factory's requirements. As the project is still in an early phase, has not been decided yet which analytic services from Mindsphere will be used or what possible extentsions below. In the following section, the available analytic services that may be used for this case are presented. More details about the platform become available at chapter 5.

3.4.1.2 Analytic services for an autonomous assembly line

Anomaly Detection Service

Idea

The Anomaly Detection Service aims to automatically detect unexpected behavior of processes and assets using time series data. For a given asset and for a specified period, the user is notified if the asset behaves abnormally in any way. Using this information the user is able to monitor their assets (e.g. by setting notification/warning thresholds).

The service enables the user to run the following applications:

- Process & Condition Monitoring
- Early warning functionality
- Detection of fault conditions without explicit definition

Access

For accessing this service you need to have the respective roles listed in Analytics roles and scopes.

Basics

The Anomaly Detection Service uses a density-based clustering approach to detect anomalous behaviour. The clustering is done in a training with historical data and leads to a model for normal behaviour of time series data. This model is evaluated for a given set of data points by checking whether the data points belong to one of the clusters formed during model training. Data points belonging to a cluster are considered normal. Data points not belonging to a cluster are given a score that indicates their distance to the nearest cluster. The higher this score, the more likely it is that the data point is an anomaly.

In addition to one or multiple time series (original sensor data in IoT model format), the service uses a specific configuration as input for the signal calculation. The configuration information consists of the following parts:

Model Training (Clustering)

The Model Training clusters the historical training data (time series) and provides a model of normal behaviour of the process / asset. The clustering is done using the density-based algorithm DBSCAN (Ester, Martin, et al. "A density-based algorithm for discovering clusters in large spatial databases with noise.").

The algorithm uses the following parameters:

- Epsilon: threshold for the distance to determine if a point belongs to the cluster (optional, default: estimated)
- minPts: minimum cluster size (points per cluster, optional, default: estimated)
- Distance measure: type of the distance measure algorithm:
    - Euclidean (default)
    - Manhattan
    - Chebyshev

If a parameter is missing, the defaults are used for executing DBSCAN. The clustering assigns each data instance a cluster ID to which it belongs, or a noise label if it does not belong to any cluster. Models have a limited lifetime. They are automatically deleted (lifetime is 1 day at least).

Model Evaluation (Scoring)

The Model Evaluation determines whether a given set of data points is anomalous or not. This is done by calculating the distance of each data point to its nearest neighbor that is part of a cluster. If this distance is smaller or equal to epsilon, the data point in question is considered normal. Such data points are assigned a score of "0". In all other cases the data point is assigned a score equal to the difference between the calculated distance and epsilon. The higher the score, the more likely it is that the data point is an anomaly.

Features

The Anomaly Detection Service exposes its API for realizing the following tasks:

- Train a model
- Get all existing models
- Get a model by ID
- Delete a model
- Detect anomalies of time series data

Event Analytics Service

Idea

The Event Analytics Service provides the essential functionality for a data-driven analysis of event data. It enables the user to get a better grasp of what is happening inside the system through statistical analysis.

The service enables the user to run the following applications:

- Alarm management

- System troubleshooting
- Root cause analysis: identification of the root causes of faults or problems

Access

For accessing this service you need to have the respective roles listed in Analytics Services roles and scopes.

Basics

Almost any programmable device emits event data, containing a reasonable message text with a corresponding time stamp of the occurring events. Event data can be leveraged in order to create a new form of economic value. By analyzing it, we can extract additional knowledge and uncover useful patterns and hidden relationships between the emitted events. In return, it can not only help us to better understand the system's behaviour and dynamics, but also drastically increase the effectiveness of system management and support reasoning. The service processes event data sets as tuples of the data types time and string. An analytic session should have a maximum data volume of 1 MB and at least 1000 events.

Statistical Analysis

The statistical analysis enables the user to identify statistically significant dependencies by displaying the most frequently occurring events (Top Events). It is a reasonable assumption that those events are of the highest interest to the user. This information can be used to determine which data is useful for further analysis (e.g., by passing the data to future Event Analytics endpoints). It also helps to reduce the cost of further analysis, as well as increase the algorithm performance (both in terms of accuracy and time required).

The analysis takes the entire data set as input, sorts it according to the number of occurrences of identical events and returns a sorted list of individual events with their occurrences.

Features

The Event Analytics Service exposes its API for realizing the following tasks:

- Find most frequently occurring events

KPI Calculation Service

Idea

The KPI Calculation Service computes Key Performance Indicators (KPIs) for an asset. It uses data source such as sensors, control events and calendar entries. The service enables a user to get a unified view on an asset and to outline the following aspects:

- Industrial procurement: reliability, availability and maintainability
- Conditions monitoring: operating characteristics
- Alarm management and inferred events
- Risk assessments
- Diagnostics applications and Root-cause analysis
- Condition-based maintenance

Access

For accessing this service you need to have the respective roles listed in Analytics Services roles and scopes.

Basics

KPI calculation

The KPI Calculation Service offers a set of computational procedures based on ISO 3977-9:1999. It works for both offline batch analysis and the online analysis of new data. The calculation uses the following inputs:

- Time period
- Sensor data as time series {TimeStamp, SensorValue} and a threshold
- Events/states as time series {TimeStamp, EventID/State}
- Maintenance calendar data {StartTimeStamp, EndTimeStamp}
- Initial and default unit states

Depending on the availability of the input data sources a specific computational procedure is applied. The calculated output varies based on the input, but always consists of a set of KPI time and factor values.

KPI states and time values

A unit spends time in various KPI states. For a considered time period (Period Hours) the following hierarchy shows the resulting time values of a KPI calculation:

- Available Hours
    - Service Hours, unit is in-service.

- Reserve Shutdown Hours, unit is available, but non-generating and not in service (default).
- Unavailable Hours
    - Forced Outage Hours, unit is unavailable due to a failure.
    - Planned Outage Hours, unit is unavailable and in a planned outage by calendar.
- Unknown Hours (NoData), not all required data is available.

KPI factors

The KPI calculation also produces values for the following KPI factors:

- availabilityFactor, probability that the unit will be usable (available without any outages).
- unavailabilityFactor, probability that the unit will be unusable (not available due to outages).
- reliabilityFactor, probability that the unit will not be in a forced outage condition.
- serviceFactor, probability that the unit will be in an operating condition.
- forcedOutageFactor, probability that the unit will be in a forced outage condition.

Features

The KPI calculation Service exposes its API for realizing the following tasks:

- Calculate a set of KPIs within a time period
- Get the all KPI states within a time period

Signal Calculation Service

Idea

The Signal Calculation Service processes time series data of an entity's sensor. The service aggregates, modifies, smoothes and transforms the original sensor data for further analysis or storage along with the original data. The service enables a user to carry out the following tasks:

- Detect missing sensor values
- Replace missing sensor values with interpolated ones
- Compute a new physical parameter from available sensor readings
- Aggregate sensor values over a sliding window

Access

For accessing this service you need to have the respective roles listed in Analytics Services roles and scopes.

Basics

Configuration Options

In addition to one or multiple time series (original sensor data in IoT model format) as input, the service uses a specific configuration for the signal calculation. The configuration information consists of the following parts:

- Operation, a common operation for processing the time series
- Parameter, parameters of the operation
- Operands, a set of operands, each defining the entity, the property set and the property of the time series data to operate on
- Result, defines the output format as an IoT model time series: Entity, property set and property of the resulting time series

The service executes the operation on the original sensor data and responds with the calculated time series (so-called soft sensors).

Operations

The Signal Calculation Service executes a very wide number of operations. All the supported operations are listed at the APPENDIX E.he following operations:

Features

The Signal Calculation Service exposes its API for realizing the following tasks:

- Apply a mathematical operation to specified properties of time series
- Filter time series by specified boolean properties
- Merge multiple time series

Signal Validation Service

Idea

The Signal Validation Service validates sensor time series data of an entity. The API offers a set of common operations for signal validation. The service enables a user to analyze data based on custom settings, like thresholds or data window sizes. The service response can be used for special post-processing or alerting tasks.

Access

For accessing this service you need to have the respective roles listed in Analytics Services roles and scopes.

Basics

Time series data are used as input for signal validation. Depending on the property set, a time series consists of a time stamp and one or more related properties and values. If a time series has multiple properties, you have to define the property to be analyzed by the Signal Validation Service.

The Signal Validation Service can perform different checks on the data. Each check requires its own set of parameters. For example, the range check has parameters for an upper and lower limit for detecting range violations.

The validation output varies with the input, but is at least an event consisting of a description and a time stamp.

Features

The Signal Validation Service exposes its API for realizing the following tasks:

- Detect range violations
- Detect spikes
- Detect noise
- Detect jumps
- Detect/interpolate gaps
- Detect bias

Trend Prediction Service

Idea

The Trend Prediction Service predicts future values for time series using linear and nonlinear regression models. It is a forecasting framework, that has many useful applications in the area of Process & Condition Monitoring.

The service enables the user perform the following tasks:

- Predictive maintenance: A component's lifetime may be reached in the short-term future.
- Monitoring of processes: Prediction of remaining time to prevent unwanted process states.

- Seasonality and trend removal as preparation for other data analytics tasks.

Access

For accessing this service you need to have the respective roles listed in Analytics Services roles and scopes.

Basics

The Prediction Service is a data-driven approach that can be applied to univariate (single input variable) or multivariate (multiple input variables) time-series data. The predicted output is always univariate and written into a single target variable.

The service provides the functionality required for estimating the relationships between the variables of a given time series in order to make predictions based on the trained model. These predictions can then be used for reasoning about the process represented by the time series.

The trained models are based on a linear or polynomial regression.

Linear regression

Linear regression assumes a linear dependence between a single target variable (y) and one or more independent input variables (e.g., time) (x1, x2, …):

y = b0 + b1x1 + b2x2 + … + bnxn

For a single independent variable x, the model is given by y = b0 + b1x1.

While not very complex, the approach has proven to be a powerful method for early detection of the potential faults. The results produced by the algorithm are both easy to interpret and to visualize.

In the case of the multivariate input, the algorithm expects that all independent variables are available at any time.

Polynomial regression

Polynomial regression assumes a polynomial (of a given degree) dependence between a single target variable (y) and one or more independent input variables (e.g., time).

For a single independent variable x, the relationship is given by y = b0 + b1x + b2x2 + … + bnxn.

For a 3 independent variables x and a polynomial degree of 2, we have a model like y = b0 + b1x12 + b2x22 + b3x32 + b12x1x2 + b13x1x3 + b23x2x3.

This approach can capture more complex relationships between the variables than a linear model at the expense of an increased computational complexity and a need of specifying the degree of a polynomial to be fitted. In the case of the multivariate input, the algorithm expects that all independent variables are available at any time.

Features

The Trend Prediction Service exposes its API for realizing the following tasks:

- Train (fit) a regression model
- Predict future values
- Train & predict in one request
- Get all regression models of an entity
- Get a regression model by ID
- Delete a regression model

3.4.2 Autonomous Injection Moulding Production Process

Philips Drachten encompasses a large suite of highly automated processes used during the manufacturing of electric shavers. Of these manufacturing processes, injection moulding is of particular importance, as this is used during the fabrication of plastic components for electric shavers. All of these plastic parts are manufactured onsite at Drachten, requiring approximately 80-90 moulding machines. Injection moulding, however, is a competitive market, making it essential for Philips Drachten to continuously improve on quality, production performance, and costs where this process is concerned.

Next to the aforementioned metrics, the application of Big Data along with seamless connectivity in the manufacturing process will result in efficient ramp-up times between different moulds, along with full traceability along the process chain all the way to the customer.

Since the entire injection moulding park is connected to a dedicated machine network, process data can be retrieved from each of these machines. One of the main challenges is to be able to generalize the data and structure across multiple models and brands of equipment. Different models, generations, brands and types of machines all have different methods and structures for transmitting data. To give some examples: some older machines only connect via serial connection, some machines can only transmit data via file transfer and each manufacturer has his own internal data structure.

From a process engineering standpoint, injection moulding processes are well-documented and the basic moulding process should be comparable. We therefore see opportunities to generalize the data structures and to build generalized data analysis models describing the many features of the process and indicating (unwanted) trends, outliers and other sort of deviations. These models should be able to deployed across the majority of the machine park and could even be usable by other partners in the same Industrial Data Spaces. Furthermore, all the data in generalized structure forms Big Data and it can be used by data scientists in order to create algorithms for making the process smarter and more efficient. One of such example algorithms are predictive models. With different known methods and models there will be technologies created which will predict (1) quality of the molding machines' outputs and (2) unexpected failures of molding machines.

By combining a generalized data framework and combining it with generalized data models, we expect to:

1. Be able to better monitor product quality (predictive quality)
2. Be able to predict upcoming machine / part failures (predictive maintenance)
3. The ability to control machine parameters and trigger actions on the physical devices (control loop)
4. Be able to have a generalized dash boarding system providing the different stakeholders with the required information. The stakeholders include tool and machine maintenance departments, production and quality management, production operators, and the manufacturing IT department.

To be able to have actionable insights that take into account all present information at the pilot site, it is necessary to generalize the data coming from the different brands and version of the moulding machines. As a first step, even before the data clean up, we will create a common generalized model. This will most definitely happen in conjunction with the data science team, responsible for the algorithms and big data models, so there is at least sufficient data available. Instead of traditional techniques that are focussed on trying to find a common ground between all different data structures, here we want to go for a canonical model. A canonical model makes sure that no data is lost when transforming from the proprietary data structure to the generalized data structure. This is important to take the future insights even further. Wherever possible we will of course interlock and build further on the Industrial Data Spaces data models.

With Big data models the goal to improve the current production performance:

- Predictive Quality – 10% improvement on Fall Off Rate (FOR)
- Predictive Maintenance – 5% less down time (OEE-A)
- Intelligent process control – Contributes to predictive quality (in process-control)
- Give operators better tools – Mean Time To Repair improvement of 5%

In order to achieve the listed goals above and to build models that perform best with available Big Data set, different data models will be assessed, tested and trained. Many models are known already in the field of predictive models, such as:

- tree models
- probabilistic models
- deep learning models
- other possible solutions

As this deliverable comes in an early state of the project, it is still under study, which models perform, as reliable and accurate as possible and will offer the best possible results for this case. More details about these processes will be presented at the second and final version of this document (M24) and the deliverables related to this pilot.

3.5 Big Data Models, Techniques and Methodologies for Smart Connected Production Pilots

In order to support the creation of Smart Connected Production pilots, the BOOST 4.0 project will use and implement innovative Big Data cognitive manufacturing processes, predictive methodologies and algorithms that enable:

- The reduction of energy consumption and the efficient system operation for automotive part prescriptive quality assurance factory 4.0
- Smart assembly planning and supply chain visualization

3.5.1 Processes for Machine Operation, Quality Control, Maintenance and Energy Consumption Reduction

3.5.1.1Overview

Gestamp as an international group dedicated to the design, development and manufacturing of metal automotive components has identified Industry 4.0 as one of its main global objectives. Gestamp wants to work on two main pillars, logistics and quality in order to test different solutions to optimize the logistics of the plant through the localization of assets and the inspection and quality control. The pilot will take place at the Automotive Intelligence Centre for the first experimentation and at the plant that the group has in Navarra (Spain) for the large scale and on-site trials. The pilot will be focused on the improvement of the plant efficiency, based on the analysis of the different sources of data such as process, logistics and quality information.

General goal of the pilot is to improve of the performance and efficiency of essential manufacturing, quality and logistics processes. By means of analysis data from different quality sources from:

- Process data sources: Data from production processes (welding process, hot and cold stamping process) acquired by different solutions, such as PiWeb tool (real-time process control tool).
- Product data sources: Measurements from manufactured parts on different quality controls (attributes controls, product audits, manual measurements and CMM devices) and on a big range of formats.

Furthermore, in manufacturing plants, logistic is a very important part of the process and very difficult to manage due to its complexity. The aim of the project is to improve current management of the logistics in a Gestamp plant by localizing in real-time the main moving assets such as forklifts, cranes and containers. This real-time information will allow to identify current inefficiencies and areas of improvement on the logistics and production processes. The expected improvements are, among others, as follows:

- OEE improvement by reducing downtimes in the production lines
- Anticipate production lines' needs
- Improvement on stock management
- Optimize movements of logistic means

It has been proven that industrial scenarios are very challenging when talking about indoor localization. For that reason, the trial of the project will be done in Gestamp Navarra plant, where many different production process are present, and with a complex logistic that includes material flow between its two locations.

Regarding the inspection and quality control, Gestamp's plants processes includes some quality check points after the key manufacturing lines (stamping and welding). Some quality checks are carried out manually by blue collar workers who write the related quality information and decisions in notebooks/papers. Doubtlessly, the low accuracy of the checks especially after some of the technologies used in the plant such as welding whose conditions are more complex as well as the loss of information in relation to the quality of the manufactured products are major concerns in the factory. Thus, main problem focuses on the lack of quality information to take relevant decisions for the maintenance and system/process management.

In addition, one critical element is data access across factory information silos like CAPTOR MES or the ERP system and the interconnection and exchange of data among them and the quality information to gain knowledge about the global status of the factory instead of the actual local and individualized view. Currently, in Gestamp's plant CAPTOR is gathering data about the manufacturing systems so it is possible to have information about the production process. On the other hand, SAP is being used independently for resource management. The big challenge will be the deployment of a trustfulness interoperable data network within the plant to allow automatic assistance and the visualization of plant's status to take decision at machine, process and factory level.

In addition, corrective and preventive actions, the approach followed in Gestamp's plant, have an individualized, partial and rigid approach which impacts in the low efficiency and sustainability from the economic and social point of view. Thus, the pilot will be focused on developing a predictive approach in order to achieve zero defects and zero breakdowns.

In conclusion, with the combination of the existing digital sources and the new ones to be deployed, Gestamp expects to obtain:

- Production process optimization: With definition of an efficient data architecture, data correlation method and its understanding in real time. Current factory situation will be monitored. Workshop performance and further surplus will be identified to optimise production process.
- Reduce inefficiencies and further time losses due to no location of assets: With the implementation of advanced technologies to locate assets in a real-time basis, the logistics processes will be optimized, gaining valuable knowledge about these assets.
- Analytical tools: Define key indicators and levers on the project to transform acquired data into added value to increase process efficiency, becoming more flexible and precise.
- Support in the decision-making process: In addition to the previous points the real-time assets location, and data sources and visualization, they are expected to help operators to take decisions. Moreover, they will display the right data on the right moment as well as to include particular alarms, events into the user interface for critical issues to focus on avoidance of product and process identification mistakes (at the production and logistics areas), reduce uncertainty for the management and optimize the stocks.

3.5.1.2 Predictive Quality Control and Real-time Asset Monitoring

Considering the above-described objectives and the current quality control processes carried out in their plants, the main challenges with regard to Big Data are:

- the speed to acquire, transmit and process data of complex parts/pieces is relatively slow,

- the volume of data collected from massive point clouds for complex products such as chassis (up to 10 million points) is rendered hard to be processed and visualised,
- the status quo that quality data are collected and stored in different formats and in separate local servers, and some are not digitalised, makes it difficult to form an interoperable big data environment required in this pilot;
- the difficulty to correlate process and product data generated from distinctive sources at machine, process, and plant levels complicates the knowledge extraction and predictive analysis; and
- the visualisation tools for complex data from multiple sources are not currently used or developed which prevents the results of data analysis from being used in a meaningful way by factory workers.

Predictive Quality Control Framework

Faced with these big data challenges, it is agreed that the technology should be able to boost performance in speed, volume, interoperability and data process power. Having this vision, a **predictive quality control framework** is to be created with the following objectives in mind:

- To optimise the speed of data acquisition, visualization and processing for massive metrological data that boosts the overall efficiency.
- To design and create a collaborative predictive analytics platform to make decision making data savvy.

Quality Data techniques and methodologies

The proposed framework is going to have the following functions and proposed components to realise the objectives aforementioned. TRIMEK will develop the required connectors based on FIWARE open source technologies to exchange and share quality data. Moreover, it will extend and adapt its M3 solution for its implementation for data management, data integration, data visualization and data analysis:

- **Rapid transmission and processing solution of metrological data**: The **M3 Software** will permit to transmit and process data from different sensors almost in real time so as to minimise latency and improve the decision-making process.
- **High-performance massive point clouds processing and visualisation**. As the factories in this pilot are manufacturing products like chassis, whose level of dimension and complexity needs powerful processing software to render massive 3D point clouds and realistic colour maps with texture. The M3 software meets the need of such processing power with new algorithms designed to realise texture mapping for complex parts and the implementation of the QIF standard to cover the

complete metrological process from the Product Manufacturing Information (PMI) to the analysis and reporting.

- **Cloud solution** for quality data management and storage. Current scattered metrological data will be centralised and stored in the cloud, ready to be connected and processed with production data, being accessible through user-friendly interface. The M3 Workspace will be adapted to harmonise different data formats to generate the comprehensive data processing, management and visualisation. Data will be cleaned and enriched. The M3 workspace enables the access to the quality across different systems. Moreover, processed data will also be stored in the cloud of the M3 Workspace. Finally, a connection/bridge between QIF and IDS will be developed to ensure the required interoperability.
- **Advanced data analysis** is going to be applied to the quality and production data to realise zero-defect and zero-break down production. The data gathered through production line will be transmitted to M3 Analytics module with IDS connector, the component enables the correlation of process and product data. This analysis is capable of knowledge extraction to form insightful reports to enhance process efficiency, develop trends and execute features and/or parameters comparison between process, products and similar parts across different plants, etc. M3 Analytics will be implemented to improve not only the plants production efficiency, but also to facilitate a flexible and dynamic decision making process.

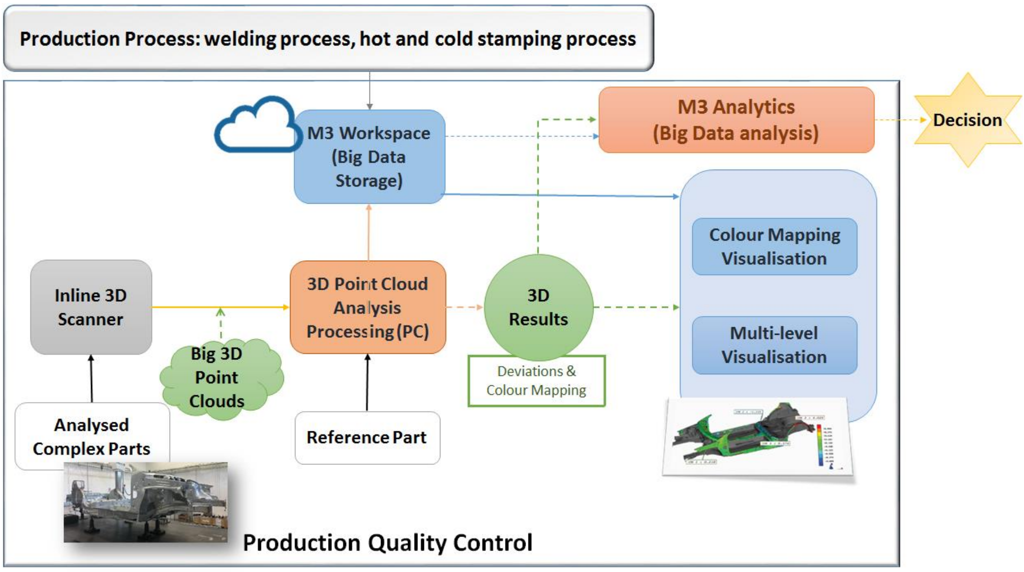

Figure 14 Predictive quality control scenario

The predictive quality control framework is going to be developed based on the M3 platform, described in chapter 5, which is poised to provide a structured solution for Metrology4.0, an edge-powered quality control analytics, monitoring and simulation system.

Real Time Asset and Network Monitoring

The real-time assets and network monitoring by ENEO and i2CAT will be designed for logistics efficiency, as it will be combined with a predictive analytic platform for logistics.

The Big Data challenges related to the overall logistics and factory management issues require their own solution. The overall objective of this approach will be to harness the capabilities of a Big Data analysis platform to improve the way in which data transmitted over Gestamp's network is collected, analysed, visualised and protected. There is a need to both enable more complex analytics of the available data as well as to provide actionable intelligence to decision makers in real time. Gestamp factory IT professionals should have a complete understanding of shop floor and network activity at all times and be able to quickly analyse available data to detect potential problems.

The approach to provide all of these solutions will be based on the adaptation of the existing redborder real-time network traffic analysis and security platform and i2track for logistic data collection and real time assets monitoring and analysis. These platforms provides the framework and modules necessary to enable the collection, enrichment, analysis and visualisation of large amounts of data in a variety of formats. The approach to achieve a big-data based smart logistics and real-time asset management strategies is composed of the following elements:

- **Data Collection**: The core objectives require a new, more streamlined approach to data collection which can be achieved with the used of the redborder platform. There is a need to evolve beyond the current scenario which involves separate data "silos" which exist between departments and manufacturing phases. The redborder platform will allow data from these sources to be integrated using a single, common approach to data collection. This unified approach to data collection will also enable the deployment of new sensors to provide currently unavailable data, such as the real-time location and activity level of manufacturing assets.
- **Data enrichment**: One of the core elements of the unified data collection and analysis strategy is the implementation of a data enrichment protocol. Once the data has been collected it will need to be enriched in order to facilitate its analysis as well as to increase the number of variables which can be measured in any given case. The enrichment process, simply put, should involve the automatic addition of context information to the data collected from each source. The benefits of the enrichment of the collected data is that it will enable data taken from any point in the manufacturing process and from any type of source to be used in the overall analysis of factory efficiency and productivity.
- **Data analysis:** The overall goal of the integration of data collection and enrichment will be to provide a more complete analysis of the available datasets.

This should enable the establishment of specific targets or goals for the data analysis process. In the initial stages this will involve using the newly collected and processed data to determine the location of factory assets in real time, protect the factory network and have a visual representation of factory and network activity at all times. This basic framework, implemented as a proof of concept during the trial period, will provide the foundation for the development of strategies and data analysis tools to solve any bottlenecks and weakness which have been previously identified as well as detect new potential problems in real time. This adoption of big data analytical techniques and specific focus on improving process efficiency, reliability and velocity will provide the main benefits of the new approach. Furthermore, the log functionality of the redborder platform will provide the means with which predictive analysis based on historical trends can be implemented.

- **Actionable, real-time intelligence**: The benefits of the big data approach to integrated data analysis throughout the manufacturing process and factory logistics management task need to reach all levels of decision makers. Furthermore, each user with access to the system should be presented with only the analysis of the data which most concerns them in an easy-to-understand way. This should all be done in real time or near real time in order to maximize the value of the tool .Therefore, the most visible result new approach will be customisable dashboards developed using the redborder platform. These customisable visualisation tools will provide factory and IT technicians with the precise information they need to make decisions to improve factory results. Examples of these dashboards are presented in the following images:

Figure 15 Factory Asset Location Dashboard

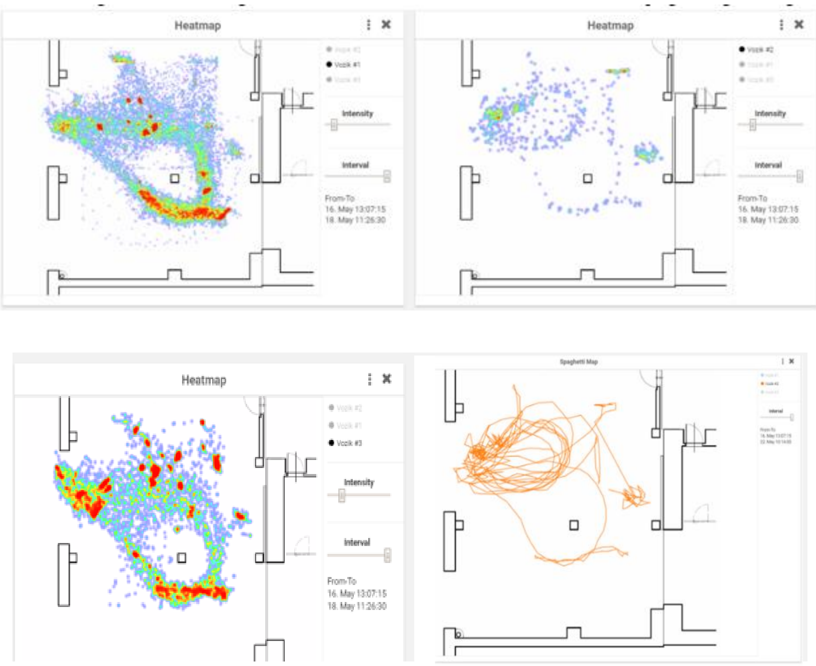

Figure 16 Examples of heat maps and route diagram of mobile assets Dashboard

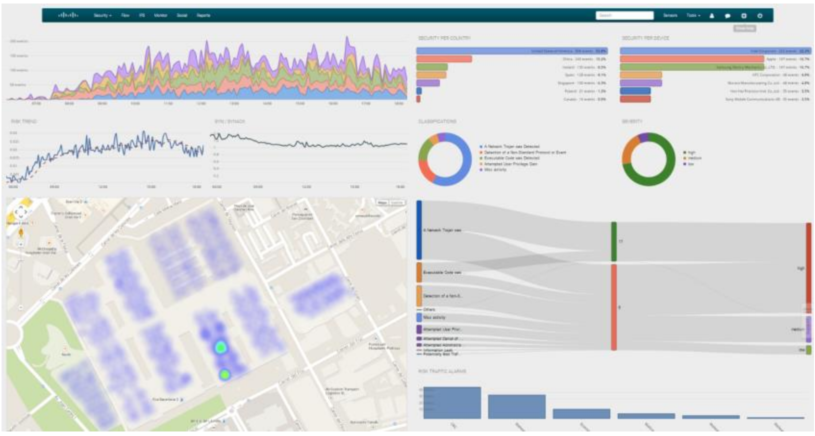

Figure 17 Factory Network Status and Security

These dashboards are the visible results of the new approach to big-data based analysis of manufacturing and logistics data. The connected factory assets are visible and information on their activity level, location and interaction with other assets is provided in a visual manner (colour coded heat maps, route diagram, trajectory lines, status symbols, etc.) which makes them easily interpretable. This is the same for the factory network information that allows IT professionals to quickly visualise the status and overall activity of the network and connected assets. These examples of this particular Big Data approach use case demonstrate how the connected factory can increase the overall efficiency of the manufacturing process and streamline logistical operations.

3.5.2 Smart Assembly Planning by Forecasting Over the Supply Chain

3.5.2.1 Overview

The production supply chain covers all activities related to planning and delivery of a new product to a customer. Material supply includes the activities such as plan material, transport material and handle material (goods receiving). It is important to have a lean and flexible material flow from supplier to plant. This helps to avoid long lead-times, tied-up capital and high transport costs.

With BOOST 4.0 solution for the VOLVO's supply chain, it is aimed at to make the entire supply chain transparent and totally trackable. In order to track and trace materials and fixtures in transit. Furthermore, it is expected that analytic functions can be developed to provide inputs for decision making e.g. production planning and material planning to increase efficiency and reduce cost. Moreover, blockchain technology will enable the visibility of the supply chain on the one hand, and will record an immutable ledger for the sake of analytics, on the other.

The future scenario is the implementation of a track and trace mechanism for the transportation of cabs in the VOLVO supply chain and the development of forecasting mechanisms that will enable the prediction of cabs' arrival time or they will detect possible delays in the aforementioned procedure of the supply chain. The main objective is to create a live time monitoring framework and a forecasting engine that will enable a more dynamic and effective production planning for the trucks assembly procedures in Tuve factory based on immutable transactions recorded on blockchain.

There are three basic steps that enable the above described scenario:

- Installation of GPS trackers/sensor and beacons edge devices
- Use of Blockchain/ Hyperledger Fabric technology
- Development of predictive analytics models and techniques that will enable the forecasting of cabs arrival time and the optimization of the assembly planning

The scope of this deliverable is related with the third step. More precisely, the research and development in these first months of the Task 3.5 and the predictive analytics models and methodologies for the VOLVO case, were focused on techniques for the forecasting of cabs' arrival time and the prediction of possible delays.

3.5.2.2 Initial Data Model

An initial data model has been designed in order to describe and conceptualize the available big data that are related to the presented use case.

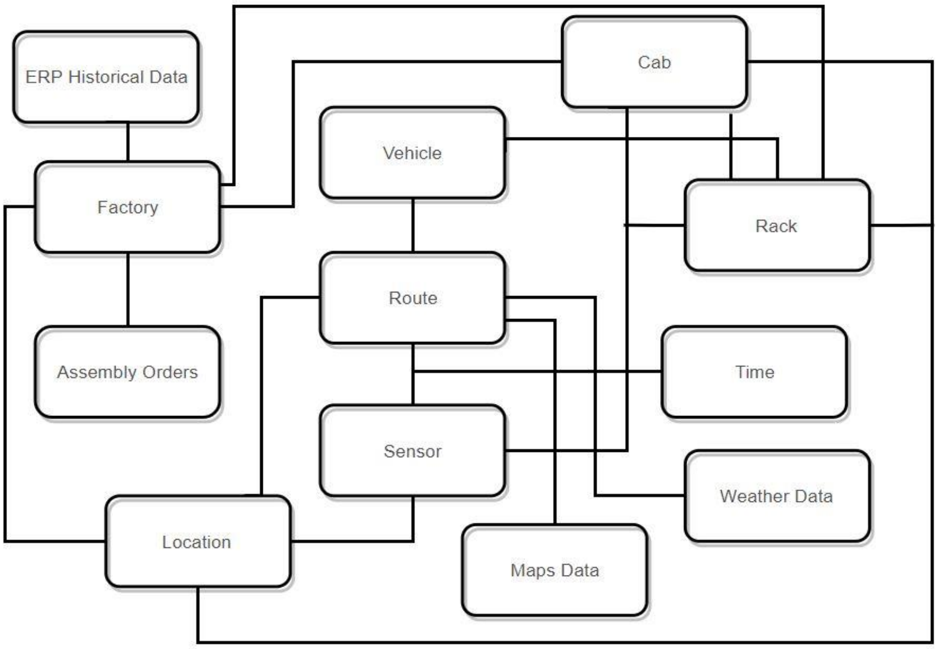

Figure 18 Initial Data Model - High Level Overview

Entity	Description
Factory	Contains the definition of a factory. A factory has a location, ERP data, orders, available cabs and racks
ERP Historical Data	Contains historical from ERP about previous transportations and assembly line details
Assembly Orders	Contains information about available or planned assembly orders in a plant
Vehicle	Represents a train or a truck, details about contained racks, cabs and sensors
Cab	Contains the definition of a cab, serial number, type details, connection with sensor and rack
Rack	Contains the definition of a rack, serial number, type details, connection with sensor and cab
Sensor	Contains information about sensor devices (GPS tracker, beacon etc.)

Location	Represent information about location data (latitude, longitude etc.)
Route	Represent the route of a transportation from one plant to another
Maps Data	Contains data related to maps (data from maps APIs etc.)
Weather Data	Contains data about weather during a transportation (weather APIs etc)
Time	Represent time information (timestamps, timetables etc.)

Table 2 Initial Data Model Entities Description

3.5.2.3 Forecasting Engine

A mechanism that will enable the prediction of cabs' arrival time or detect possible delays in the procedure of the supply chain, is presented at this point. The data of the monitoring of cabs and racks that will be available from the track and trace mechanism for the transportation of cabs in the VOLVO supply chain will lead to the development of an engine that will enable the forecasting of cabs arrival time and will enhance the decision making over the supply chain. The predictive analytics tool will be based especially on logged data that this tool will be able to query from the blockchain and a big data storage. In addition, many other factors will be taken into consideration by the analytics tool based on a vast amount of historical data coming from VOLVO. A large amount of statistical and machine learning methods will be applied in order to provide the highest possible level of accuracy of the forecasting estimations. The description of the prediction engine contains the following sections:

- Route Modelling
- Feature Selection
- Data Training and Prediction Model
- Abnormal Detection Techniques
- Model Evaluation

The conceptual architecture of the forecasting engine is depicted in Figure 16. Each element is described comprehensively in the following sections.

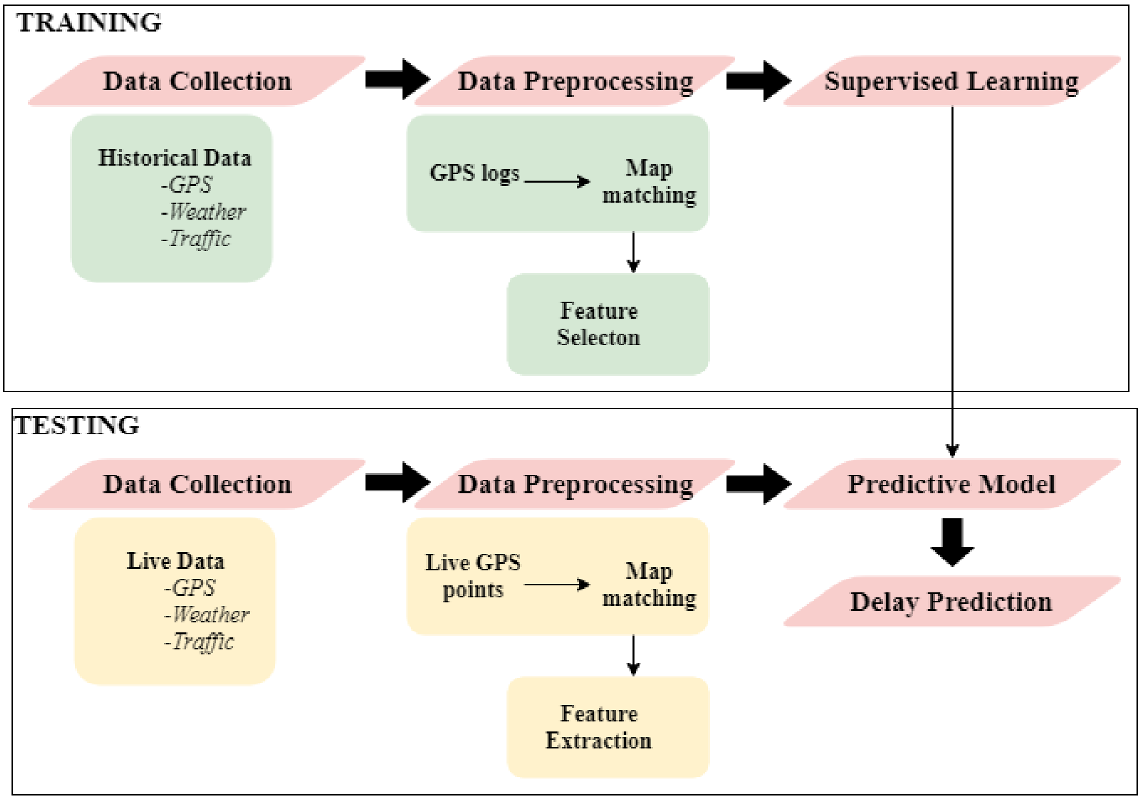

Figure 19 The architecture of the forecasting engine.

Route Modelling

In order to simplify and model the transportation route, the Umea to Tuve path will be segmented in smaller trajectories to achieve a better analysis and detection of the possible transportation delays. Distinct trajectories will be created based on the transportation route.

The GPS logged data will be partitioned and road segments will be created according to the longitude and latitude coordinates of each trajectory. Then, the separate GPS trajectories should be matched to the corresponding road segment. In this respect, **map matching algorithms** will be used to map each GPS point of a trajectory to the actual road segment and the GPS data will be translated into manageable information.

A range of intelligent transport system applications and services such as route guidance, fleet management, road user charging, accident and emergency response, bus arrival information, and other location based services require location information and use map matching algorithms [87]. Map-matching algorithms use inputs generated from positioning technologies (such as GPS) and supplement this with data from a high resolution spatial road network map to provide an enhanced positioning output. The general purpose of a map-matching algorithm is to identify the correct road segment on which the vehicle is travelling and to determine the vehicle location on that segment.

These algorithms can be divided into: matching position [88] and track curves match. The commonly used position matching methods include direct projection, probability statistics and the fuzzy logic method and so on. On the other hand, the commonly used track curve matching methods include geometric matching, correlation coefficient method and so on [89].

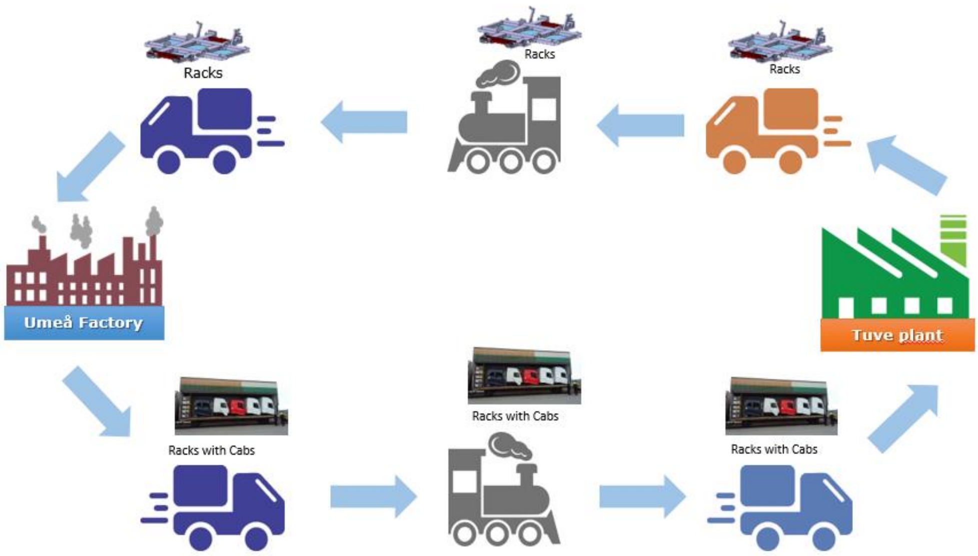

Figure 20 The transportation process flow. Each step represents a road segment.

The transportation route is depicted in Figure 16. The segmentation will occur based on the following diagram and the separate parts correspond to the following processes:

- Cabs for transportation are put on racks, transported from Umeå plant to station and be loaded on train.
- Cabs and racks are transported from Umeå station to Tuve.
- Cabs are put in cab trim line in Tuve plant.
- Cabs are transported from cab trim line to final assembly at Tuve plant.
- Racks return from Tuve to railway yard at Arendal Gothenburg.
- Racks return from Umeå railway to Umeå plant.
- Racks are put in repair process at Umeå plant.

The data analytics methods described below will be implemented in each part of the route, taking into consideration several factors that could affect the transportation time. Delays will be detected in each road segment that will lead to the overall delay prediction in the transportation flow.

Feature Selection

During the transportation of cabs and racks from Umea factory to Tuve plant and back, several events may happen and delay the whole process. The arrival time of the cabs depends upon the following factors:

- Lack of material in Cab production in Umeå
- Lack of racks
- Machine breakdown
- Product damage during transportation
- Weather conditions, e.g. snow on the rail
- Traffic

In order to proceed in the data analysis, distinctive features will be extracted from the above factors. This way the machine learning algorithms that are going to be developed could train faster, the complexity of the model will be reduced and the accuracy of the model will be improved. The aim is to reduce the number of features whilst maintaining high performance by discarding those least useful. Several methods are proposed in literature regarding feature selection and one of them is going to be implemented. Specifically, in Filter methods the features are selected on the basis of their scores in various statistical tests for their correlation with the outcome variable. In wrapper methods, a set of features is chosen, then the efficacy of this set is determined and some perturbation is made to change the original set and the efficacy of the new set is evaluated. Embedded methods combine the qualities' of filter and wrapper methods [90].

Data training and Prediction Model

For the development of prediction models, several state of the art techniques and methodologies will be used and tested for their predictive performance and efficiency. These methodologies are families of Artificial Neural Networks (ANNs) and Support Vector Machines (SMVs), along with Classification and Regression Trees (CART) (decision trees) and their ensemble approach (Random Forest), Markov chain (MC) and Hidden Markov Models (HMMs).

Artificial Neural Networks

Artificial Neural Networks (ANNs) is a popular approach to address complex problems, derived from sectors like health, industry, finance, transportation etc. Neural networks can be hardware-based (neurons are represented by physical components) or software-based (computer models), and can use a variety of topologies and learning algorithms. One popular supervised model is the Multi-Layer Perceptron trained with variations of the back-propagation algorithm (BPN). BPN is a feed-forward model with supervised learning [91].

Support Vector Machines

Support Vector Machines (SVMs) classifier is one of the most convenient and widespread classification and regression algorithm and it is first proposed by Boser, Guyon and Vapnik in 1992 [92]. SVM is a machine learning technique based on risk minimization. The main objective of SVM is to construct a hyperplane as a decision boundary as the maximum margin between classified classes based on Kernel functions. Several kernel functions have been deployed so as to improve the predictive performance of the SVM. In our work, we apply two Kernel functions: Polynomial and Radial Basis Function.

The Support Vector Machine - Radial Basis Function case relies on the Gaussian Radial Basis function kernel with its form given by: $K(x, y) = \exp(-\gamma\|x - y\|^2)$, where $\|x - y\|^2$ is the Euclidean distance between the feature vectors x and y and $\gamma = \frac{1}{2\sigma^2}$ is a positive constant with σ to be a free parameter. Along with σ, RBF kernel has another free parameter, constant C.

The Support Vector Machine - Polynomial case relies on the Polynomial function kernel. Its form is given by: $K(x, y) = [x^T y + \theta]^p$, where p is the degree of the polynomial and θ is a free parameter that usually takes its values form integer space, although $\theta = 1$ is preferable as it avoids Hessian matrix to become zero. As in RBF, a free parameter C is defined.

Decision Trees

Decision Tree learning, is a technique for approximating discrete-valued functions, in which the learned function is represented by a decision tree (or Classification Tree or Learning Tree). Decision trees can also be re-represented as sets of if-then rules so as to improve human readability. This tree-shaped structure is capable of generate classification rules for the tested dataset [93]. CART, is a recursive partitioning method that builds classificacion and regression tres for predictin continues dependent variables (regression) and categorical variables (classification). The classic CART algorithm was popularized by Brieman et al. 1984.

Random Forest

Random Forest also known as random decision forest, is an ensemble of decision trees and each decision tree is constructed by using a random subset of the training data, while the output class is the mode of the classes decided by each decision tree [94].

Markov Chain

A Markov chain is “a stochastic model describing a sequence of posible events in which the probability of an event depends only the state attained in the previous state”. Markov process is a stochastic process that satisfies the Markov property described above. A markov chain is a type of markov process that has either discrete state space or discrete index set. Markov chains have many applications as statistical models of real-world processes [95].

Hidden Markov Models

Hidden Markov Model is a statistical Markov model in which the system beeing modeled is assumed to be a Markov Process with unobserved, hidden, states. The Hidden Markov Model can be represented as the simplest dynamic Bayesian Network and it's developed by Baum et al. [96].

Abnormal Detection Technique

The analysis of trend in times series from sectors such as industry and transportation is a filed with increasing interest and it is researched a lot in nowadays. The prediction of the course of the linear trend of the monitored time series, even for a small time interval in the near future can be proved very significant in terms of machine breakdowns (predictive maintenance) and energy and cost savings. This section provides a brief description of a methodology called Slope Statistic Profile (SSP) for detection of the change point(s) of the linear trend in real time.

The SSP method estimates the change point (or the breakpoint *T*) from the profile of a linear trend test statistic, computed on consecutive overlapping time windows along the time series. The selected test statistic for linear trend estimation it is showed that gives high power compared to other test statistics for both correlated and white noise residuals. In the SSP approach, a first candidate breakpoint $\boldsymbol{T}$ is the time point at which the profile crosses the threshold line of rejection of the null hypothesis of no trend at $\pm\boldsymbol{t_{w-2,1-a/2}}$, where *a* is the significance level, *w* is the size of the sliding window and *t* follows the Student distribution with *w-2* degrees of freedom. The search of $\boldsymbol{T}$ is confined in a time interval corresponding to the profile segment bounded by $\boldsymbol{t_{w-2,1-a_1/2}}$ and $\boldsymbol{t_{w-2,1-a_2/2}}$ for positive trends and by $\boldsymbol{-t_{t_{w-2,1-a_2/2}}}$ and $\boldsymbol{-t_{w-2,1-a_1/2}}$ for negative trends, where significance levels $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ for two side test are 0.20 and 0.05, respectively [17]. The t-statistic for the parametric linear trend test that is used in SSP approach is $\boldsymbol{t = \frac{\hat{\beta}}{s(\hat{\beta})} \sim t_{w-2}}$, where $\boldsymbol{\hat{\beta}}$ is least square estimator for the trend parameter and $\boldsymbol{s(\hat{\beta})}$ is the estimated standard error of $\boldsymbol{\hat{\beta}}$. The null hypothesis of no trend is rejected at a significance level $\boldsymbol{a}$ if $\boldsymbol{|t| \geq t_{w-1,1-a/2}}$ [97].

Model Evaluation

In order to evaluate the forecasting models we will use some of the well-known evaluation techniques found in the field of machine learning.

k- Fold Cross-Validation

Cross-validation is a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single parameter called k that refers to the number of groups that a given data sample is to be split into. As such, the procedure is often called k-fold cross-validation. When a specific value for k is chosen, it may be used in place of k in the reference to the model, such as k=10 becoming 10-fold cross-validation.

Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model. It is a popular method because it is simple to understand and because it generally results in a less biased or less optimistic estimate of the model skill than other methods, such as a simple train/test split.

Receiver Operating Characteristic (ROC) Curve

A Receiver Operating Characteristic (ROC) Curve is a way to compare diagnostic tests. It is a plot of the true positive rate against the false positive rate. A ROC plot shows the relationship between sensitivity and specificity. It also expresses the test accuracy and the likelihood ratio.

4.4.2.4 Assembly Line Optimization

Since now, there were no information that could be provided to production planning departments for replanning and re-optimisation of cab trim and final assembly at the final assembly plant at Tuve and every minute of production stop resulted in a hugevast loss in terms of cost along the whole end-to-end supply chain. The forecasting engine described above will give the option for rescheduling and planning alternative solutions for the assembly line considering that the transportation delays will be predicted.

There is a vast amount of data that can be used to the assembly line optimization process, although they are not available yet so as to prepare a fully adequate optimization model. At this point, the problem of **Assembly Line Balancing** is mentioned along with the methods of addressing it that can be found in literature. In general, the assembly line balancing is the procedure of deciding how to assign tasks to workstations. The goal is to obtain task groupings that represent approximately equal time requirements in order to minimize idle time and result in a high utilization of labor and equipment. Several surveys focus on cost and profit based assembly line balancing. Problems, approaches and analytical models on assembly line balancing that deal explicitly with cost and profit oriented objectives are analysed in literature. In [98], cost components, analytical models and the solution algorithms are comprehensively discussed. A more concrete study is presented in [99], an investigation of the potential use of Differential Evolution on the simple assembly line balancing problem. This observation stems from a literature review performed with the aim of positioning DE metaheuristic within the relevant research field. Another approach is presented in [100]. The hybrid grouping genetic algorithm incorporates the concepts of constructive heuristics and GGA to enhance intelligent iterative search of the solution space. The hybrid algorithm, though primarily developed for solving the simple assembly line balancing problem, can be extended to several other variants, with minor adjustments.

3.6 Big Data Models, Techniques and Methodologies for Smart Maintenance & Service Pilots

In order to support the creation of Smart Maintenance and Service pilots, the BOOST 4.0 project will use and implement innovative Big Data cognitive manufacturing processes, predictive methodologies and algorithms that enable:

- The production planning, storage and management of spare parts
- Large-scale predictive maintenance

3.6.1 Production Planning for Whitegoods Spare Part

Big Data models and analytics methodologies for the Whirlpool production planning aims to the optimization of the spare parts planning and the distribution process in the EMEA region through the full adoption of a prediction tool and techniques that can help the organisation to better estimate when and where a spare part is needed . The procurement of the components, the production and distribution of a spares is currently triggered by a statistical forecast based only on historical demand of the spares without taking into consideration other important endogenous variables that can contribute to explain what is really happening in the market.

The relevant information identified and needed should be derived from big data coming from different departments of the organisation:

1) from every single factory that is producing the appliances we can retrieve the test data, estimate the life duration of the appliance and intercept every change on the components that may affect the defectiveness of the different models produced; Moreover every change applied into the production of an appliance to fix defectiveness is directly connected to a decries of spare part consumption in the future;

2) from the market operations we can retrieve the sell in and the data coming from the usage of the connected appliances in order to initiate a predictive maintenance approach and estimate the overall installed base;
3) from consumer service field function we can retrieve the data registered after the service calls in order to estimate the real defectiveness of our appliances in the field and also feedback factory and engineers to take immediate corrective actions.

3.6.1.1Spare Parts Data Model

The data model is designed to manage data about any device that is composed by several spare parts. If a predictive model needs data that are specific of a particular device (ex: Absorbed electric current for a washing machine), there is a specific entity (Device Attributes represented in green in the diagram below).

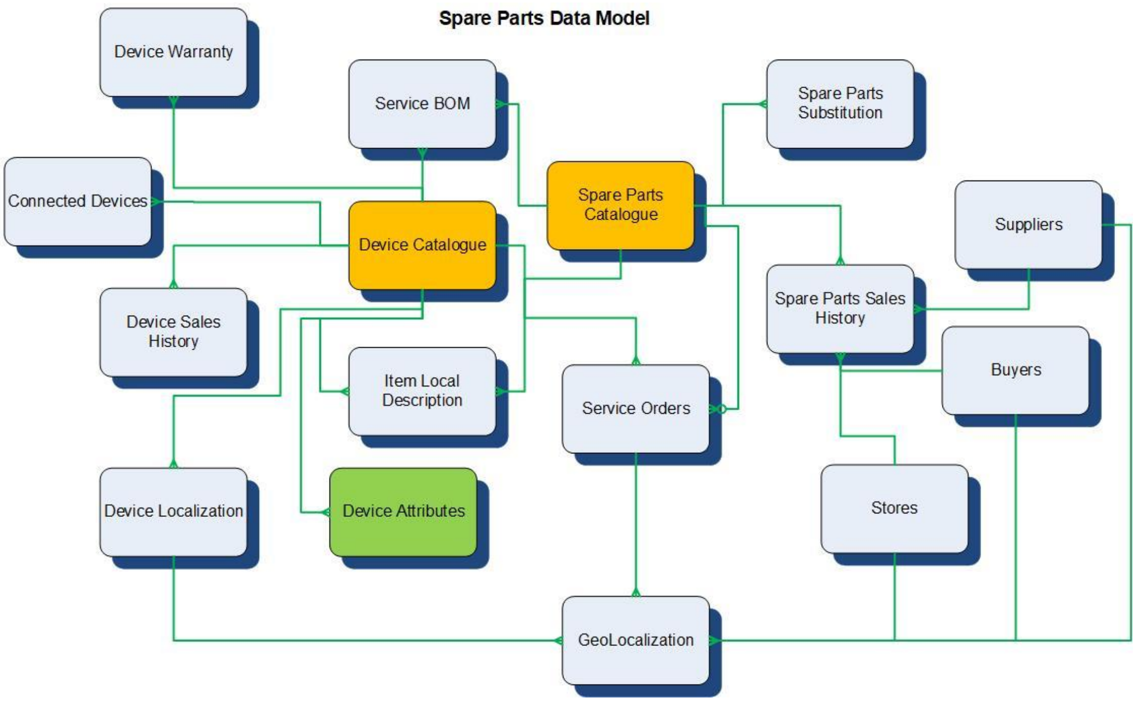

Figure 21 Data Model for Spare Parts

This Data Model is designed for the following goals:

- Spare parts sales forecasting. The forecasts are based mainly on spare parts historical sales data, but they may consider also:
  - Device historical sales data
  - Service Order historical data
  - Environmental context (based on geolocalization)
    - Predictive maintenance on connected devices. These are predictive models used to detect potential anomalies on devices. They are

based on physical parameters measured by sensors on connected devices.

Entity	Description
Spare Parts Catalogue	Contains the definition of a Spare Part
Spare Parts Sales History	Contains historical sales data about spare parts
Item Local Description	Contains entities local language description. This useful when companies are multinational
Spare Parts Substitution	Spare parts substitution is a service table to manage identifier substitution of a spare part over its lifetime
Device Catalogue	Contains the definition of a device
Device Sales History	Contains historical sales data about devices
Device Warranty	Contains information about type and duration of warranty period, if applicable.
Service Bill of Material	Contains the relation between a device and its spare parts
Service orders	Contains the maintenance activities requested on a device, including a spare part substitution
Connected devices	Contains data generated by device sensors and sent to the company through an internet connection
Device Localization	Contains, when available, a reference to its geo-localization
Device Attributes	Contains the definition of specific attributes of a device, which are related to generic field names inside the database
Supplier Catalogue	Contains the definition of all suppliers of Spare parts
Buyer Catalogue	Contains the definition of all buyers of Spare parts
Store Catalogue	Contains the definition of all stores providing Spare parts

Geo-localization	It is a reference table containing information about geographic areas, which may have influence on predictive models
Spare Parts Forecast	Contains forecasts generated by predictive models
Forecast Performance	Contains metrics about forecast accuracy over time

Table 3 Spare Parts Data Models Entities

3.6.1.2 Data Management Techniques

Data Management tools provided by SAS consists of:

- A set of database connectors to most of commercial database on the market, including Hadoop distributions (Cloudera, Hortonworks).
- SAS Data Management Languages (BASE SAS, SQL, FedSQL), which provides hundreds of functions for data processing and transformation
- Data Management Applications (Data Integration Studio, Enterprise Guide, etc.), which provides user interfaces for data processing flows generation, without code writing.

In the BOOST 4.0 project the Data Management tools are used for:

- Extract, transform and load data from Whirlpool Data Lake into Spare Parts Data Model
- Extract, transform and load data from Spare Parts Data Model into Analytical Base Tables used by Analytical and Reporting Tools
- Load results from predictive models into Spare Parts Data Model

SAS DIStudio provide a visual interface to design data integration processes. SAS supports data integration in the following ways:

- Connectivity and metadata. A shared metadata environment provides consistent data definition across all data sources. SAS software enables you to connect to, acquire, store, and write data back to a variety of data stores, streams, applications, and systems on a variety of platforms and in many different environments. For example, you can manage information in Enterprise Resource Planning (ERP) system, relational database management systems (RDBMS), flat files, legacy systems, message queues, and XML
- Data cleansing and enrichment. Integrated SAS Data Quality software enables you to profile, cleanse, augment, and monitor data to create consistent, reliable

information. SAS Data Integration Studio provides a number of transformations and functions that can improve the quality of your data

- Extraction, transformation, and loading. SAS Data Integration Studio enables you to extract, transform, and load data from across the enterprise to create consistent, accurate information. It provides a point-and-click interface that enables designers to build process flows, quickly identify inputs and outputs, and create business rules in metadata, all of which enable the rapid generation of data warehouses, data marts, and data streams
- Migration and synchronization. SAS Data Integration Studio enables you to migrate, synchronize, and replicate data among different operational systems and data sources. Data transformations are available for altering, reformatting, and consolidating information
- Data federation. SAS Data Integration Studio enables you to query and use data across multiple systems without the physical movement of source data. It provides virtual access to database structures, ERP applications, legacy files, text, XML, message queues, and a host of other sources. It enables you to join data across these virtual data sources for real-time access and analysis. The semantic business metadata layer shields business staff from underlying data complexity.

SAS data integration projects have a number of advantages over projects that rely heavily on custom code and multiple tools that are not well integrated.

- SAS data integration reduces development time by enabling the rapid generation of data warehouses, data marts, and data streams.
- It controls the costs of data integration by supporting collaboration, code reuse, and common metadata
- It increases returns on existing IT investments by providing multi-platform scalability and interoperability
- It creates process flows that are reusable, easily modified, and have embedded data quality processing. The flows are self-documenting and support data lineage analysis.

Execution Environment

For technical experimentation, data will be provided on flat files. All extraction and transformation processes will be executed by SAS Base Engine, while Analytical tools will make use of CAS in-memory engine.

For large scale experimentation, data will be provided on HDFS file system. All extraction and transformation processes will be executed by Hadoop (Hive or Impala) engine invoked by SAS programs. Data transfer between Hadoop and CAS will be performed through massive parallel process, that will take advantage of SAS libraries on Hadoop and the co-located architecture. Analytical tools will make use of CAS in-memory engine.

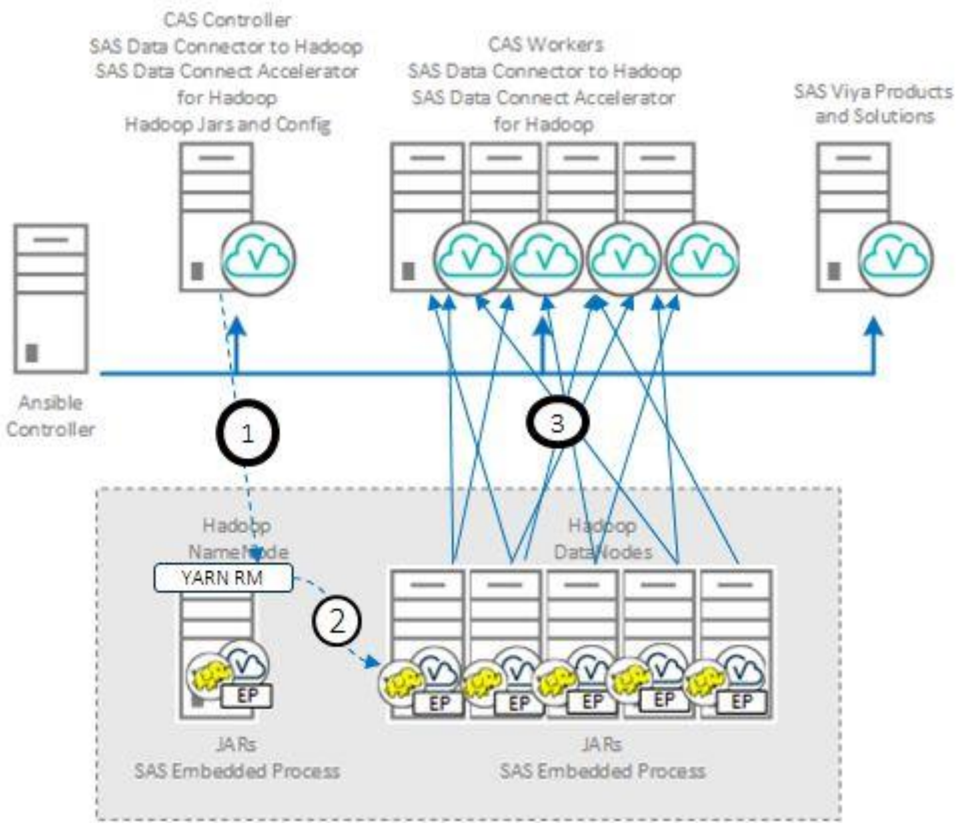

Figure 22 SAS platform and Hadoop Integration

3.6.1.3 Analytical Methodologies for the Production Planning

Different types of analytical techniques will be used to solve the problem of an effective production planning. In particular, the analysed entity (the spare parts) has been clustered on the basis of their information. From historical data of spare sales, time series are created. Time series of the dependent variable (sold quantity) are related to other input variable (regressors). Some explorations of data will be done to understand the quality and find insights. According to the type of spare, different types of forecast models will be used.

It will be possible to apply different anomaly detection models for data coming from connected devices. In order to identify unusual patterns that do not fit expected behaviour, unsupervised or supervised machine learning models will be used.

The Large-scale time series analysis and forecasting models available for the use case are listed below:

Time series analysis

- Autocorrelation analysis
- Cross-correlation analysis
- Seasonal decomposition and adjustment analysis
- Count series analysis
- Diagnostic tests for seasonality, stationarity, intermittency and tentative ARMA order selection.

Time frequency analysis

- Windowing functions
- Fourier analysis for real and complex time series
- Short-time Fourier analysis
- Discrete Hilbert transform
- Pseudo Wigner-Ville distribution

Time series modeling

- ARIMA models (dynamic regression and transfer functions)
- Exponential smoothing models
- Unobserved component models
- State-space models
- Intermittent demand models with Croston's method

Neural network/machine learning modeling strategy nodes

- Panel series neural network framework for generating features and training a neural network
- Multistage (neural network/regression + time series) framework for creating a forecasting methodology that combines signals from different types of models
- Stacked model (neural network + time series) forecasting addresses problems that have both time series characteristics and nonlinear relationships between dependent and independent variables

Singular spectrum analysis (SSA)

- Univariate SSA decomposition and forecasting
- Multivariate SSA
- Automatic SSA

The unsupervised machine learning models are listed below.

- k-means & k-modes clustering
- Principal component analysis
- Standard PCA
- Robust PCA
- Moving Window PCA
- Support Vector Data Description

The supervised machine learning models available for the use case are listed below.

- Logistic regression models
- Ordinary least squares models
- Generalized linear
- Partial least squares regression models
- Quantile regression models
- Nonlinear regression models
- Neural network models
- Decision tree
- Forest
- Gradient boosting
- Support vector machine
- Factorization machine models
- Boolean rule model

3.6.2 Large-scale Predictive Maintenance

3.6.2.1 Overview

The focus of this use case is Smart Maintenance for the Benteler hot forming line. The main task of the hot forming line is stamping of sheet metal into a three-dimensional shape. The metal is heated before stamping, and rapidly cooled down during stamping. This causes the material to be hardened, which is important to structural components for the automobile industry. A fully automated On-Loading process places pre-cut sheet metal plates onto a conveyor belt. They are transported into a furnace with defined temperature profiles in order to get a controlled heating process. Heated plates are loaded by a feeder into the hydraulic press, where the main forming takes place, and handled by an offloading feeder thereafter.

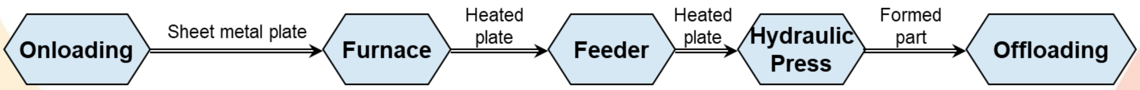

Figure 23 System components of the hot forming Line

Within the hot forming line, two sub-systems are considered for closer examination. The first Business Process considered is the maintenance process in case of hydraulic oil leakage. Oil leakage is considered as any involuntary loss of oil within the hydraulic subsystem of the hydraulic press. This can be caused by three failure modes. Minor leaks at sealings can be repaired by replacement of sealings, o-rings etc. This can also be prevented by early replacement and regular maintenance processes of these parts. Minor leaks are characterized by gradual loss of oil within the hydraulic subsystem. Major leaks are caused by structural damages of machine components, e.g. intake and exit oil lines, press head or hydraulic components as valves and pumps. These defects are much more difficult to repair and mostly characterised by disruption or sudden change of oil flow. A third possible failure class are spillages caused by the press tool. These are characterized by product-specific, minor gradual oil losses. Since this failure mode is dependent on the tool, a change of oil loss may be observed after product change. It can be prevented by maintenance of the tool before or after tool change.

Main effect of the failure mode is an oil loss within the hydraulic subsystem. The amount of oil loss depends on the failure mode and can range from minor loss which lies below the threshold detectable by level sensors, up to breakage of oil lines which lead to significant oil loss and malfunctioning of the system. Major consequence is an unplanned downtime of the press line for repair of the cause of oil loss as well as for oil refill. The different failure modes and their severity lead to different repair times, ranging from simple maintenance routines (replacement of sealings) up to major repairs with an extent that is hard to estimate. If the hydraulic system operates at low oil level for a significant amount of time, this can also lead to potential damages of hydraulic components.

This business process is heavily affected by the installment of predictive and smart maintenance processes. Main objective of smart maintenance algorithms is the detection (condition monitoring) and prediction of oil leakages. Maintenance processes can be triggered much faster or even in advance. Maintenance workers are also supported by detailed description and diagnosis of the failure, giving details about the location of failures or maintenance activities, reasons for faults and type of failure mode. Maintenance and repair activities can be planned more efficiently and manual diagnosis is prevented.

Failure mode:	Hydraulic oil leakage ▪ minor leaks at sealings ▪ Major leaks caused by cracks in machine components ▪ Spillage caused by tool
Failure effect:	▪ oil loss within the hydraulic subsystem ▪ Downtime of line for repair / maintenance (oil refill) ▪ Potential damage of hydraulic components
Future Scenario:	▪ Detect and Predict oil leakage ▪ Describe oil leakage (spillage, minor/gradual, major/sudden) ▪ Diagnosis (machine leakage or tool leakage?)

Figure 24 Summary of failure in case of hydraulic oil leakage

The scrap belt is connected to several lines and runs underground the BENTELER factory hall. Any scrap metal that is accumulated during the production process along the hot forming line is placed through a funnel onto the scrap belt. Scrap metal parts are then transported from the production line to a scrap metal container, where it is then taken to recycling.

The correct functioning of the scrap belt is crucial to production, since a halt of the scrap belt means a potential halt of several production lines. Restoring the function of the scrap belt is a tedious process, since the cause has to be detected along the whole scrap belt which spans several production lines. Fault diagnosis takes place under harsh conditions. Examination has to take place manually in the underground tunnel with limited access and restricted space. Failure of the scrap metal belt is caused by pile-up of excess metal during the transport. Also, scrap metal pieces can be jammed between the scrap belt and walls or ceiling of the underground tunnel.

Primary effect of the failure is an increased motor force which leads to an increased current and higher energy consumption. If this failure remains undetected, the condition of the scrap belt can be severe, leading to an unexpected halt of the scrap belt. The failure depends on the amount of scrap metal produced at each line, as well as the shape of scrap metal parts and the placement onto the belt.

In a future scenario, predictive maintenance algorithms detect pile-up of scrap metal in advance, and thus allow to take counter-measures in order to prevent a halt of the scrap metal belt. In the case of a complete halt, maintenance can be triggered much faster by continuous condition monitoring. In addition, diagnostic algorithms can give maintenance or repair advice, giving possible locations for the cause of the failure. This significantly limits the manual fault diagnosis in the underground tunnel, and allows a faster restoring of the regular belt function.

Failure mode:	▪ Pile-up of scrap metal during transport ▪ Jamming of scrap metal pieces between scrap belt and underground tunnel
Failure effect:	▪ Increased motor current for scrap belt ▪ Unexpected halt of scrap belt
Future Scenario:	▪ Detect and Predict pile-up of scrap metal ▪ Maintenance advice

Figure 25 Summary of failure in case of scrap pile-up of the scrap metal belt

3.6.2.2 Big Data Models, Techniques and Methodologies for the Use Case

As it is presented in the previous section, the Hot-forming Hydraulic Press and the Scrap Belt, are the components of the Benteler's production line, that are going to be supported by advanced analytics, machine learning and data mining algorithms for predictive maintenance and fault detection. Figure 23 gives an overview of the software architecture for the smart maintenance algorithms alongside the process that the technology providers of this trial are going to follow, in order to achieve a successful implementation. As it is depicted in the Figure, the technology providers are going to follow several steps, starting from the "Analysis of the Maintenance Processes" up to actual "Model Deployment" in the production line, in order to enhance the maintenance process of the Hydraulic Press and the Scrap Belt. Despite the fact that the two components (i.e. Hydraulic Press and Scrap Belt) differ on the physical layer, the applied maintenance policy is identical for both of them.

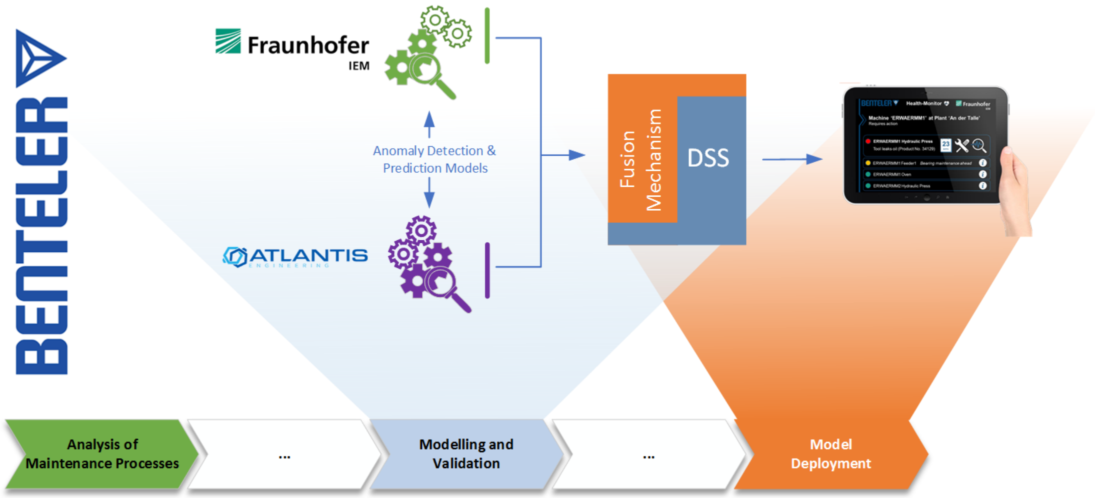

Figure 26 Overview of the enhanced Benteler's maintenance processes with Machine Learning algorithms

Maintenance processes will start from analysing the maintenance processes and technical systems. For this, processes and systems will are modelled in a semi-formal, graphical notation. This allows an easy understanding of systems and processes. Possible faults and their propagation through systems and processes can be easily identified and traced. This makes root cause analysis easier.

Considering the "Modelling and Validation" step, the technology providers are going to contribute in the predictive maintenance and anomaly detection (for fault detection) processes, proving appropriate models. The provided models will be combined using a fusion mechanism to provide a single recommendation of an action or a set of actions through a Decision Support System (DSS) Solution. Fraunhofer IEM will provide rule-based fault detection methods based on the knowledge from the previous analysis of processes and systems, that will be augmented by online learning through data during runtime of the software. In the context of Benteler's use case, ATLANTIS will provide a fault detection tool (FDT). This tool will be supported by an outlier detection algorithm on streaming data, to analyse data from IoT devices and machine data and detect potential faults in the production line at real time. ATLANTIS will also provide a predictive maintenance tool (PMT) to predict failures in the production process, based on prior knowledge.

For the FDT, a state-of-the-art algorithm for outlier detection on streaming data [101] will be used. This algorithm is able to analyse efficiently Big Data. The algorithm is provided in different programming languages like C++, Java and Scala. The latter is compatible with the Flink analytics and Big Data software [102]. The outlier detection algorithm will be able to utilize domain knowledge to pre-process/clear the incoming data and provide more accurate results. A rule-based mechanism will be also coupled with the outlier detection algorithm, as a backup mechanism, to further prevent unwanted events. The focus will be on solutions that are as unsupervised as possible from the machine learning perspective.

For the PMT, planned to be applied an approach similar to [103], where a supervised machine learning algorithm is used (i.e. random forests), together with a risk function to form a regression problem. In [103], authors analyse failure events from known failure modes. To provide a more flexible solution, in cases where the fault events are not available, technical provedrs are planning to utilize motif detection and/or outlier detection algorithms to produce such events that will be fed to an ensemble predictive algorithm. The solutions will be tailored to real-world Industry4.0 settings, where the fault events are very rare compared to the normal events, and there is no crisp knowledge as to which factors lead to/cause a problem.

A semi-automated online-learning is under research. The specified failures for both the business cases (i.e. oil leakage for the Hydraulic Press and metal jamming for the Scrap Belt) are rare events. Supervised learning methods are very difficult to train in the early phases, since no labels are available for training data sets. Rules can be deducted from the expert knowledge gathered in the analysis phase, in order to created unsupervised, rule-based fault detection methods. The output of these methods, combined with additional assessment by experts, can then be used for labeling of the time series data. The resulting labelled data can then be used to train supervised classificators during runtime of the algorithm.

The proposed models need to be able to handle online streaming data in order to provide up-to-date input and early alerts to the maintenance staff. Batch processing solutions can also be applied in a periodic manner, as a secondary mechanism in order to strengthen the results dependability. The streaming processing of imbalanced data is an open research challenge. The proposed solution is important to be able to adapt to changes in the application environment and to be easily applied on new production sites. Unsupervised and Reinforcement learning provide self-training capabilities that address this demand.

A fusion mechanism that will combine the output from different predictive and/or anomaly detection models, taking into consideration the specificities of each business case and through a Decision Support System, will provide useful output to the maintenance experts. The ensembled solution will include combinations that consider the hybrid usage of different models. The outputs of multiple models can be fused by assigning a particular probability to each individual model, or using a parametric meta-model. The evaluation of the predictive models needs to take into consideration the imbalanced domain.

After the model deployment, research will be contacted on model transfer. This includes transfer of the trained models on the algorithmic level, but also by means of the underlying development process. The previously described steps will contribute to design a standardized developing process for the overall implementation of smart maintenance application. This will further decrease implementation time for future scenarios and allows to transfer knowledge and algorithms to similar processes and technical systems. Data-driven modeling techniques for transfer learning will be researched, which further allows the re-utilisation of previously trained models for similar plants or similar machines. This will also decrease development and implementation time for other End users. The designed processes can be validated by implementation of smart maintenance processes in other Benteler plants.

4 Big Data Analytics Platforms

This Section aims at providing an overview of the commercial Big Data Analytics Platforms that will be extended & integrated with the Boost4.0. It covers functionalities offered, discusses needs for data governance and existing governance mechanisms and also outlines integration needs and potential extensions to enable applicability in the manufacturing domain.

The Boost 4.0 big data analytics platforms embrace the idea of architectural plurality and hybridation. The big data platforms are applied based on the value required by the implementation of the specific business processes.

	Central Architectures (e.g. Data Lakes)	Federated Architectures (e.g. Industrial Data Space)	Distributed Architectures (e.g. pure Blockchain)
Data Ownership	Central or distributed	Distributed	Distributed
Data Stewardship	Central or distributed	Distributed	Distributed
Data Capture and Creation	Distributed	Distributed	Distributed
Data Storage	Central	Distributed	Distributed, redundant
Data Enrichment and Data Preprocessing	Central	Distributed	Distributed
Data Integration and Fusion	Central	Central (e.g. through Linked Data and Data Space approaches)	Distributed
Data Sovereignty	Central (if any)	Distributed	Distributed
Data Provenance	Central (if any)	Central	Distributed
Data Brokering, Clearing, Billing	Central	Central	Distributed

Figure 27 Main architectural features for different data exchange and sharing patterns.

Every Boost 4.0 pilot holds a clear business value and the extensions described in the following Sections aim at bringing that value forward.

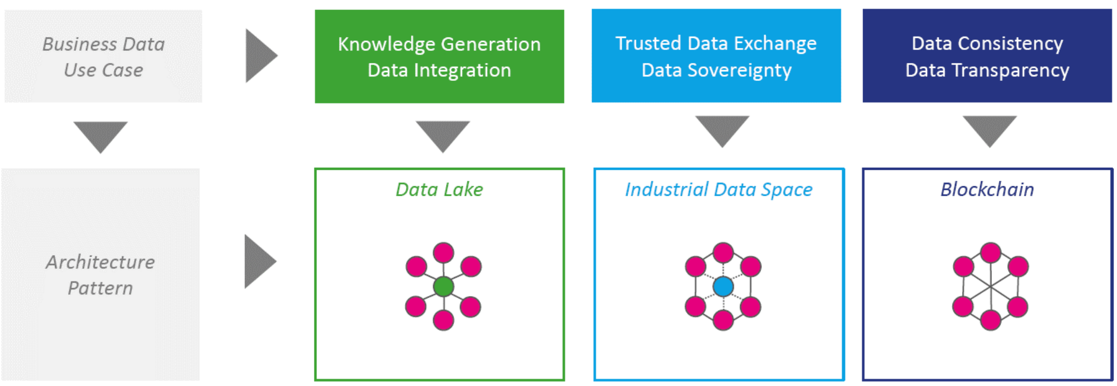

Figure 28 Business value targeted by each architectural pattern

4.1 Networked engineering big data analytics platforms

4.1.1 RISC framework data analytics platform

4.1.1.1 Overview

The RISC framework includes the tools CALUMMA and Collibri mentioned in the application.

Collibri focuses on the integration of large data volumes as well as continuous data streams. Further developments outside of the project Boost 4.0 can be read in the big data stack afterwards.

CALUMMA is an ontology-based research infrastructure for domain expert driven knowledge discovery. The main principle behind the system is to put the domain experts with their knowledge and experience in the centre of the knowledge discovery process. Therefore the domain experts models his domain of research as a domain ontology. CALUMMA adapts itself to this domain ontology at runtime and appears to the users like an individually programmed software for their domain. Furthermore, the elaborate meta-information from the domain ontology is used to actively support the domain experts (who are not necessarily IT experts) in complex data science tasks, such as data preparation, data selection and filtering, data cleaning and plausibilization, data visualization and analytics.

CALUMMA is mainly based upon Java enterprise technology and uses MariaDB as internal data storage system. The system is also able to access data from external system either by importing it into its internal data storage or accessing the external data at runtime. It the curse of the BOOST project CALUMMA can be used in manifold ways. On the one hand, it can be directly used as data analytics platform by the domain experts at FILL. On the other hand, it can be used as a (backend) component for a bigger data analytics framework.

CALUMMA supports a large variety of data visualization techniques. The all have in common that they are easily accessible by the user and strongly automated due to the elaborate meta-information from the underlying domain ontology. All visualization are highly interactive. The user is always enabled to select and process the data behind each single visualization element (bar in a barchart, dot in a scatter chart, etc.). The system also supports a number of dimensionality reduction algorithms which allows the user to project a high dimensional research data frame into a two dimensional display in order to explore the data.

The *big data stack of the RISC framework* is designed as follows:

Big data streams (e.g. sensor data of the production process or machine operation) are stored in the following way: At the backend, Apache NiFi is being used to receive the online data, which is being stored into an HBase NoSQL data store , which is being accessed through Phoenix for query speed. To enhance the flexibility and applicability of the system additional input data connectors will be implemented relying on standard importers and protocols such as OPC UA.

Model information of the machines is stored in the form of ontologies in a generic database system. The ontologies are generated by means of a graphical editor as well as further processing. Additional data such as, for example, ERP data and machine configurations are stored in the generated data model by means of the ETL process.

RISC Software relies on the open source software collection Apache Hadoop for data storage and computation. Running on a cluster of computers, Apache Hadoop provides a software framework for distributed storage and processing of big data using the MapReduce programming model. Being originally developed for commodity hardware, it has also found use on clusters of higher-end hardware. All modules were designed assuming that hardware failures occur frequently and should be automatically handled by the framework, making it a robust and reliable solution.

The core component for storing data is the Hadoop Distributed File System (HDFS). Files are split into blocks and distributed across nodes in the cluster. HDFS achieves reliability by replicating the data blocks across multiple hosts, and thus RAID systems are not advisable for the data storage partitions.

To process data in parallel, Hadoop sends packaged code to each node and takes advantage of data locality, meaning that each node only manipulates the data it has access to. This was found to be more efficient than conventional super computer architectures that rely on parallel file systems and distribution via high-speed networking. Beside the base modules, RISC software frequently uses Apache HBase, Apache Phoenix, Apache Hive, and Apache Spark.

HBase is an open-source, distributed NoSQL database running on top of HDFS. As a column-oriented key-value data store it can handle large quantities of sparse data. It is well-suited for fast read and write operations on large datasets with high throughput and low latency. It also features data compression and in-memory operation. Tables in HBase can serve as input and output for MapReduce jobs that run in Hadoop.

Apache Phoenix is a SQL layer on top of HBase. It provides a JDBC driver that allows our customers to conveniently access and modify their data with various analytics and business intelligence applications via SQL.

Apache Hive provides SQL layers for various file formats and file storage methods as well as NoSQL Databases that integrate with Hadoop, including Apache Phoenix. This means the unified SQL language HiveQL can be used to send SQL queries against a wide variety of data stores based on HDFS.

RISC Software relies on Apache Spark for cluster-computing. It provides high-level APIs in Java, Scala, Python and R and the machine learning library MLlib. Spark keeps track of the data that each of the operators produces, and enables applications to reliably store this data in memory. This is the key to Spark's performance, as it allows applications to avoid costly disk accesses.

Spark programs can be submitted from outside to the cluster using Livy-Sessions, for example from within the web-based IPython interpreter Jupyter. This way, after the cluster has done some heavy lifting, results can be plotted immediately using python modules such as matplotlib. This allows for a rapid development of data analysis algorithms (e.g. for condition-monitoring, predictive maintenance, etc.)

4.1.1.2 Data Sovereignty

CALUMMA is mainly based upon Java Enterprise technology. Its internal meta-data-model is stored in a relational database system, Maria DB. The CALUMMA Core component encapsulates the data access layer and the business logic. All other modules are built upon the core module. The CALUMMA Management Tool is used for modelling the domain ontology, manager users with usergroups and privileges, and for data processing and data analytics.

The web interface is designed for interaction with human users, while the data integration model is able to integrate structured data from electronic data sources. The REST API is used to connect CALUMMA at runtime with external systems and applications.

4.1.1.3 Extensions/Integration

As lots of the data of the FILL pilot will be time series, the goal will be to find anomalies in the data. Services will be added related to Time Series Analysis, Feature and Model Selection, Data Preparation and Benchmark Evaluation. All these services were described in details at the previous chapter.

<u>(European) Industrial Data Space</u>

In the context of the *RISC framework*, components from the FIWARE ecosystem can be extended as a basis for implementation to increase its flexibility, while keeping the implementation effort low.

In addition, a connection to the European Industrial Data Space (EIDS) is planned to extend the applicability of the RISC framework. The connection with the EIDS will clarify the possibilities and the value of the RISC framework as a Visual Analytics Tool. The connection to the EIDS can be realized by a special certified connector that can exchange data between the RISC framework and a EIDS node. This extension corresponds to the modular character of the RISC framework, where some data connectors, converters and memory modules as well as components for modeling machine behavior and visual analytics belong.

Another important aspect of interoperability between RISC framework and the EIDS is implementing the data sovereignty policies as given by the EIDS, and therefore ensuring the adherence to the applicable data sovereignty rules. The connector itself will (at least partially) be is based on the implementation of FIWARE to increase the speed of development and interoperability.

Within the FILL pilot an integrated business process for smart digital engineering using big data extends the V-model. (The V-model describes the development of mechatronic systems based on a systematic analysis of the requirements and a distribution of the requirements and loads among the individual disciplines.)
Within the project Boost 4.0 the above described EIDS connector for the data harvesting process and the model development.

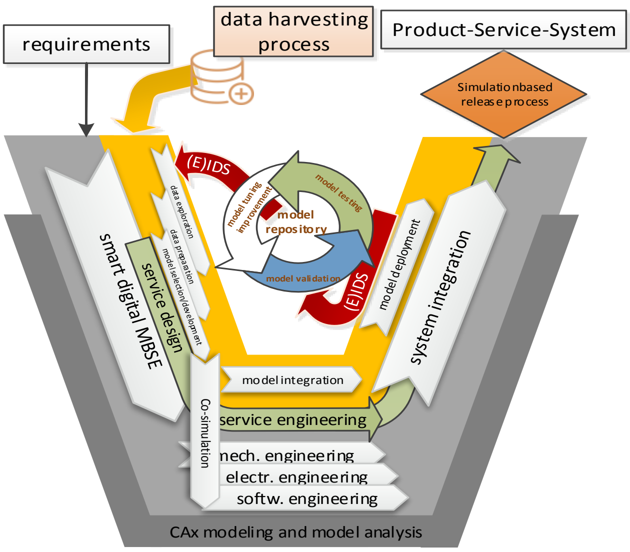

Figure 29 (E)IDS and RISC framework integration

4.1.2 ESI hybrid twin cloud platform

4.1.2.1 Overview

ESI delivers Virtual Prototyping as the foundation for the Hybrid Twin™, synchronized through the Internet of Things (IoT) and Big Data to receive real-life product feedback to maximize the longevity of a product through intelligent predictive maintenance. With ESI Cloud, its new web-based platform, ESI also aims to deliver Virtual Prototyping and data analytics.

The ESI Cloud platform will be accessible from a common WEB portal. This WEB client will be used for simulation data management and services.

Each of these server applications are accessible through exposed APIs via Web services for ESI Cloud platform.

To enable the models architect to perform simulations, the platform shall also be interfaced with 3rd party applications, the solvers in particular. In other words, the platform will manage a list of applications, and will be able to launch any of these registered applications.

In terms of data management, all the data as well as the relationships between these data will be stored in ESI Cloud platform.

As for the simulation data, all the simulation models will be managed by the ESI Cloud platform application that will also keep track of the simulations performed with these models and the related results.

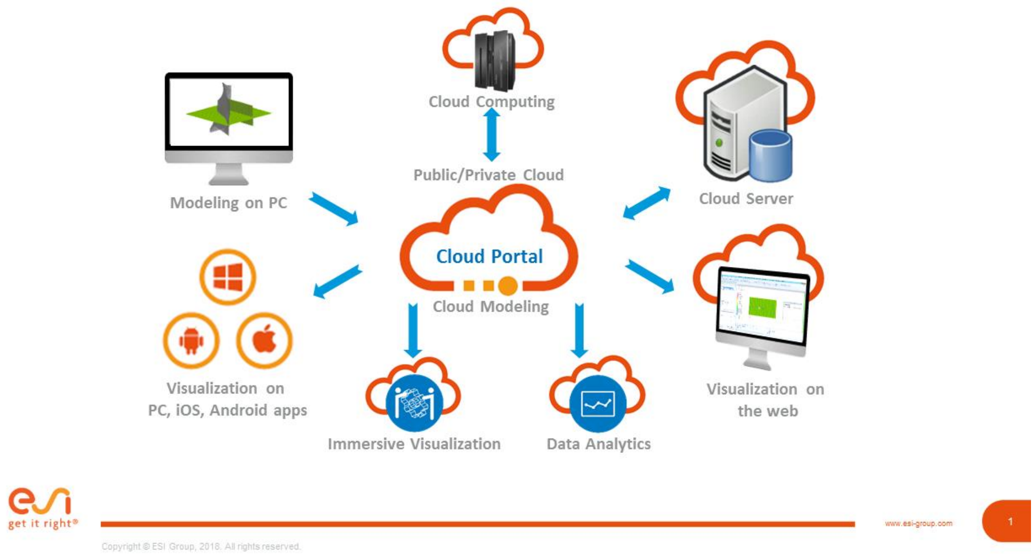

Figure 30 A Collaborative Cloud Modelling Platform for Virtual Prototyping

This application will be used in cooperation with VW to develop the CAST use case.

4.1.2.2Extensions/Integration

In the context of the VW pilot (WP4) dedicated to Smart Digital Engineering, we plan to:

- extend our ESI Cloud platform in order to fulfil business scenarios presented by the pilot:
    - Zero Defects Production
    - Feedback to mould design

    The extension will concern the design of a casting workflow which will embed the Hybrid Twin™ approach.

- make our ESI Cloud platform compatible with EIDS.

4.2 Cognitive production planning big data analytics platform

4.2.1 Visual Components 4.0

Visual Components 4.0 is a 3D Simulation and Visualization platform from Visual Components Oy, which allows create, visualize, validate, optimize and virtually commission production systems. The platform provides an intuitive interface to easily build any factory layout at different levels from a simple machine to the entire factory plant. The digital twin created in Visual Components mirrors the real factory layout to simulate and visualize all the production flows, logistics, automation and robotics.

Visual Components 4.0 provides two application APIs, which allows to the users tailor and configure its own solutions. In addition, the Communication interface enables the interoperability with robotic systems (such as UR and Staubli) as well as automation systems through OPC UA.

4.2.1.1 Overview

Components 4.0 is a desktop 3D Simulation and Visualization platform that allows creating virtual factory layouts at different level of complexity, from a simple machine to the entire factory. The digital twin created can be interconnected with the real factory systems through the communication interface to visualize and validate production flows in the virtual environment.

The virtual layout is created using virtual components, which typically represents factory floor equipment. These components can be added from the pre-defined library provided by Visual Components (eCatalog) or created by the user from the beginning using the original CAD files. These components and their operation and interaction define the factory floor operations (see Figure below).

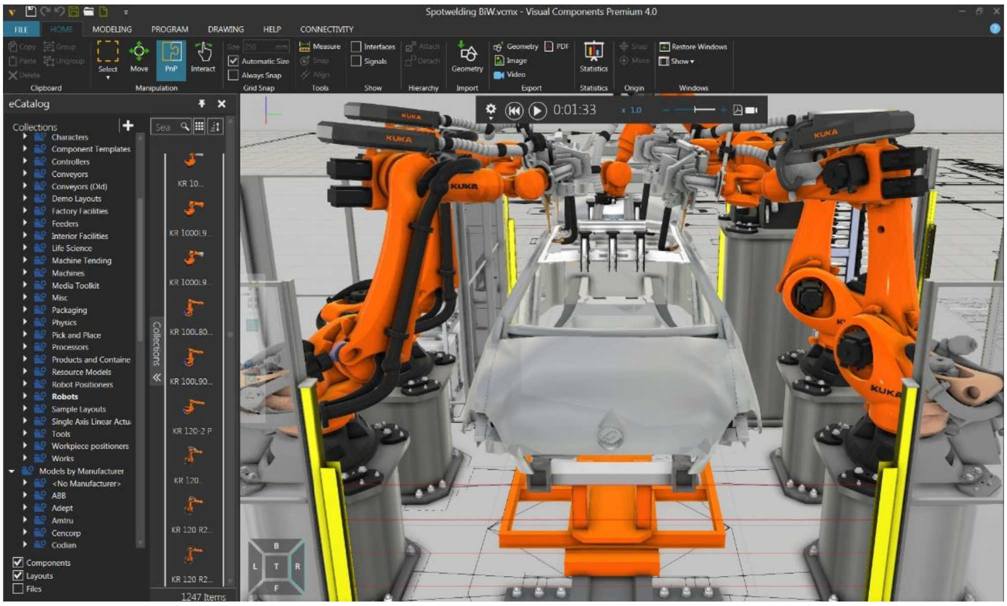

Figure 31 Screenshot of Visual Components 4.0 showing the GUI and the simulation of a robotic welding cell

The GUI of Visual Components 4.0 (Figure 1) provides access through the GUI to all the necessary tools to create the factory layout, model the components, program the robotic systems, connect with external controllers and generate engineering documentation.

Visual Components 4.0 is vendor independent, which allows to integrate in the same layout solutions from different vendors and arrange easily different configurations to obtain the best configuration and facilitate the deployment and commissioning of the systems in the real factory.

Visual Components 4.0 platform provides access to the different parameters of the simulated components, retrieving these parameters during the simulation allow to analyze and evaluate changes in the production during the simulation, which enables evaluating the different configurations, created in the virtual world and enhancing efficiency when configurations are transferred to the real lines.

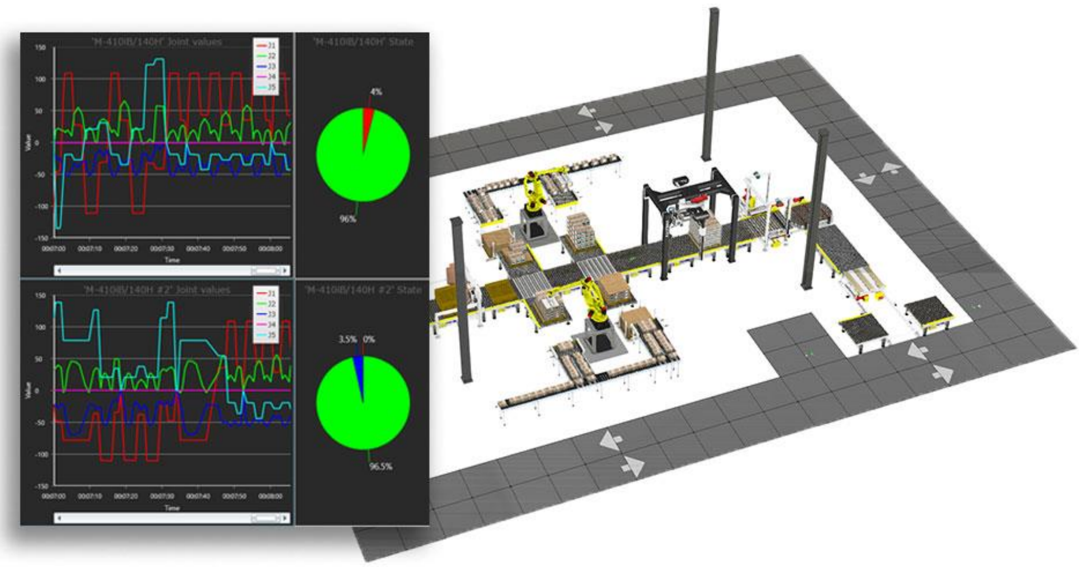

Figure 32 Screenshot of simulated factory layout and analytics obtained while virtual runtime

4.2.1.2APIs & Communication Interfaces

Visual Components 4.0 provides two different application interfaces and one communication interface to interact with the simulation layout:

- Python Interface is a resource interface within our 3D simulation platform, it allows configure components and processes inside the simulation.
- .Net interface provides all the operations available in the GUI as well as access to the layout model. It is designed primarily to allow client applications to create and manipulate components and layouts.
- Communication interface allows communicating with external automation components such as robot controlers (Staubli and UR) and automation systems using OPC UA.

4.2.1.3 Extensions / Integration

The Python and .Net interfaces open the possibility of creating tailored applications depending on the final user. Example of tailored applications created for Visual Components 4.0 can be found in the Visual Component's forum (http://forum.visualcomponents.com/forums/forum/vc40/net-addon-programming/).

BOOST pilots, particularly FILL and VWAE pilots, will bring new extensions to handle big data and communicate in next to real time cases. The solutions to be developed will consider the applicability to other verticals considered inside BOOST.

It is planned to be integrated the serialization of the big data generated in the simulation into data models to be handled by AI engines. This serialization will allow introducing automatic (or semi-automatic) reconfiguration of the simulation layouts to reach maximum performance. The data generated within the simulation will be merged with historical data to improve productivity.

4.3 Autonomous production automation big data analytics platform

4.3.1 SIEMENS MindSphere platform extensions for 3rd parties

MindSphere is the open, cloud-based IoT operating system from Siemens that lets you connect your machines and physical infrastructure to the digital world. It lets you harness big data from billions of intelligent devices, enabling you to uncover transformational insights across your entire business.

MindSphere from Siemens offers a scalable cloud Platform as a Service (PaaS) that's perfect for developing apps. Designed as an open OS for the Internet of Things, it lets you seamlessly connect with your machines so you can improve the efficiency of plants by harnessing the huge volumes of data that your assets generate.

4.3.1.1Overview

MindSphere is a distributed two-sided open platform to connect the IoT and leverage its combined power through services and Apps. The platform consists of three distinct layers, the MindSphere Application Platform which provides a managed Platform as a Service (PaaS) to host your applications directly on MindSphere, the MindSphere Services Platform which allows to use our services via public APIs in your own solutions and the MindConnect Elements which provide you with plug and play hardware and customizable software components to get your data into the platform.

MindSphere Application Platform

The MindSphere Application Platform is powered by Cloud Foundry and provides you with all of its PaaS capabilities and a deep integration into the MindSphere ecosystem.

Operational aspects like deployment, scaling, monitoring are getting easier since these features are baked into the core of the MindSphere Application Platform.

MindSphere Services Platform

The MindSphere Services Platform provides a variety of supporting services to make application development easier. Through our APIs you can automate almost every process, manage users and assets, retrieve and store Industrial IoT data, run analytics and much more.

MindConnect Elements

With MindConnect Elements, Siemens offers numerous possibilities for connecting machines, plants and worldwide fleets to MindSphere, regardless of the manufacturer.

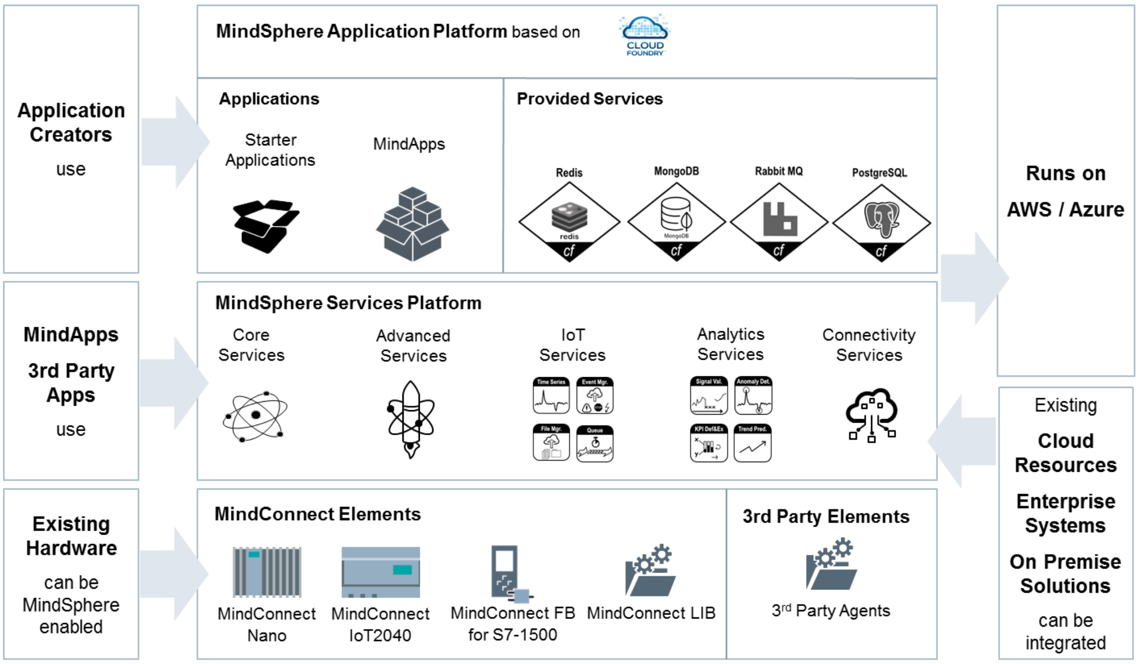

Figure 33 MindSphere overview

4.3.1.2 APIs & Services

In this section you will find all currently published services that can be used; the services are grouped in the following cluster:

- Platform Core
- IoT and Storage
- Connectivity
- Advanced Services
- Analytics Services

4.3.1.3 Platform Core

In this section shows the platform core MindSphere services, from more details please see APPENDIX A: MindSphere Platform Core at the end of the document.

Identity Management

Use Customer Management APIs to manage your customers (e.g. add new ones or change existing) for own tenants only.

OAuth Authorization Server

Use OAuth Authorization Server to request authentication and authorization.

Tenant Management

Use Tenant Management to manage: subtenants, tenant information and legal information.

Usage Transparency Service

Usage Transparency Service offers a UI giving insight on your resource consumption on the MindSphere platform. For Developers it as offers an API to track a metric defined by the developers. This metric can also be retrieved via an API of UTS.

4.3.1.3.1 IoT and Storage

In this section shows the IoT and storage MindSphere services, from more details please see APPENDIX B: IoT and Storage at the end of the document.

IoT File

File Service to read, write and delete files: upload, update and delete files associated to assets; store metadata information, and search for files by metadata.

IoT Time Series

Use Time Series to create, read, update and delete dynamic data. Since time series data are always related to an asset, the instance of an asset must have been created by you beforehand.

IoT TS Aggregates

Use Time SeriesAggregates Service to read aggregated time series values. Retrieve the following aggregated values per interval: Count, Sum, Average, Minimum, Maximum, First Value, and Last Value.

4.3.1.3.2 Connectivity

In this section shows the connectivity MindSphere services, from more details please see APPENDIX C: Connectivity at the end of the chapter.

Agent Management

Use Agent Management to create, edit or remove MindConnect elements, on-board and off board agents and set relations to assets.

MindConnect API

Use MindConnect API to develop custom agents and connect your device or application to MindSphere.

4.3.1.3.3 Advanced Services

In this section shows the advanced MindSphere services, from more details please see APPENDIX D: Advanced Services at the end of the document.

Asset Management

Represent physical assets from your site in MindSphere. Use models and create instances, set relations to others and create structures such as hierarchies.

Event Management

Manage standardized and customized events. Acquire events from the field & other applications.

Data Flow Engine

Provides workflow orchestration. Create custom rules based on incoming data, apply them on the fly & define actions based on this information.

Notification Service

Use APIs or graphical user interface to send information to your users & customers via e-mail, SMS or push/scheduled notification.

4.3.1.3.4 Analytics Services

In this section shows the advanced MindSphere services, from more details please see the APPENDIX E: Operations of Signal Calculation Service at the end of the document

Anomaly Detection

The Anomaly Detection Service aims to automatically detect unexpected behavior of processes and assets using time series data.

Event Analytics

The Event Analytics Service provides the essential functionality for a data-driven analysis of event data. It enables the user to get a better grasp of what is happening inside the system through statistical analysis.

KPI Calculation

The KPI Calculation Service computes Key Performance Indicators (KPIs) for an asset. It uses data source such as sensors, control events and calendar entries.

Signal Calculation

The Signal Calculation Service processes sensor time series data of an entity. The service aggregates, modifies, smoothens and transforms the original sensor data for further analysis or storage along with the original data.

Signal Validation

The Signal Validation Service validates sensor time series data of an entity. The API offers a set of common operations for signal validation.

Trend Prediction

The Trend Prediction Service predicts future values for time series using linear and non-linear regression models. It is a forecasting framework that has many useful applications in the area of Process & Condition Monitoring.

4.4 Collaborative manufacturing network big data analytics platform

4.4.1 TRIMEK M3 Platform

4.4.1.1 Overview

As explained in Chapter 3.7, the predictive quality control framework is going to be developed based on the M3 platform which is poised to provide a structured solution for Metrology4.0, an edge-powered quality control analytics, monitoring and simulation system. This solution is used for the organization, analysis and reporting operations of the metrological information, taking advantage of the storage and computational capabilities of the cloud to carry out advanced operations and provide smart added value services. Figure 35 depicts the M3 global architecture.

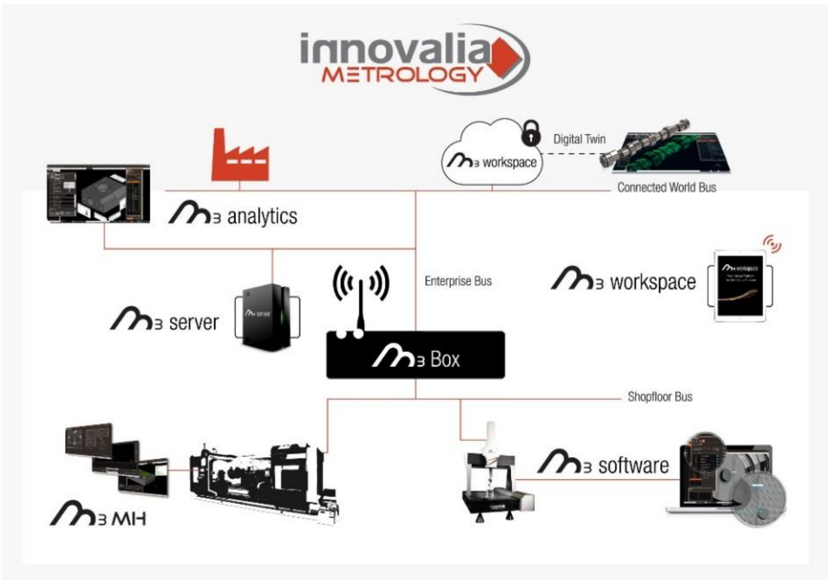

Figure 34 M3 architecture

Within the range of the proposed framework, M3 Software, M3 Workspace and M3 Analytics will be adapted, extended and implemented to serve the purpose of the pilot.

M3 Software

The M3 software is a high performance software for capturing and analysing point clouds. This module allows to scan and to capture point clouds of the real pieces, in a versatile, agile and powerful way. The M3 software in combination with the 3D optical scanners can be used to develop precise and highly accurate point cloud images that can then be converted to different 3D design and modelling software. These scanned images can then be cross referenced with the original designs or with other scanned objects allowing for quick and accurate comparison and discovery of deformation or other dimensional discrepancies. The M3 software covers the entire spectrum of metrology, regardless of device, brand or model. It works locally but is powered up by the use of the edge-powered technologies included in the global solution.

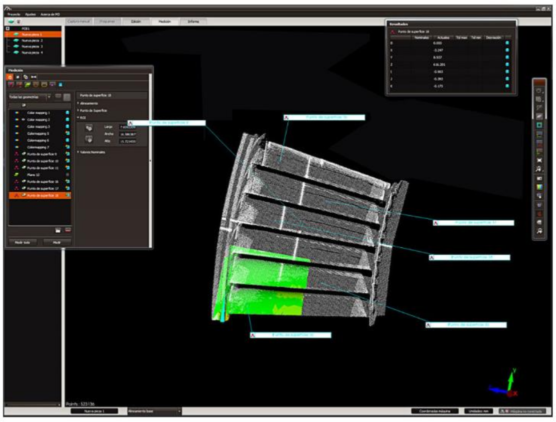

Figure 35 M3 software

M3 Analytics

The M3 analytics is a powerful tool that enables the visualization, the statistics analysis and the reporting operations related to all the data stored in the cloud by means of several algorithms and computational components. As this tool makes use of the memory and computational resources available in the cloud, it is possible to use it anytime and anywhere and by means of a simple computer or tablet with low computational capacities. Its main features are:

- Create control dashboards
- Combine production data with measurement data
- Automation of reports
- Histograms
- Cp and Cpk Statistics
- Custom Filters

M3 Workspace

M3 Workspace is a cloud based metrology software that synchronizes with the main M3 Metrology software which allows for the automated uploading of any metrology results straight to the cloud. M3 workspace is web based and allows users with access to visualize the results (Point clouds, CAD models, Color mapping, Reports, etc.) using any smart device. In addition, M3 Workspace acts as a sort of repository where all the results that come from the measurements are located and can then be easily downloaded for further analysis. It permits the massive management of digital parts and point clouds, storing and sharing the metrological information.

4.4.1.2 Data Sovereignty

As aforementioned, TRIMEK will use its M3 Platform to carry out one of the business scenarios of Gestamp pilot, the development of a predictive quality control framework. The objective of this business scenario is to optimise the speed of data acquisition, visualization and processing for massive metrological data and to design and create a collaborative predictive analytics platform to make decision making data savvy. Therefore, in the end, the overall efficiency of production process will be increased.

In this context, two different Data Owners can be defined:

- TRIMEK: provides, processes and analyses metrological data from different sources.
- Gestamp: provides production data and quality data related to different products.

In this case, data consumer will be the predictive quality control solution developed by TRIMEK based on its M3 Platform.

The use of the IDS connector at the factory will bring the necessary warranty to TRIMEK that only the measurements and processed data which have been approved can be delivered to the predictive quality control solution. Likewise, Gestamp will be sure that only the relevant product and process data is going to be delivered and published.

TRIMEK is deploying and configuring a M3 FIWARE System Adapter in the IDS Connector. Thus, the metrological data coming from the Coordinate Measuring Machines (CMMs) is sent to the M3 Platform to analyse the performance of the CMM and ensure zero defects. Likewise, the IDS-ready connector could also send product quality measurements and process information to the predictive quality control solution, enabling advanced analysis for zero defects and zero breakdowns.

In short, only TRIMEK and Gestamp will be able to exchange data and be certified to request specific data under data owner's usage policy. The terms and conditions of the Data sovereignty will be detailed throughout the development of the project.

4.4.1.3 Extensions/Integration

Regarding the Big Data M3 platform, TRIMEK will extend the M3 Software, M3 Workspace and M3 Analytics in order to adapt them to fulfil the objectives of the Gestamp pilot. In spirit of this, proper and advanced algorithms and functionalities will be implemented to cover:

- Colour mapping with textures for massive point clouds.
- Implementation of QIF standard, covering PMI.
- Harmonization of different data formats
- Collaborative cloud environment for data from different sources and locations
- Connection between QIF and IDS to ensure interoperability
- Correlation of process and product data for predictive purposes
- Knowledge extraction
- Develop trends and execute features and/or parameters comparison between process, products and similar parts across different plants, etc.

Moreover, M3 platform needs to be integrated with IDS. TRIMEK will developed and deployed the corresponding FIWARE System Adapter in the IDS Connector for metrological data sharing.

4.4.2 Industrial IoT Real Time Location System And Network Monitoring

4.4.2.1 Overview

As explained in Section 3, i2track: Predictive logistics platform:

- Real time assets tracking and routes optimization based on assets flow analysis
- Predictive logistics tool development to optimize logistic process prior to a line's stop or to improve operations.

The logistics optimization process will affect the following scenarios:

- Raw material traceability
- Dies position and status control
- Containers flow control

The first step in the future business scenario starts with monitoring the logistic flow of containers around the factory. With the objective of removing paper labels from the containers, a sophisticated solution based on RFID needs to be put in place. This process can be separated in the next steps:

- It is needed to select he container that will be monitored, taking into account that some stay in-house while others can spend a considerable amount of time of premises (in the OEM client).
- Determine the key information to store
- Determine the number of containers to locate
- Install tracking devices in mobile assets (forklifts, tow tractors…)
- Integration with ERP / MES: Definition of the data to be written in tags and the interface with line operators, drivers, etc.
- Determine positioning of stacked containers (vertical axis)
- Determine the best implementation for inbound/outbound containers, regarding the risk of misuse by other actors.

The second step in the future business scenario is the tracking and monitoring of metal sheet cutting dies. With this objective in mind we separate the process in the next steps:

- Design of the optimal tag + reader solution
- Determine the height of the bridge crane in respect to the cutting die
- Determine the optimal tag location that can resist the piling up of the dies.

Finally the third step in the future business scenario is to complete the traceability of the manufactured parts up to the raw material. That is to associate each sheet metal rolls to the parts produced form them. We identify the following steps:

- Identify the received rolls. To do so it is important to determine which actor will attach the ID tag to the metal roll, the metal provider or Gestamp
- Design specific RFID solution for coils integrated with the supplier and plant information systems

See below a diagram regarding the RTLS solution performance and IT integration.

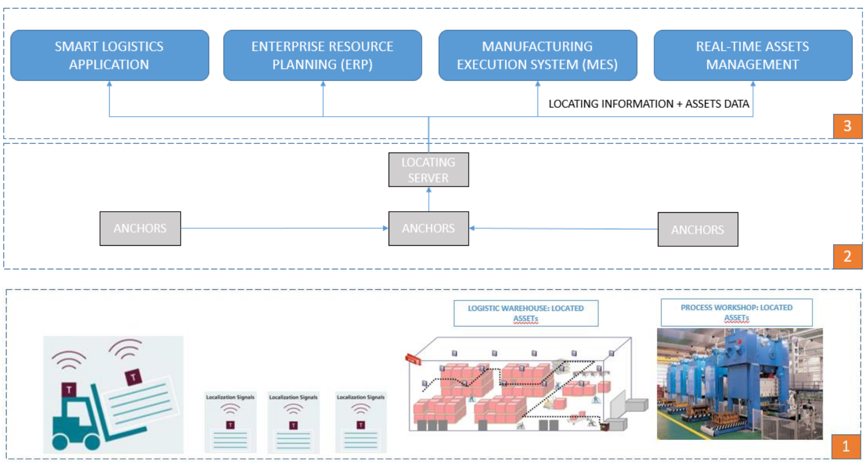

Figure 36 RTLS Solution and IT Integration

4.5 Full equipment & product availability big data analytics platforms

4.5.1 SAS analytics

4.5.1.1Overview

SAS Analytics provides a set of different tools and functionalities that can be useful to achieve the needs of the Whirlpool Use Case. In the scope of the project, SAS will implement the functionalities described in WP8 with the following Tools:

- SAS Viya
- SAS Visual Forecasting
- SAS Visual Data Mining and Machine Learning
- SAS Event Stream Process.

4.5.1.1.1 SAS Viya

SAS Viya is a cloud-enabled, in-memory analytics engine that delivers everything you need for quick, accurate and consistent results. Elastic, scalable and fault-tolerant processing addresses complex analytical challenges and effortlessly scales to meet future needs.

SAS Viya provides:

- Faster processing for analytics
- A standardized code base that supports programming in SAS and other languages, like Python, R, Java and Lua
- Support for cloud, on-site or hybrid environments. It deploys seamlessly to any infrastructure or application ecosystem.

SAS Viya Capabilities

Cloud-enabled, elastic and scalable

SAS Viya is built to be elastic and scalable for both private and public clouds. Complex analytical, in-memory calculations are optimized for unconstrained environments and automatically adjust in constrained environments. The elastic processing effortlessly adapts to burst processing using available resources – scaling computing capacity as needed. This elasticity allows quickly experiment with different scenarios and apply more sophisticated approaches to increasing amounts and speeds of incoming data.

Open analytics coding environment

Data scientists and statisticians are provided with breadth of analytics capabilities that are easily available from the coding language of their choice. Whether it's SAS, Python, Java, R or Lua, analytical professionals can access the power of SAS for data manipulation, interactive data interrogations and advanced analytics. SAS Viya also includes public REST APIs to all underlying functionality, so software developers can add proven SAS Analytics to applications. And all analytical assets are united within a common environment to provide a single, managed inventory across your organization.

Fast, distributed in-memory processing

SAS Viya provides highly available, distributed processing crafted to handle multiple users and complex analytical workloads. Computing operations are automatically distributed across the cores of a single server or the nodes of a massive compute cluster. If a node fails, fault-tolerant processing ensures blazingly fast speed is retained.

All data, tables and objects are held in memory as long as required and for whomever is using them. Independent sessions ensure optimized processing for everyone. The SAS Viya engine provides a secure, scalable multiuser environment for concurrent access. Users can collaborate to simultaneously explore the same data, probe findings and identify analytically sound actions.

In addition, code written in the distributed environment is portable. It can be defined once, run anywhere and scaled to solve any size data problem. Code built on SAS Viya runs in stream, in database, in memory, in Hadoop, in public and private clouds, and even in device.

Resilient architecture with guaranteed failover

All analytical computations should run without interruption. The fault-tolerant design of SAS Viya automatically detects server failure, even across clusters. Processing is optimized and redistributed as needed. SAS Viya also manages several copies of data on the computing cluster. If a node in the cluster becomes unavailable or fails, the required data is retrieved from another block. These self-healing mechanisms ensure high availability for uninterrupted processing and automated recovery.

4.5.1.1.2 Visual Forecasting

SAS Visual Forecasting provides a resilient, distributed and optimized generic time series analysis scripting environment for cloud computing. This solution includes automatic forecast model generation, automatic variable and event selection, automatic parameter optimization, automatic model selection and automatic forecast generation. It also provides advanced support for time series analysis (time domain and frequency domain), time series decomposition, time series modelling, signal analysis and anomaly detection (for IoT). With SAS Visual Forecasting, you pick up the data once and run everything you need, taking advantage of in-memory, large-scale distributed processing. A scripting language optimizes and compiles your forecast based on where it is running.

SAS Visual Forecasting automatically analyses large numbers of time series, so forecasters don't have to diagnose each series. The software determines the forecasting models that are most suitable for the historic data. When doing hierarchical forecasting, holdout samples can be specified so that forecasting models are selected not only by how well they fit past data, but how well they are likely to predict the future. An appropriate model is generated for each entity being forecast, based on user-defined criteria. Model parameters are automatically optimized. Any number of business drivers and events can be supplied and will be considered for inclusion in the models.

SAS Visual Forecasting provides a resilient, distributed and optimized generic time series analysis scripting environment. It supports fast, in-memory time series analysis. By nature, distributed systems break up large files and process each piece separately. This is problematic for time series analysis where the ordering of data is crucial. Time series analysis typically requires that time series data is stored contiguously in memory and in sorted order.

SAS Visual Forecasting shuffles the data so that each time series (or BY group) is copied into the memory of a single computing node. Each time series is executed on one thread of a node, and each node executes the compiled script for each of its assigned series. This makes time series analysis and forecasting possible on an enormous scale. And the scripting language is optimized and compiled for the machine it is running on, so users don't have to rewrite code for different machines.

Highly flexible forecast override

SAS Visual Forecasting adds a powerful, new capability that enables manual overrides to be made to a specific filter or group of time series defined by attributes, not just by hierarchical variables. For example, an analyst in the apparel Key Features industry may want to adjust a forecast for all products of a certain colour that is expected to be popular. Colour is not typically a level of the forecasting hierarchy. With the override capability, a custom filter can be defined of products meeting the colour attribute. Without this feature, if you wanted to apply an override to all products of a certain colour, you would have to manually enter the override to each product. Another example is sentiment, determined by text analysis of online reviews or user surveys. An analyst may want to increase (or decrease) forecasts for all products that have favourable (or unfavourable) sentiment. Creating filters saves a lot of time and manual effort when overriding non-hierarchical variables.

API support for working with open source:

While SAS Visual Forecasting has a broad range of forecasting models built-in, users can create their own customized models that perform well with their data. Also, with REST APIs, other applications can call SAS forecasting models.

Hierarchical reconciliation

Each series in the hierarchy is modelled and forecasted individually. Forecasts are then reconciled at multiple levels of the hierarchy in a top-down fashion. Users can adjust a forecast at any level and apportion it to lower levels, so the hierarchy maintains consistency, and individual forecasts (by products, locations, etc.) roll up to the top number. Without reconciliation, lower-level forecasts won't add up to the top-level forecast.

4.5.1.1.3 Visual Data Mining and Machine Learning

SAS Visual Data Mining and Machine Learning offers an exciting, new end-to-end visual environment that covers all aspects of machine learning and deep learning – from data access and data wrangling to sophisticated model building and deployment. In-memory, distributed processing handles large data and complex modelling, providing faster answers and efficient use of resources.

Flexible and approachable visual environment for analytics

Multiple users can currently analyse any amount of structured and unstructured data with the easy-to-use visual interface. Each project (goal) is defined by visual pipelines that break the analytics life cycle into a series of steps presented in a logical sequence. Branching can execute asynchronously. The visual interface (Model Studio) provides an integrated environment for the most common machine-learning steps: data prep, feature engineering, exploration, model building and deployment. Interactive tasks provide an easy way to apply sophisticated algorithms to large and complex data. These interactions also generate SAS code that can be save for later automation of tasks. In addition, code snippets and best practice templates are easily shared. Model Studio provides a highly collaborative environment for building, expanding and sharing models.

Highly scalable, in-memory analytical processing

This solution provides a secure, multiuser environment for concurrent access to data in memory. Data and analytical workloads operations are distributed across nodes, in parallel, and are multithreaded on each node for very fast speed. All data, tables and objects are held in memory as long as required, allowing for efficient processing. With built-in fault tolerance and memory management, advanced workflows can be applied to data, ensuring that processes always finish. You get dramatically reduced runtimes for large data and analytical processing, reduced network traffic and can take full advantage of modern, multicore architectures to find solutions much faster. The visual pipeline approach provides a collaborative, efficient environment for creating and deploying sophisticated machine-learning and deep learning models

Innovative statistical, data mining and machine-learning techniques

SAS Visual Data Mining and Machine Learning delivers an incredibly broad set of modern statistical, machine learning, deep learning and text analytics algorithms within a single environment. Analytical capabilities include clustering, different flavours of regression, random forests, gradient boosting models, support vector machines, natural language processing, topic detection and more. These powerful methods drive the identification of new patterns, trends and relationships between data attributes in structured and unstructured data. The solution also provides matrix factorization for building customized recommendation systems. With its ability to process high velocity and high-volume data sets, SAS Visual Data Mining and Machine Learning is uniquely suited for deep learning techniques. Deep learning algorithms include deep neural networks, convolution neural networks for image classification and recurrent neural networks for improved text analysis. Complex learning algorithms, such as neural networks, gradient boosting and random forest, can be automatically tuned for optimal performance, saving time and resources.

Integrated data preparation, exploration and feature engineering

To overcome usually time-consuming analytical data preparation activities, the drag-and-drop interface enables data engineers to quickly build and run transformations, augment data and join data within the integrated visual pipeline of activities. All actions are performed in memory to maintain a consistent data structure. Discover data issues and fix them with advanced analytical techniques. Quickly identify potential predictors, reduce the dimensions of large data sets and easily create new features from your original data.

Integrated text analytics

Designed with big data in mind, you can examine extremely large collections of text documents. Explore all of your textual data, not just a subset, to gain new insights about unknown themes and connections. Combining structured data with text data uncovers previously undetected relationships and adds even more predictive power to analytical models.

Model assessment and scoring

Test different modelling approaches in a single run and compare results of multiple supervised learning algorithms with standardized tests to quickly identify champion models. Then, operationalize analytics in distributed and traditional environments with automatically generated SAS score code.

Accessible and cloud-ready

Whether it's Python, R, Java or Lua, modelers and data scientists can access SAS capabilities from their preferred coding environment. And with SAS Viya REST APIs, you can add the power of SAS to other applications. You can also deploy SAS Visual Data Mining and Machine Learning where it makes the most sense for your organization: on-site, in a private cloud via technologies such as Cloud Foundry or in public clouds, including Amazon Web Services and Microsoft Azure. You can also access this software via the pre-deployed and preconfigured managed software-as-a-service offerings provided by SAS.

5.1.1.1 SAS Event Stream Process

SAS Event Stream Processing ingests large volumes of streaming data quickly – millions of events per second – so companies can understand events in the data while it's in motion. No data stream is too big or fast. It is possible to integrate, visualize, transform and analyse IoT data across the entire ecosystem – edge devices, data centres or the cloud. The solution's processing speed is bounded only by the hardware environment's limitations.

Incoming data is read through adapters and connectors, which are part of a publish-and-subscribe architecture. Event data publishes into a source window of an event stream processor. A visual interface makes it easy to define the windows, procedures and operators. In turn, it's simple to define continuous queries through which the data will stream. Streaming data is examined for patterns and can be intelligently filtered to store anomalies that demand deeper investigation. Or, if no relevancy is detected, the data can be discarded. Downstream applications subscribe to receive streaming analysis results with prescribed actions. This approach allows you to respond quickly to changing conditions and position your business for new IoT market opportunities.

For detailed information about SAS tools refer to:

- Visual Forecasting: https://www.sas.com/en_us/software/visual-forecasting.html
- Visual Data Mining and Machine Learning: https://www.sas.com/en_us/software/visual-data-mining-machine-learning.html
- Visual Analytics: https://www.sas.com/en_us/software/visual-analytics.html
- Event Stream Process: https://www.sas.com/content/dam/SAS/en_us/doc/factsheet/sas-event-stream-processing-106151.pdf.

4.5.1.2 *Data sovereignty*

In WP8 Architecture, data on SP Data Model will be accessible only through SAS applications. More specifically, data will be loaded by data apps into memory, in order to be processed by SAS in-memory engine. In this way, data sovereignty will be implemented through SAS Authorization system.

Authorization is the aspect of security that determines which resources are available to which users. The SAS Viya authorization layer consists of two authorization systems:

- Cloud Analytic Services (CAS) authorization system
- General authorization system

Each system uses a distinct model to protect a distinct class of resources.

- **CAS Authorization system** is responsible to define authorization templates on data that is loaded into cas engine.
- **General authorization system** is used to define authorization templates on user objects such as folders, reports, models.

CAS authorization manages access to the following CAS objects:

- caslibs
- CAS tables and columns
- CAS action sets and actions

CAS authorization requirements do not apply in the following circumstances:

- The requesting user has assumed a role that is exempt from all applicable authorization requirements. For example, the user has assumed the Superuser role and the request is to add a caslib.
- The target object is not potentially sharable. For example, the target is a table in a personal caslib, a session caslib, or the session scope of a global caslib.

In the CAS authorization system, memberships, inheritance, and row-level filters can influence access.

Key features for CAS Authorization are:

Feature	Definition
Access control	A composite of authorization elements. Example: An access control grants ReadInfo to groupA on caslibA.
Target	A resource. Examples: tableA, caslibA

Principal	The user, group, or construct to which an access control is assigned. Examples: UserA, GroupA, Authenticated Users
Permission	A type of access. Values: ReadInfo, Select, LimitedPromote, Promote, CreateTable, DropTable, DeleteSource, Insert, Update, Delete, AlterTable, AlterCaslib, ManageAccess
Setting	An indication of whether (and to what extent) access is provided. Values: Grant, Row-Level Grant, Deny
Filter	In a row-level grant of the Select permission, the constraint expression. Example: User='SUB::SAS.Userid', sales>1000
Effective access	A context-neutral description of the net result of all relevant access controls. Values: Authorized, Not Authorized, Row-Level
Access outcome	In an access request, the authorization decision. Values: Authorized, Not Authorized, Row-Level Authorization

In the general authorization system, information about the requesting user, the target resource, and the environment can influence access. Each access request has a context that includes environmental data such as time and device type. Environmental constraints can be incorporated using conditions.

4.5.1.3 Extensions/Integration

SAS is an open analytics platform which supports, embrace and extend 3rd party engines such as Python, Java, Lua, R and REST APIs.

An example can be viewed using this link: https://developer.sas.com/apis/cas/actions.html

There are situations where users might want to include 3rd party engines due to the following reason for example: reading a novel method in an academic publication and keen to prototype it using a 3rd party engine.

New prototype method can be compared with SAS in-memory algorithms. In addition, SAS have a strong in-memory data processing capabilities which means data scientist can take advantage of this when working with 3rd party solutions. We offer that ability to extend out to the edge. SAS does this on an even deeper level, allowing you to leverage the embedded process technology, so rather than lift off sources, we are able to take the compute capacity and score at the source. This also allows us to do model retraining – key to machine learning at the point of generation. To improve models over time as conditions/data change, unlike other technologies.

We also provide the ability to import specific deep learning models from 3rd party systems like Caffe and others.

4.5.2 CERTH IoT Platform

4.5.2.1 Overview

CERTH IoT Platform is a cloud-based platform developed by CERTH/ITI. The platform:

- Enables Big Data Storage
- Supports real-time monitoring services
- Supports multiple continuous connected IoT devices
- Offers Real-time Predictive Analytics services

It is originally designed to support a smart home equipped with plenty IoT devices and analytic tools for the data coming continuously from these devices. The platform's core is written in Java. The interfaces of the IoT platform are developed using AngularJS.

The platform uses two different databases for storage. A document based NoSQL database, MongoDB for storing 'static' data (devices, IDs, users etc.) and an InfluxDB for 'dynamic' data storage (measurements, locations, timestamps etc.)

The CERTH IoT platform offers different types of connectivity. Both HTTP/REST and MQTT interfaces are supported by the platform. Moreover, web sockets are used for live reloading (e.g in maps visualization)

Finally, for the security part, HTTPs is used alongside with Basic authentication.

4.5.2.2 Extensions/Integration

In the context of the VOLVO pilot dedicated to assembly line visualization and forecasting over the supply chain, we plan to:

- extend IoT Cloud platform in order to fulfil business scenarios presented by the pilot and cover the needs of data storage, maps visualization and predictive algorithms
- make the platform compatible with EIDS and IBM Hyperledger Fabric

4.5.3 ATLANTIS DSS for smart maintenance

4.5.3.1 Overview

Traditionally, the core of a Decision Support System (DSS) is a rule-based mechanism, which handles manually defined actions based on pre-specified static threshold violations. Such tools are used to assess the performance of the machines, to diagnose failures and overall to improve the maintainability and the operational efficiency of the production line.

In the context of the BOOST4.0 project we plan to extend the traditional notion of the DSS, by leveraging data mining processes for Big Data analysis, IoT and machine learning techniques, in order to provide a DSS tool that is compatible with the key Industry4.0 element called smart maintenance. The goal is to provide a tool that is able to diagnose, handle (to a specific extend autonomously) and potentially predict future unwanted incidents (like machine failures) to ensure that the required equipment is fully functional at all times. The basis will be an existing ATLANTIS tool that is currently implemented as finite state machine.

The vision is to combine the input from maintenance experts, fault detection and potentially prediction algorithms using an extendable fusion mechanism inside the DSS tool to provide easily interpretable and useful results. Maintenance expert's knowledge, plays a pivotal role in the initialization of the system. Based on their experience, the initial set of rules will be created to instantiate the rule-based mechanism. The provided rules can be refined (the provided thresholds of the rules) based on prior knowledge (e.g. users' feedback). Horizontally to the rule-based mechanism, outlier detection and motif detection algorithms can be used for fault detection at real time and with the appropriate pre-processing for input to a predictive maintenance approach to enhance the flexibility of the DSS tool.

4.5.3.2 Extensions/Integration

To follow the project's demands a version of the tool will be compatible with the IDS specification, implementing the needed IDS connectors, incorporating the specified vocabularies and providing the appropriate communication and security proto

5 Conclusions and Future Work

In conclusion, this deliverable describes the effort spent from M4 to M9 for Task 3.5 - Big Data Models for Cognitive Manufacturing and and from M7 to M9 for Task 3.6 - Big Data Analytics Platform and presents the first results of these tasks. The complete work and the outcome of the two task will be presented in D3.4 on M24.

In the first steps of Task 3.5, the research was focused on a thorouth analysis of the state-of-the-art in cognitive manufacturing based on Big Data, and the connection of available methodologies with the BOOST 4.0 pilots as the project and Task 3.5 aim to contribute to AI and Smart Manufacturing domains. The planned outcome will be cognitive factories that will be flexible, adaptive and reliable in the momentary situations and unforeseen conditions, in order to derive an efficient production scheme and aid the decision makers. Furthermore, these smart factories will be able to reorginize production and assembly planning based on early detection of machine failures or on prediction of delays in the supply-chain. Based on chapter 4 of this document, Tasks 3.5 and 3.6 will contribute to the BOOST 4.0 Online Collaborative Analytics Service Marketplace, which is based on IDS architecture, services and apps related to:

- Model-based Fault diagnosis methods for the detection of deterioration rate of production machines and their root cause
- Predictive modeling and early detection algorithms for the detection of defect products (machine learning, data mining, information theory, statistics, complex event processing, trend analysis)
- Advanced data visualization and visual analytics. New ways of interaction and display technologies to support analytical reasoning and collaborative decision making

Furthermore, the main Big Data Platforms that will be used for the pilot cases' implementation are presented in this document as well. All platforms' extensions in order to cover the pilot needs and the project's architecture will be described in the next version of the document at M24.

6 References

[1] Z. Duric et al., "Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction," in Proceedings of the IEEE, vol. 90, no. 7, pp. 1272-1289, Jul 2002.

[2] R. Anderson, John & Matessa, Michael & Lebiere, Christian. (1997). ACT-R: A Theory of Higher Level Cognition and Its Relation to Visual Attention. Human-Computer Interaction. 12. 439-462. 10.1207/s15327051hci1204_5.

[3] F. Zaeh, M & Beetz, Michael & Shea, K & Reinhart, G & Bender, K & Lau, C & Ostgathe, M & Vogl, W & Wiesbeck, M & Engelhard, M & Ertelt, C & Rühr, Thomas & Friedrich, M & Herle, S. (2009). The Cognitive Factory. Changeable and Reconfigurable Manufacturing Systems. 355-371. 10.1007/978-1-84882-067-8_20.

[4] Cognitive manufacturing & Industry 4.0. https://www.ibm.com/blogs/internet-of-things/manufacturing-industry-4-0/

[5] Alfonso Sánchez Boza, Rodolfo Haber Guerra, Agustín Gajate, Artificial cognitive control system based on the shared circuits model of socio-cognitive capacities. A first approach, Engineering Applications of Artificial Intelligence, Volume 24, Issue 2, 2011, Pages 209-219, ISSN 0952-1976.

[6] D. Bruckner, H. Zeilinger and D. Dietrich, "Cognitive Automation—Survey of Novel Artificial General Intelligence Methods for the Automation of Human Technical Environments," in IEEE Transactions on Industrial Informatics, vol. 8, no. 2, pp. 206-215, May 2012.

[7] M. Agyemang, K. Barker, R. Alhajj, "A comprehensive survey of numeric and symbolic outlier mining techniques", Intelligent Data Analysis 10 (6) (2006) 521-538.

[8] V. Chandola, A. Banerjee, V. Kumar, "Anomaly detection: A survey", ACM Computing Surveys, 41

[9] V. Hodge, J. Austin, "A survey of outlier detection methodologies", Artificial Intelligence Review 22 (2) (2004) 85-126.

[10] M. Markou, S. Singh, "Novelty detection: a review-part 1: statistical approaches", Signal Processing 83 (12) (2003a), 2481-2497.

[11] M. Markou, S. Singh, "Novelty detection: a review-part 2: neural network-based approaches", Signal Processing 83 (12) (2003b) 2499-2521.

[12] A. Patcha, J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends", Computer Networks 51 (12) (2007) 3448-3470.

[13] C. O. Reilly, A. Gluhak, M. A. Imran, S. Rajasegarar, "Anomaly detection in Wireless Sensor Networks in a non-stationary environment", IEEE Communications surveys & Tutorials 16 (3) (2014) 1413-1432

[14] D. Savage, X. Zhang, P. Chou, Q. Wang, "Anomaly detection in online social networks", Social Networks 39 (2014) 62-70.

[15] Y. Zheng, S. Rajasegarar, C. Leckie, M. Palaniswami, "Smart car parking: temporal clustering and anomaly detection in urban car parking", 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), IEEE, Singapore (2014) 1–6.

[16] Lars Baumgärtner, Christian Strack, Bastian Hoßbach, Marc Seidemann, Bernhard Seeger, and Bernd Freisleben. 2015. Complex event processing for reactive security monitoring in virtualized computer systems. In Proceedings of the 9th ACM International Conference on Distributed Event-Based Systems (DEBS '15). ACM, New York, NY, USA, 22-33.

[17] T. Vafeiadis, E. Bora-Senta, D. Kugiumtzis, "Estimation of linear trend onset in time series", Simulation Modelling Practice and Theory, vol. 19, 2011, pp. 1384–1398.

[18] T. Vafeiadis, S. Krinidis, C. Ziogou, D. Ioannidis, S. Voutetakis, D. Tzovaras, "Robust malfunction diagnosis in process industry time series", Proceedings of IEEE 14th Conference of Industrial Informatics 2016;111-116.

[19] WU, E, DIAO, Y., AND RIZVI, S. 2006. High-performance complex event processing over streams. In Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD'06). ACM, New York, NY, 407–418.

[20] Suhothayan, Sriskandarajah & Gajasinghe, Kasun & Loku Narangoda, Isuru & Chaturanga, Subash & Perera, Srinath & Nanayakkara, Vishaka. (2011). Siddhi: A Second Look at Complex Event Processing Architectures. GCE'11 - Proceedings of the 2011 ACM Workshop on Gateway Computing Environments, Co-located with SC'11. 10.1145/2110486.2110493.

[21] H. Alkhadafea, A. Al-Habaibehb, A. Lotfia, Condition monitoring of helical gears using automated selection of features and sensors, Volume 93, November 2016, Pages 164–177

[22] F. Schnizler, T. Liebig, S. Mannor, G. Souto, S. Bothe, H. Stange, Heterogeneous Stream Processing for Disaster Detection and Alarming, 2014 IEEE International Conference on Big Data, pp. 914-923.

[23] A. Motamedi, A. Hammad, Y. Asen, Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management, Automation in Construction vol. 43, 2014, pp. 73–83.

[24] Yao, J.T., Y.L. Li and C.L. Tan, \Forecasting the Exchange Rates of CHF vs USD Using Neural networks", Journal of Computational Intelligence in Finance, Vol.5, No.2., pp7-13, 1997.

[25] Y. Freund, R.E. Schapire, A desicion-theoretic generalization of on-line learning and an application to boosting, in: Computational Learning Theory, Springer, 1995, pp. 23–37.

[26] https://homes.cs.washington.edu/~marcotcr/blog/lime/

[27] 1. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin, "why should i trust you?": Explaining the predictions of any classifier, Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA), KDD '16, ACM, 2016, pp. 1135–1144.

[28] Wei Koh, Pang & Liang, Percy. (2017). Understanding Black-box Predictions via Influence Functions.

[29] D. E. Goldberg, "Genetic algorithms in search, optimization and machine learning, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA ©1989.

[30] LeCun, Yann. "LeNet-5, convolutional neural networks". Retrieved 16 November 2013.

[31] Z. Gao, C. Cecati and S. X. Ding, "A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis. With Model-Based and Signal-Based Approaches," in IEEE Transactions on Industrial Electronics, vol. 62, no. 6, pp. 3757-3767, June 2015.

[32] Solé-Simó, Marc & Muntés-Mulero, Victor & Rana, Annie & Estrada, Giovani. (2017). Survey on Models and Techniques for Root-Cause Analysis. arXiv:1701.08546

[33] R. Beard, "Failure accommodation in linear system through self-reorganization," Ph.D. dissertation, MIT, Cambridge, MA, USA, 1971.

[34] J. Furnkranz, "Separate-and-conquer rule learning," Artif. Intell. Rev., vol. 13, no. 1, pp. 3–54, Feb. 1999.

[35] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," SIGMOD Rec., vol. 22, no. 2, pp. 207–216, Jun. 1993.

[36] I. Aydin, M. Karakose, and E. Akin, "Generation of classification rules using artificial immune system for fault diagnosis," in Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, Oct 2010, pp. 343–349.

[37] S. Muggleton and L. de Raedt, "Inductive logic programming: Theory and methods," The Journal of Logic Programming, vol. 19, pp. 629 – 679, 1994.

[38] S. Augier, G. Venturini, and I. Y. Kodratoff, "Learning first order logic rules with a genetic algorithm," in Proc. of the First Int. Conf. on Knowledge Discovery & Data Mining, 1995, pp. 21–26.

[39] P. Liggesmeyer and M. Rothfelder, "Improving system reliability with automatic fault tree generation," in Fault-Tolerant Computing, 1998. Digest of Papers. Twenty-Eighth Annual International Symposium on, June 1998, pp. 90–99.

[40] A. Majdara and T. Wakabayashi, "A new approach for computer-aided fault tree generation," in Systems Conference, 2009 3rd Annual IEEE, March 2009, pp. 308–312.

[51] L. Bottou and C.-J. Lin, "Support Vector Machine Solvers," in Large Scale Kernel Machines, L. Bottou, O. Chapelle, D. decoste, and J. Weston, Eds. Cambridge, MA, USA: MIT Press, 2007, pp. 301–320.

[52] Z. Sahri and R. Yusof, "Support vector machine-based fault diagnosis of power transformer using k-nearest-neighbor imputed DGA dataset," J. Comput. Commun., vol. 2, no. 9, pp. 22–31, Jul. 2014.

[53] T. Sorsa and H. N. Koivo, "Application of artificial neural networks in process fault diagnosis," Automatica, vol. 29, no. 4, pp. 843 – 849, 1993

[54] A. Alaeddini and I. Dogan, "Using Bayesian networks for root cause analysis in statistical process control," Expert Syst. Appl., vol. 38, no. 9, pp. 11 230–11 243, Sep. 2011.

[55] J. E. Cook and A. L. Wolf, "Discovering models of software processes from event-based data," ACM Trans. Softw. Eng. Methodol., vol. 7, no. 3, pp. 215–249, Jul. 1998.

[56] W. M. P. van der Aalst and B. F. van Dongen, "Discovering petri nets from event logs," Trans. Petri Nets and Other Models of Concurrency, vol. 7, pp. 372–422, 2013.

[57] E. Myk and O. Unold, "Mining fuzzy rules using an artificial immune system with fuzzy partition learning," Applied Soft Computing, vol. 11, no. 2, pp. 1965 – 1974, 2011, the Impact of Soft Computing for the Progress of Artificial Intelligence

[58] B.-S. Yang and K. J. Kim, "Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals," Mechanical Systems and Signal Processing, vol. 20, no. 2, pp. 403 – 420, 2006.

[59] Z. Peng, M. Xiaodong, Y. Zongrun, and Y. Zhaoxiang, "An approach of fault diagnosis for system based on fuzzy fault tree," in MultiMedia and Information Technology, 2008. MMIT '08. International Conference on, Dec 2008, pp. 697–700.

[60] O. Kuzelka, J. Davis, and S. Schockaert, "Learning possibilistic logic theories from default rules," in Proceedings of the Twenty- Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016, S.

[61] Z. Zhang and M. I. Jordan, "Bayesian multicategory support vector machines," in UAI '06, Proceedings of the 22nd Conference in Uncertainty in Artificial Intelligence, Cambridge, MA, USA, July 13- 16, 2006. Corvallis, Oregon: AUAI Press, 2006

[62] B. Ye, Z. Luo, W. Zhang, and C. Piao, "Fault diagnosis for power circuits based on svm within the bayesian framework," in Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on, June 2008, pp. 5125–5129.

[63] D. Lowd and P. Domingos, "Naive bayes models for probability estimation," in Proceedings of the 22Nd International Conference on Machine Learning, ser. ICML '05. New York, NY, USA: ACM, 2005, pp. 529–536.

[64] M. Kearns and Y. Mansour, "Exact inference of hidden structure from sample data in noisy-or networks," in Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, ser. UAI'98. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998, pp. 304–310.

[65] N. Friedman, L. Getoor, D. Koller, and A. Pfeffer, "Learning probabilistic relational models," in In IJCAI. Springer-Verlag, 1999, pp. 1300–1309

[66] M. Richardson and P. M. Domingos, "Markov logic networks," Machine Learning, vol. 62, no. 1-2, pp. 107–136, 2006.

[67] S. Kok and P. M. Domingos, "Learning markov logic network structure via hypergraph lifting," in Proceedings of the 26th Annual International Conference on Machine Learning, ICML 2009, Montreal, Quebec, Canada, June 14-18, 2009, ser. ACM International Conference Proceeding Series, A. P. Danyluk, L. Bottou, and M. L. Littman, Eds., vol. 382. ACM, 2009, pp. 505–512

[68] C. P. Robert, T. Rydn, and D. M. Titterington, "Bayesian inference in hidden markov models through the reversible jump markov chain monte carlo method," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 62, no. 1, pp. 57–75, 2000.

[69] A. Nath and P. M. Domingos, “Learning relational sumproduct networks,” in Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA., B. Bonet and S. Koenig, Eds. AAAI Press, 2015, pp. 2878–2886.

[70] W. M. P. van der Aalst and B. F. van Dongen, “Discovering petri nets from event logs,” Trans. Petri Nets and Other Models of Concurrency, vol. 7, pp. 372–422, 2013.

[71] C. B. Dutta and U. Biswas, “Failure diagnosis in real time stochastic discrete event systems,” Engineering Science and Technology, an International Journal, vol. 18, no. 4, pp. 616 – 633, 2015.

[72] H. Chen and S. Lu, “Fault diagnosis digital method for power transistors

in power converters of switched reluctance motors,” IEEE Trans. Ind. Electron., vol. 60, no. 2, pp. 749–763, Feb. 2013.

[73] N. Freire, J. Estima, and A. Cardoso, “Open-circuit fault diagnosis in PMSG drives for wind turbine applications,” IEEE Trans. Ind. Electron., vol. 60, no. 9, pp. 3957–3967, Sep. 2013.

[74] M. Benbouzid, “A review of induction motor signature analysis as a medium for faults detection,” IEEE Trans. Ind. Electron., vol. 47, no. 5, pp. 984–993, Oct. 2000.

[75] S. Nandi, H. Toliyat, and X. Li, “Condition monitoring and fault diagnosis of electric motors—A review,” IEEE Trans. Energy Convers., vol. 20, no. 4, vol. 20, no. 4, pp. 719–725, Dec. 2005.

[76] S. Nandi, T. Ilamparithi, S. Lee, and D. Hyun, “Detection of eccentricity faults in induction machines based on nameplate parameters,” IEEE Trans. Ind. Electron., vol. 58, no. 5, pp. 1673–1683, May 2011.

[77] Y. Gritli et al., “Advanced diagnosis of electrical faults in wound-rotor induction machines,” IEEE Trans. Ind. Electron., vol. 60, no. 9, pp. 4012–4024, Sep. 2013.

[78] Reliability theory and practice. / Igor Bazovsky. Bazovsky, Igor.

https://babel.hathitrust.org/cgi/pt?id=mdp.39015000449549;view=1up;seq=29

[79] S. Guodong , S.J. Shin, S. Jain, Data analytics using simulation for smart manufacturing, In Proceedings of the 2014 Winter Simulation Conference, pp. 2192- 2203. IEEE Press, 2014.

[80] B. Dean, Data In Motion: The Next Frontier For Manufacturers, available at http://goo.gl/RcFqXk

[81] C. Weidig,P. Galambos, A.B. Csapó, P. Zentay, P.Z. Baranyi, Future Internet-based Collaboration in Factory Planning, Acta Polytechnica Hungarica 11(7) (2014) 157-177.

[82] M. Matzopoululos, Dynamic Process Modeling: Combining Models and Experimental Data to Solve Industrial Problems. In Process Systems Engineering, Wiley, Germany, 2010.

[83] B. Brad, M. Chui, J. Manyika, Are you ready for the era of ‘big data, McKinsey Quarterly 4 (2011) 24-35.

[84] W. Ribarsky, D.X. Wang, W. Dou, W.J. Tolone, Towards a Visual Analytics Framework for Handling Complex Business Processes, In System Sciences (HICSS), 2014 47th Hawaii International Conference on, 1374-1383. IEEE, 2014.

[85] M. Abish, R. Maciejewski, T.F. Collins, D.S. Ebert, Visual analytics law enforcement toolkit, In Technologies for Homeland Security (HST), 2010 IEEE International Conference on, 222-228. IEEE, 2010.

[86] S.G. Eick, Geospatial visualization with VisTracks, Wiley Interdisciplinary Reviews: Computational Statistics 2 (3) (2010) 272-286.

[87] M. A. Quddus, W. Y. Ochieng, R. B Noland, Current map-matching algorithms for transport applications: State-of-the art and future research directions. Transportation research part c: Emerging technologies, 15(5), 312-328, 2007.

[88] K. Zhao, Y.H. Yang and B.Z. Qu ,GPS / DR group and the navigation system map matching algorithm based on position points matching. Guidance and fuses, 24 (3) : 22-27,2003.

[89]L. X. Q. Liu, M. L. Z. Liu, (2007). Map matching algorithm and its application. Proceedings on Intelligent Systems and Knowledge Engineering (ISKE2007).

[90] I. Guyon, A. Elisseeff, An introduction to variable and feature selection. Journal of machine learning research, 3(Mar), 1157-1182 , 2003.

[91] R. Hecht – Nielsen, III.3 Theory of back propagation neural network, HNC, Inc. and University of California, San Diego, 1992, p. 65-93.

[92] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers. In: 5th Annual Workshop on Computational Learning Theory, ACM, (1992) 144–152.

[93] S.J. Lee, K. Siau, A review of data mining techniques. Ind. Manage. Data Systems: 101(1), (2001) 41–46.

[94] L. Breiman, Random forests – random features. Technical Report 567, Statistics Department, University of California, Berkeley, ftp://ftp.stat.berkeley.edu/pub/users/breiman.

[95] B. Hajek, Random processes for engineers, Cambridge University Press, ISBN 978-1-316-24124-0, 2015.

[96] L.E. Buam, T. Petrie, Statistical Inference for probabilistic functions of finite state Markov Chains, The Annals of Mathematical Statistics, 37(6) 1966, 1554-1563

[97] T. Vafeiadis, E. Bora-Senta, D. Kugiumtzis, Estimation of linear trend onset in time series, Simulation Modelling Practice and Theory, 19, 2011, 1384–1398.

[98] Ö. Hazır, X. Delorme, A. Dolgui, A survey on cost and profit oriented assembly line balancing. IFAC Proceedings Volumes, 47(3), 6159-6167, 2014.

[99] A. C. Nearchou & S. L. Omirou, Assembly Line Balancing Using Differential Evolution Models, Cybernetics and Systems, 48:5, 436-458, 2017.

[100] M. Mutingi, C.Mbohwa, Assembly Line Balancing. In: Grouping Genetic Algorithms. Studies in Computational Intelligence, vol 666. Springer, Cham, 183-197, 2017.

[101] D. Georgiadis, M. Kontaki, A. Gounaris, A. N. Papadopoulos, K. Tsichlas, Y. Manolopoulos. "Continuous outlier detection in data streams: an extensible framework and state-of-the-art algorithms." In SIGMOD International Conference on Management of Data, pp. 1061-1064. ACM, 2013.

[102] P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, K. Tzoumas. "Apache flink: Stream and batch processing in a single engine." Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 36, no. 4 (2015).

[103] P. Korvesis, S. Besseau, M. Vazirgiannis. "Predictive Maintenance in Aviation: Failure Prediction from Post Flight Reports." In Data Engineering (ICDE), pp. 1414-1422, 2018

APPENDIX A: MindSphere Platform Core

Identity Management

Idea

The Identity Management Service manages tenants, users and groups within the MindSphere platform. It enables customers to access the User Accounts and Authentication (UAA) service used within MindSphere for identity management and authorization.

Access

For accessing this service you need to have the respective roles listed in Identity Management roles and scopes.

Basics

Tenants: a tenant is an organization-specific logical environment for your data. A tenant typically represents a legal entity, such as a company or corporation. MindSphere is a multi-tenant architecture.

A tenant comprises up to two zones for identity management and access control:

- **User zone**: each tenant has a user zone. It enables users of the tenant to log in and use applications the company has subscribed to. Administrators of tenants can manage the users and assign roles in the user zone to provide users with access to subscribed applications.
- **Provider zone**: the provider zone represents an additional environment required for web application and API development and operation. Each DevOps plan tenant has an additional provider zone, which stores all information related to access management required for either development or operations (e.g., roles, permissions, applications, technical users).

As tenant names are global resources, they must be unique across all MindSphere tenants.

Each user tenant needs to define at least one administrator for managing the users and roles of this tenant.

Users and Groups

Every tenant has its own users and has roles available depending on the applications it is subscribed to. A role represents a grouping of permissions required to access an application. By modeling roles as SCIM groups, the User and Roles management in MindSphere follows the SCIM standard (System for Cross-domain Identity Management).

Currently, within the user zone, SCIM groups are only used to represent roles.

In addition, SCIM groups may represent user groups (for managing sets of users), data groups (for managing end customer access to data, assets, etc.) and permissions (for managing more fine-grained access to resources, only within provider zone of a tenant),

Roles and Scopes

The following information is relevant for tenants with a provider zone only. If you expose an API for your web application, scopes define the application specific permissions. Scope names typically reflect these permissions in a syntax like:

{apiName}.{permission/action}

The following list shows examples for the CRUD-permissions of an IoT service:

- iot.c (permission to create objects in IoT)
- iot.r (permission to read objects in IoT)
- iot.u (permission to update objects in IoT)
- iot.d (permission to delete objects in IoT)

Scopes are mapped to a specific role. A role name has the following syntax:

mdsp:{tenantName}:{application/apiName}.{roleName/action}

Thus, all scopes above could be mapped to a role called

mdsp:core:iot.admin

Application-specific roles and scopes are defined in provider tenants and can be managed within the developer cockpit application.

OAuth Client

The following information is relevant for tenants with a provider zone only.

An OAuth client (also called technical user) allows your application to acquire a token to access protected resources without the need of an interactive user to currently use your application.

An OAuth client is useful, for example, for doing regular background activities (batch activities) within your application or if your application is not hosted behind MindSphere Gateway and therefore does not find interactive user tokens in request headers. OAuth clients are defined within the provider zone of your tenant and comprise a client ID and client secret, which allow to obtain a token using the client credentials grant (RFC 6749).

Features

The Identity Management Service exposes its API for realizing the following tasks:

- List all users of a tenant
- Create, get, update, delete users of a tenant
- Get all roles assigned to the own user
- List all SCIM groups of the user zone of a tenant
- Create, get, update, delete SCIM groups of the user zone of tenant
- List, add, remove members of a SCIM group of the user zone of tenant

Example Scenario

The administrator of a brewery wants to prepare the tenant for the new developers of their web application. Use the Identity Management Service to populate the tenant with new users and assign them the roles required for development (e.g., mdsp:core:StandardUser, mdsp:core:Developer).

OAuth Authorization Server (API)

The OAuth Authorization Server provides APIs for authentication and authorization. In particular, it provides APIs for requesting access tokens and retrieving the public key used for signing Json Web Tokens issued by the OAuth Authorization Server.

Tenant Management Service

Idea

Tenant Management is a central service within the MindSphere platform that provides the possibility to manage a tenant, its related subtenants as well as related data.

Access

For accessing this service you need to have the respective roles listed in Tenant Management roles and scopes.

Basics

Tenants: a tenant is an organization-specific logical environment for your data. A tenant typically represents a legal entity, such as a company or corporation. In this context MindSphere is a multi-tenant architecture.

Subtenants: a subtenant is an organization-specific logical environment for the data of a cooperating legal entity. A subtenant typically represents a subset of the tenant's data which the tenant wants to share with one or more of its customers.

The tenant information stores the following parameters:

- Country: the home base of your legal entity
- displayName: the brand name which is shown in the MindSphere environment
- name: the name which is shown in the URL of the tenant
- prefix: an ID which is used by developers to identify relevant code components
- type: the tenant type: USER, DEVELOPER or OPERATOR

Legal Information of a Tenant

A tenant can display its own legal/branding information, which is defined as legal links of type www, phone, and mail. When tenant branding is enabled, it automatically creates the region global and sets the default language to English. Tenants can be located in different regions with different legislation, which might require to display different legal links depending on their location. For this case you can define custom regions and using country codes specify which countries belong to each region. One country can only belong to one region. You can also define different languages and add them to multiple regions. For each region you can define individual legal links in multiple languages.

The customized legal links are displayed for all users of the tenant except tenant admins, who still see the original legal links of MindSphere. The tenant's location and the language settings of the web browser determine, which legal links are displayed.

Features

The Tenant Management Service exposes its API for realizing the following tasks:

- Retrieve tenant-specific information
- Manage your own legal information
- Create, retrieve and delete subtenants

Example Scenario

Providing Legal Information: the administrator of a machine manufacturer updates the original Siemens legal information of the tenant with the company's own relevant legal information in order to provide this information to other tenants and subtenants.

Creating subtenants: the administrator of a machine manufacturer wants to create a subtenant for one of their customers in order to share some relevant machine data with this customer.

Usage Transparency Service

Idea

Usage Transparency Service offers a UI giving insight into your resource consumption on the MindSphere platform. For developers it also offers an API to track and retrieve metrics defined by the developers. The Usage Transparency Service (UTS) gathers various consumption metrics on the MindSphere platform to fulfill requirements regarding:

- API resource consumption by available tenants
- Apps consumption by available tenants
- User details for a particular time period
- History details of subtenants
- Number of device requests by tenants

Access

For accessing this service you need to have the respective roles listed in Usage Transparency Service roles and scopes.

Basics

Admin: the admin has all the administrative privileges. These privileges include viewing and listing data of all the tenants, subtenants and the users.

Tenant: client to an organization. For example: ABC is a tenant to Siemens, where Siemens is an organization and ABC is the client/tenant.

SubTenant/T2 tenants: client to a tenant. For example: PQR is a client to ABC, where ABC is the tenant and PQR is the subtenant.

User: all the email account holders belonging to the subtenants and tenants.

Tenant: Tenant admin.

Features

The Usage Transparency Service exposes its API for realizing the following tasks:

API/ Service

- Get details of all transactions
- Monitor data consumed in the transactions
- Get endpoints information which are used by consumers to fetch data from MindSphere
- Monitor the relative data of all the API calls passing through MindGate
- Monitor the relative API details including the dependent API details

Apps

- Monitor consolidated data consumed by applications
- Monitor apps used by users

Users

- Get total number of subtenants
- Get total number of users

Asset

- Monitor inbound data size (request size)
- Monitor devices registered with the tenants

Example Scenario

The manager of a wind turbine production plant wants to analyze the wind speed for energy production. The wind farm has several turbines. A thermocouple is connected to the turbines. The frequency usage of this device and the number of users accessing the device are to be monitored. The device pushes the data for the display using the southbound endpoints. A notification service email is required to be sent to the maintenance team for regular checking of the turbines. API calls are made to the applications to monitor the temperature of the turbines. This helps in maintaining and stabilizing the turbines to be intact.

APPENDIX B: IoT and Storage

IoT File Service

Idea

The IoT File Service provides the user with file management for files related to assets. The service enables the user to carry out the following file management tasks:

- Read, upload, update and delete files associated with assets
- Save metadata of uploaded files
- Search for files by metadata
- Upload files to a temporary storage and manage there
- Download files to specified agents

Access

For accessing this service you need to have the respective roles listed in IoT File Service roles and scopes. A user can only interact with files within their tenant and subtenants.

Basics

The service implements CRUD (create, read, update and delete) functionality and searching for files.

Each file is related to an asset. An asset can have multiple files attached. The asset ID is required for each request.

Each file has a set of standard properties as metadata:

- File Id
- File name
- File path
- File type (e.g., image)
- File size
- Description

- Time stamps (created, updated)

The content of these files is not parsed by MindSphere and it requires custom applications or analytical tools to interpret and visualize the data. The maximum file size that can be stored or retrieved depends on the client's network speed. The current timeout is one minute, and any file that takes longer than one minute to read or write will result in an error 502 "Bad Gateway".

Versioning: the service does not provide explicit versioning. If different versions of a file need to be stored, put a version identifier in the file name to make it unique.

File deletion: files are always deleted physically. There is no logical delete. Files of an asset are automatically deleted if the asset is deleted physically.

Search filter syntax: files can be queried by filtering on any of their available file properties. Complex filters using multiple clauses and various operators are possible. The allowed operators in a filter clause are:

- eq or =
- gte or >=
- gt or >
- lte or <=
- lt or <

Filters can use the wildcard character * on any field. Multiple filters can be specified using the and operator (or is not supported).

The result set of a filter request can be sorted with the order parameter. When sorting by name, the result is automatically sorted by path as well. Sorting is only allowed on exactly one property per request and sorting can only be done on the following properties:

- File Id
- File update timestamp
- File path
- File name

Note that time stamps are stored as strings in ISO 8601 format in the UTC timezone. Be aware that if an entity time stamp of "2017-09-21T14:08:22.345+02:00" is written in the file property, searching for this file by time stamp would have to be a search for "2017-09-21T12:08:22.345Z".

Features

The IoT File Service exposes its API for realizing the following tasks:

- Create a new file
- Update (overwrite) an existing file
- Read a file
- Delete a file
- Search for files by metadata
- Count and order files by metadata

Example Scenario

The application developer of a brewery wants to store files associated to assets in order to provide additional meta information (e.g., images, technical descriptions, manuals, etc.). The developer uses the IoT File Service to upload files to an asset.

IoT Time Series Service

Idea

The IoT Time Series Service is used to create, read, update, and delete time series data. Time series data is stored against an asset (entity) instance and an aspect (property set). Within the IoT Time Series Service you can store and query time series data with a precision of 1 millisecond.

Access

For accessing this service you need to have the respective roles listed in IoT Time Series Service roles and scopes. A user can only read data within his tenant and subtenants.

Prerequisites

The prerequisites depend on the use case:

• Ingesting time series data from a field device: an onboarded asset (entity) is connected to MindSphere and produces time series data. A valid mapping must exist from device data points to asset variables.

• Ingesting time series data from an application: an asset (entity) with respective variables must exist.

Basics

A time series record consists of a timestamp, one or more values, and an optional quality indicator for each value, as defined in the aspect type definition. If there are multiple variables in the aspect, they are expected to arrive in the same payload. Writing a time series record with the same timestamp, asset (entity) instance, and aspect as a record overwrites the old record. There is no versioning of data. The time stamp may be specified with millisecond precision. Measurement values must match the type that is defined in the aspect type. Time series data can be of type int, long, double, boolean, string, big_string (blobs), or timestamp. The maximum sizes of strings and big strings are defined in the aspect type, and can be up to 255 and 100,000 respectively.

In the aspect type definition, there is a qualitycode field that indicates whether the time series data is accompanied by quality code data. If the qualitycode value is true for a variable (property), the quality code property name is the same as the variable (property) name but with a _qc suffix. For example, the quality code property for the temperature variable (property) is temperature_qc.

Multiple records for an asset (entity) instance and aspect can be written in one call to the service. Records can be read and deleted by identifying the asset (entity) instance, aspect (property set), and time range.

When writing or deleting data, the default behavior of the service is to queue the incoming request and perform the physical writes to the underlying data store asynchronously. In this mode, if a user writes time series data and immediately tries to read it back, it may not yet be present in the data store and won't be returned. If a user deletes data and then reads that time range, the data may not have been physically deleted yet, and could be returned in the response. Requests for an asset (entity) sent in queued mode are processed in the order they are received.

Features

The IoT Time Series Service exposes its API for realizing the following tasks:

• Read time series data for a single asset (entity) and aspect (property set)

- Return data for a specified time range
- Return the latest value if no range is provided
- Write or update time series data for a single asset (entity) and aspect (property set)
- Overwrite existing time series data
- Delete time series data for a single asset (entity) and aspect (property set) within a given time range

Example Scenario

A wind turbine is connected to MindSphere. The blades move with a constantly changing speed. The sensor measures the speed and sends the data via MindConnect to MindSphere.

IoT Time Series Aggregates Service

Idea

The IoT Time Series (TS) Aggregates Service is used to read aggregated interval data that is calculated based on raw time series data. The IoT Data Service automatically creates aggregated summaries of all numeric time series data. This allows applications to retrieve smaller data sets that cover a longer period of time with much better performance than processing all the raw time series data. For example, an application could request daily summary data for a month, obtaining 30 records for an aspect (property) instead of ~2.5 million records of 1 second data for that same period.

Access

For accessing this service you need to have the respective roles listed in IoT Times Series Aggregates Service roles and scopes. A user can only read data within his tenant and subtenants.

Basics

The IoT TS Aggregates Service automatically creates pre-calculated aggregates and stores them for later retrieval to enhance performance on aggregate queries. The following pre-calculated intervals are available:

- 2 minute

- 1 hour
- 1 day

Intervals smaller than the smallest pre-calculated interval are also available and generated on the fly. When retrieving aggregate data, the client needs to specify the desired interval duration. The following durations are supported:

- 1-120 seconds
- 1-60 minutes
- 1-24 hours
- 1+ days
- 1+ weeks
- 1+ months

For best performance, the aggregation interval should be an exact multiple of a pre-calculated aggregation interval, such as 10 minutes, 2 hours, 5 days, 3 weeks, or 2 months. The service also supports other requested aggregation intervals and uses raw time series and pre-calculated aggregations to calculate results where needed. For example, an application could ask for 30 second intervals or 15 minutes intervals. These requests take longer to process.

The interval durations available have some overlap that provides flexibility to the application. For example, 1 hour and 60 minutes represent the same time period, but allow different behavior. A 1-hour interval duration allows a user to request data from 5:00 – 10:00. A 60-minute interval duration can be used for that same time period, but can also be used for 5:30 – 10:30. The 1-hour duration would be rejected for 5:30 – 10:30 because the times don't align with hourly boundaries. Similarly, if the tenant is configured with the first day of the week being Monday, an interval duration of 1 week must start and end on a Monday. An interval duration of 7 days can start on any day of the week.

Each aggregated record contains an assortment summary of the following information:

- first value and time of first value of the interval
- last value and time of last value of the interval
- minimum value and time of the minimum value during the interval

- maximum value and time of the maximum value during the interval
- the count of the values during the interval
- the sum of the values during the interval
- and the average of the values during the interval.

The aggregated data is wall clock aligned based on the asset's (entity's) time zone.

- Two-minute data ends at two minutes after the hour, four minutes after the hour, and so on.
- Hourly data ends at the top of the hour.
- Daily data ends at midnight in the entity time zone.

Time zones that are 15 or 45 minutes off a UTC hour are not supported, such as Nepal standard time (UTC+05:45). Time zones that are 30 minutes off a UTC hour are supported, such as India (UTC+05:30). Time zones that are an even number of hours off of GMT are supported.

The data summarized in an aggregated interval is the data with time stamps greater than the start time of the interval and less than or equal to the end time of the interval. This ensures that count type of data is correctly reflected in the aggregated interval. For example, if a device provides a wall clock aligned count at the end of each five minutes, an hourly aggregation includes the data with time stamps from five minutes after the hour until the end of the hour (3:05 to 4:00 for example).

Aggregates are created based on arrival of data and time logic.

- Two-minute aggregated data is created within two minutes after the two-minute interval has finished.
- Hourly data is created within eight minutes after the end of the hour.
- Daily data is created within fifteen minutes after midnight in UTC time for this release.

Aggregated data is only created if time series data exists for that period. Aggregates are created based on the data that is available at the time. Late arriving data is not included in the initial aggregated results, but triggers a recalculation after a delay.

If a variable (property) is configured to contain quality code data, the service uses it to calculate the aggregates. The quality code is compared to the good and uncertain thresholds defined on the tenant (per default the values are set according to the OPC UA standard). Good data is defined as being anything less than or equal to the goodThreshold value. Uncertain data is defined as being anything greater than the goodThreshold value and less than or equal to the uncertainThreshold value. Bad data is defined as anything greater than the uncertainThreshold. Good and uncertain data is used to calculate the summary values listed above. Bad data is ignored for these calculations. Each aggregated record contains counts of the three types of data, countgood, countuncertain, and countbad. If no quality code data is present on a time series record, it is considered good.

Features

The IoT TS Aggregates Service exposes its API for realizing the following tasks:

- Read aggregated interval data
- Specify the time range
- Specify the interval duration
- Select only the summary information you are interested in

Example Scenario

A wind turbine produces time series data for the speed of the blades. The speed of the blades changes only every hour. The IoT TS Aggregates Service API allows you to collect just the accumulated time series of one hour.

APPENDIX C: Connectivity

Agent Management Service

Idea

The Agent Management Service is used to onboard, offboard, update and delete agents. It provides connectivity functions to enable communication with the MindSphere Platform.

The Agent Management Service is typically used by application developers or machine builders (OEMs). Agent Management API provides agent provisioning and configuration functionality.

Access

For accessing this service you need to have the respective roles listed in Agent Management roles and scopes.

Users can interact only with agents onboarded within their tenant.

Basics

Agents: an agent is the primary actor within the MindSphere environment. Every action is directly or indirectly related with an agent. As an example an agent uploads data, retrieves its events, changes its configuration etc. The very first step to use MindSphere APIs is to create an agent.

When the agent is created, an initial access token is generated (IAT). This IAT is a JSON Web Token (JWT) that holds various information about the agent. This token needs to be downloaded to the agent and provided during the onboarding step.

Agents need to "onboard" to MindSphere prior to using any services. Onboarding simply means registering an agent, so that MindSphere authorizes and authenticates the agent. During this onboarding step, the agent needs to provide the IAT to MindSphere to validate its identity. MindSphere validates the agent by checking the signature of the IAT.

The agent communicates its credentials based on one of the following security profiles:

- SHARED_SECRET
- RSA_3072.

SHARED_SECRET Security Profile: for agents with this security profile, MindSphere generates a secret for the agent and stores it in its persistent storage. This secret is returned to the agent in the onboarding response during the onboarding step.

RSA_3072 Security Profile: agents with this security profile first send their public key to MindSphere when onboarding. MindSphere stores the public key in its persistent storage.

If onboarding is successful, MindSphere responds with a Registration Access Token (RAT), which is used to renew registration to update credentials when the agent's credentials are expired.

Access Token: after onboarding, agents need to acquire an access token. Acquiring an access token involves creating a self signed JWT, and sending this self signed JWT to MindSphere. Upon receiving the access token request MindSphere validates the signature of the JWT with the stored credential of the agent. Based on the security profile either shared secret or public key will be used to validate the agent's self signed JWT.

If the self signed JWT is valid, MindSphere responds with an access token. An access token is a time restricted JWT token that holds various information and the agent's scopes (access rights). Agents can use the access token to consume the MindSphere services until expiration. After expiration, the agent needs to acquire a new access token to continue using MindSphere services.

Data Source Configuration: it involves configuring Data Points for an agent. This configuration is mandatory for MindSphere to interpret uploaded data. Without this configuration MindSphere does not understand the data an agent uploads. When the agent is first created, its Data Source configuration is empty. This configuration must be updated via the update endpoint.(/agents/{id}/dataSourceConfiguration).

A Data Point represents measurements done by either a sensor or a device. For example, "Temperature" and "Torque" can be Data Points for a Data Source configuration.

Boarding: the Boarding API provides the following functionality:

- retrieve onboarding configuration
- get boarding status of an agent
- offboard and agent.

The boarding configuration consists of onboarding data, the Initial Access Token (IAT) and a registration URL. The boarding configuration is used by the agent to onboard to MindSphere.

MindSphere users can retrieve an agent's boarding status. The boarding status contains three different statuses, namely NOT_ONBOARDED, ONBOARDED and ONBOARDING. They are defined as follows:

- NOT_ONBOARDED: Agent is not authorized to exchange data with MindSphere.
- ONBOARDING: Onboarding configuration for the agent is ready and MindSphere is waiting for the agent to trigger the onboarding process.
- ONBOARDED: Agent is authorized to exchange data with MindSphere.

The Boarding API also provides an interface to offboard the agent. When the agent is offboarded, a brand-new IAT is generated by MindSphere. Using this new IAT the agent can be onboarded again.

Register: the registration process takes place in accordance with the OAuth 2.0 Authorization protocol (RFC 6749). The boarding configuration retrieved from the /agents/{id}/boarding/configuration is required to register (onboard) the agent to MindSphere. In order to be registered, the agent needs to send a request with the Initial Access Token (IAT) from its boarding configuration. The structure of a request varies slightly depending on the Agent's Security Profile. The IAT Key is valid for one week (7 days). The /register endpoint returns a Registration Access Token (RAT), which is valid indefinitely and required for key rotation. Agent credentials have to be updated every 7 days, regardless of the security profile. In the update process the agent has to provide the RAT instead of IAT to update its credentials, otherwise the process is the same as for initial registration. After the registration is completed for an agent, its boarding status is set to ONBOARDED. The key rotation process is performed via the /register/{id} endpoint, and enables agents to change their key (symmetric or asymmetric). The RAT is required in the request.

Token: each registered (onboarded) agent is required to get an access token in order to use any of the services offered by MindSphere. MindSphere grants access tokens to onboarded (registered) agents. In order to get an access token, the agent needs to create a Json Web Token (JWT) which holds information such as agent ID, tenant name etc., and sign it with its shared_secret/private_key based on its security profile. The /OAuth/token_key endpoint which returns the public key of the server, is provided to the agents in order to enable them to verify any access token granted by MindSphere.

Features

- Create, edit, remove agents
- Onboard and offboard agents
- Acquire agent onboarding configuration
- Define asset - agent relations
- Define an agent's Data Sources
- Acquire access tokens to consume Connectivity Services

Example Scenario

The application developer of a brewery wants to programmatically on- and offboard MindConnect devices connected to the production lines.

The developer uses the Agent Management Service to register and offboard the desired devices.

MindConnect API

Idea

The MindConnect Service exposes an API that enables shop floor devices to send data securely and reliable to MindSphere. It opens the MindSphere platform to custom applications to collect and send data which shall be stored and used by applications in the cloud. The MindConnect Service enables the development of custom data collectors, also referred as custom agents. This software applications act as data sources that upload the collected data into MindSphere.

Access

For accessing this service you need to have the respective roles listed in MindConnect API roles and scopes. The custom agent needs a field-side network infrastructure to forward and route outbound HTTPs requests to the Internet. MindConnect supports multiple agent device classes: strong hardware platforms as well as resource constrained devices. All target agent platforms must comply to the following minimum requirements:

- HTTP processing
- TLS
- JSON parsing
- JSON Web Token (JWT) generation
- HMAC generation (preferably SHA2 based hashing)

Basics

Data Source Configuration: if an agent is to upload data to MindSphere, MindSphere needs additional configuration to know how to interpret the agent's data stream. This configuration requires the following definitions within MindSphere:

- Data Source Provisioning
- Property Set Provisioning
- Mapping for Data Source and Property Set

A Data Source is a logical group that holds so called Data Points. Data Points hold metadata about a specific metric that the agent generates or measures. For example, if an agent measures ambient temperature and pressure data, each of these two measurements need to be defined as a separate Data Point:

- Data Point 1: Temperature measurement
- Data Point 2: Pressure measurement
- A Data Source on the other hand is an encapsulating/grouping object for the Data Points.

Data Points are set by a Data Source Configuration. MindSphere provides the dataSourceConfiguration endpoint of the Agent Management Service. For more details refer to Creating a Data Source Configuration.

Standard data type: the MindConnect service uses so called standard data types. Standard in this context means:

- The API defines how standard data types have to be transmitted, e.g. how metadata and production data needs to be formatted as HTTPs payloads.

- For each of the standard data types, there are predefined routing mechanisms which allow an automated parsing and storing of that information to (virtual) assets within MindSphere.

- For each of the standard data types, there is a preconfigured mass data storage available.

- Data of standard types can be accessed and queried in a standardized way by applications and analytical tools in MindSphere.

In contrast to custom data types, there is no additional configuration or any coding required for parsing and storing data provided by a custom agent. It is a fully automated functionality provided by MindSphere.

MindSphere supports the following standard payload data types for production data:

- Time Series are simple Data Point values that change constantly over time, e.g. values from analog sensors like a temperature sensor. This also applies to any other measured values that have an associated time stamp.

- Events are based on machine events, e.g. emergency stop or machine failure occasions. However, this mechanism can also be used to propagate custom agent driven notifications, e.g. if you do on-site threshold monitoring and want to report a broken threshold.

- Files with this data type you can upload files of up to 9MB per exchange call. The files are attached to the corresponding (virtual) asset, e.g. device log files or complex sensor structures. Files that are uploaded can be referenced by their parent (virtual) asset in MindSphere. The content of these files is not parsed by MindSphere and it requires custom applications or analytical tools to interpret and visualize the data.

- Data Models: description of the agent-side asset hierarchy and configuration including measurement points. For some custom agents it is more convenient to upload the data model to MindSphere directly. This data is used by MindSphere to dynamically create (virtual) assets, aspects, variables or mappings.

Data Point Mapping: a Data Point Mapping needs to be defined for MindSphere to interpret the data flowing from the agent to MindSphere. The Data Source configuration holds metadata about the agent side where a Property Set holds metadata about the IoT side. Finally, MindSphere needs information to map from Data Point meta to Property meta. This configuration is called Data Point Mapping and defines a mapping for each Data Point to a property.

Data Exchange: the MindSphere exchange endpoint of the MindConnect API provides the agent with the capability of uploading its data to MindSphere. This data can be of type:

- Time Series
- File Upload
- Event Upload

The format conforms to a subset of the HTTP multipart specification, but only permits nesting of 2 levels.

Features

A custom agent consumes the MindConnect Service for realizing the following tasks:

- Upload time series
- Upload files
- Poll for events
- Create, receive and acknowledge business events (Alarms)
- Describe and upload asset data models
- Upload data of custom data types for custom handling in MindSphere
- Download files from MindSphere repository (e.g. for firmware or configuration updates)

Example Scenario

The manager of a wind farm wants to collect sensor data of a wind turbine. A developer writes a field application (agent) which collects the sensor data. The data is sent to MindSphere via the MindConnect API.

APPENDIX D: Advanced Services

Asset Management Service

Idea

An asset can be a digital representation of a machine or an automation system with one or multiple automation units (e.g. PLC) connected to MindSphere. Aspects are a data modeling mechanisms for assets. Aspects group the data points based on logical sense. The Asset Management Service is used to create an asset. Within the Asset Management API you can read the asset data and manage your assets. You can retrieve all aspects and variables of an asset, change the location, assign files. The service is also used for creating asset types (templates for assets). Aspect types (template for group of variables) can be used in asset types, to help create the same properties multiple times.

Access

For accessing this service you need to have the respective roles listed in Asset Management roles and scopes.

Basics

The API divides the functions into categories, for each category you can use a different controller. This section describes the controllers:

Aspect types: it is a template for creating multiple aspects with the same variables. They can be used in asset types. You can configure, read or delete your aspect types. An aspect type is accessed by its ID. There are predefined aspect types, which are available for all users, but cannot be modified or deleted.

Asset types: it is a template for creating multiple assets with the same variables. You can configure, read or delete your types. An asset type is accessed by its ID. There are predefined asset types, which are available for all users, but cannot be modified or deleted.

Assets: creates the following types of assets:

- Device types: represents a machine, or any object, from which the data should be collected of.

- Agent types: represents the agent (software or physical device) measuring and collecting data of devices, machines, etc.

- Hierarchy types: represents the hierarchy levels of an organization.

- Application types: used by other services to collect data of their applications (e.g. Edge Analytics Application).

You can configure, read and manage an asset or the root asset. An asset is accessed by its ID. Assets can be deleted, except for root assets, which are deleted when their tenant is deleted. When reading an asset, its static variables and aspects are filled with their values. In addition, you can assign existing files to assets. Assets inherit their parent's type assignments, but not their assets. Assets can also overwrite the inherited type assignments.

Structure: the structure of an asset can only be read. It will show all aspects and variables without their values. You can configure the variables' values by updating the asset.

Location: you can manage the location of your assets and update or delete the location data.

Files: you can upload files, and assign these files to assets or asset types. One file can be used multiple times, or not at all.

Features

The Asset Management exposes its API for realizing the following tasks:

- Create types for creating multiple assets with the same structure
- Create an asset in four main topics
- Get a list of all available assets
- Get a list of all available asset types
- Get a list of all available aspect types
- Manage all aspects of an asset
- Manage the location of your assets
- Upload and assign files to assets

Example Scenario

A brewery has moved his conveyor belt to a different location. A fixed installed camera takes images of the conveyor belt every 2 minutes. The Asset Management API allows you to change the location of the conveyor belt. You can define a new aspect for the conveyor belt like production volume

Event Management

Idea

The Event Management Service manages different types of events created by users in MindSphere. It enables users to create, update and delete events associated to entities.

Access

The Event Management is exposed to developers as a REST API. For accessing this service you need to have the respective roles listed in Event Management roles and scopes.

Basics

Users can use predefined event types or create their own templates for generating new events. The user is able to filter, update or delete the events or event types.

Events: there are two different kinds of events: standard events and custom events. Both inherit the following fields from a common ancestor:

- id
- correlationId
- timestamp (required)
- entityId (required)

The following additional fields are available:

- Standard Events.

A standard event is created from a predefined event type, for which the following base properties are available:

- description
- severity

- code
- source
- acknowledged

• Custom Event: a custom event is created from a user-defined event type. The available properties are defned in the event type.

Event types: custom events are created from user-defined event types. These define a set of custom fields which hold the properties of the custom event. Each custom field of an event type has a data type and can be defined to be required and/or filterable. An event type has a global or local scope and a time to live (TTL). Both are applied to all instantiated events. Event types can be derived from other (parent) event types. An inheritance hierarchy of up to 4 levels is allowed. Every event type inherits all fields from its ancestor event types.

Filtering and sorting: events can be filtered and sorted by all assigned base properties except description. Additionally, if the event type is set as filter parameter, events can be filtered and sorted by all custom fields specified as filterable. Events can also be filtered and sorted by all inherited custom fields. The filter syntax can be customized using operators like and, or, not equals, contains or startsWith. Be aware that only a single, not negated event type can be used as filter parameter.

Updating: Events can be updated. There are two ways to update an event: A PUT request overwrites an event and a POST request creates a history. This means the original event remains in the database and a new event is created with the same entityId and correlationId but a different id. Listing events returns only the latest of an event's history. The whole history of the event is listed with history query parameter set to true.

Deleting: events can be deleted based on a custom filter asynchronously via POST request.

Features

The Event Management Service exposes its API for realizing the following tasks:

- Create an event.
- Get a single event.
- List filtered and sorted events.
- Update an existing event (PUT).
- Create a historic event while updating an event (POST).
- Create, update, delete an event type.
- Get a single event type.

- List filtered and sorted event types.
- Delete events asynchronously based on custom filter.
- Create events asynchronously.

Example Scenario

The manager of a brewery wants to track machine faults of the brewery's production line. The manager uses the Event Management Service to create an event with an error code.

Data Flow Engine

Idea

The Data Flow Engine (DFE) enables customers to create workflows based on platform events and ingested data. These workflows can be used for example to automate KPI calculations in 3rd party apps.

Access

For accessing this service you need to have the respective roles listed in Data Flow Engine roles and scopes.

Basics

Workflows defined by DFE can be used to filter and process ingested data, react on events, calculate KPIs, aggregate data, modify the data model, write data into platform data lake and to send notifications or create platform events. These workflows are called streams. They consist of a chain of predefined apps. These apps are autonomous, deployed processing units and can be combined and configured to do real time computation by consuming and processing an unbound amount of data. The number of apps is continually enhanced and can be extended by custom apps in the future. Once a stream is deployed, it will automatically start to ingest the configured data flow of the tenant that has deployed it.

The DFE is implemented based on the Spring Cloud Data Flow architecture. Refer to the Spring Cloud Data Flow documentation to understand the idea behind the engine. Similar concepts can be found under different terminologies, such as like triggers - sources, rules - processors, actions - sinks.

DFE streams

The DFE uses message based streams with configurable apps. It differentiates between three types of apps:

- Source: Input data in order to trigger a flow.
- Processor: Filter or manipulate the input data.
- Sink: Execute an action based on the flow's result.

A working stream must contain one source app and one sink app. It can also contain one or more processors between them. MindSphere can store any number of stream definitions, but the number of concurrently deployed streams is restricted according to the purchased plan.

Using these components the model of a DFE stream can be outlined as here:

The apps listed below are already available, but DFE can be enhanced based on general usage or custom needs.

Sources

Source apps are used to listen to different source entities or events. This type of application can only stand at the beginning of a stream. A source may also provide an opportunity for filtering to decrease the number of messages flowing through the stream.

TimeSeriesSource: it monitors time series uploads and triggers a stream. This is the entry point for time series data into a DFE stream.

TimerSource: it sends timestamp messages based on a fixed delay, date or cron expression.

Processors: processor apps can be placed inside a DFE stream. Processors can interrupt the stream, but they leave the input and output data (JSON) untouched.

ChangeDetector: this app filters out all messages that have the same value in the specified triggering variable field as the previous message. The differentiating key can finetune the granularity of the filtering by arranging the messages with the same differentiating key into a filter group. That way for example one can use a single flow and perform separate filterings for every asset.

FilterProcessor: this processor interrupts the stream if the configured expression is not fulfilled. This app can be used to filter messages in a DFE stream with custom expressions defined in Spring Expression Language.

MessageSampler: the MessageSampler provides the "limit requests in time" functionality. The purpose of this element is to prevent the overload of the downstream flow by suppressing messages from the same device for a given time period. It is important to note that the MessageSampler works based on server time and ignores the timestamp sent with the time series data. Suppose we have a MessageSampler set up to allow messages through every 2 minutes and the following time series data given:

1. Message arrives at 11:27:00

{

{someKey} : {someValue},

"_timestamp" : "2018.06.07T11:06:00.000Z"

}

2. Another message arrives at 11:28:00 (one minute after the first one)

{

{someKey} : {someValue},

"_timestamp" : "2018.06.07T11:21:00.000Z"

}

According to the timestamp, 15 minutes have passed, but the MessageSampler ignores this value. The second message is not sent through. Even though the timestamp indicates otherwise, the messages arrived only 1 minute apart from each other (according to server time), not 15 minutes apart. This is a valid scenario because of the asynchronous nature of the system.

Hysteresis

Hysteresis can be used to react to events with a little slack space. For example, someone wants to know if an engine's RPM goes above a predefined threshold, but they don't want a lot of messages because of oscillations around the threshold. They can choose to suppress messages until some condition is met, like the RPM drops below another threshold value.

Here is an example hysteresis definition:

Message emitting expression: "MAINMOTOR.rpm > 100"

Suppression releasing expression: "MAINMOTOR.rpm < 80"

This way Hysteresis only emits one message if the value is above 100 and then suppresses any further messages until the value drops under 80, instead of emitting a message each time the motor oscillates between 99 and 101.

If the suppression expression is not defined, the app emits a message for each event which fulfills the message emitting expression.

The --message-emitting-expression and --suppression-releasing-expression parameters can be defined as Spring Expression Language expressions.

Sinks

Sinks are the ending apps of a DFE stream. They only have inputs and perform the last step of a stream. This can be a notification or a change inside the platform.

EventSink

The EventSink receives a message, converts it to an event based on the configured parameters and forwards it to the Event Management Service.

EventSink supports field mapping, which means that a JSON can be defined to map attributes from the incoming message to the event to be created. The parameters description, source, code, severity and acknowledged also can be defined as expressions, like custom fields in the field-mapping-json. Of course, they can be defined as static values as well. For more details check the field-mapping-json parameter in EventSink parameters list.

The following picture shows the decision process to determine whether the system should use an existing event type or create a new one:

Data Flow Server

The Data Flow Server (DFS) acts as a management service for the Data Flow Engine, it provides functionalities for listing and modifying streams and their components. To deploy a new DFE stream, a user has to call the API endpoints on the Data Flow Server.

Features

Use the Data Flow Engine API for realizing the following tasks:

- Create or delete streams
- Get the list of existing streams or find an exact stream by name
- Deploy or undeploy a stream

- Get the list of runtime apps or find an exact app by name
- Get the specific app instances

The Data Flow Engine is able to:

- Subscribe to time series data flows from MindSphere
- Send events when a specific stream has been triggered
- Filter messages containing irrelevant information

Example Scenarios

A client wants to monitor an engine in a factory. The engine is a critical part of the manufacturing process and has a history of overheating (Simple Scenario). Conclusions about the engine's condition can be drawn from how much it resonates during operation (Extended Scenario).

Simple Scenario

The client creates a Data Flow Engine stream with a TimeSeriesSource, a FilterProcessor and an EventSink. The TimeSeriesSource listens to a heat sensor placed on the engine. The FilterProcessor is set to only forward messages of temperatures higher than an allowed threshold. In the EventSink chooses a built-in event, instead of customizing one. The sink stays dormant for a while, because the filter does not forward messages until the temperature threshold is reached. Once it is, the client receives notifications about the engine overheating.

Extended Scenario

The same client wants to monitor the resonance of the engine. The client sets up the TimeSeriesSource and the EventSink as for monitoring the temperature, but inserts a FilterProcessor and a MessageSamplerFilter between them. The FilterProcessor is set to only forward messages when the resonance is higher than a threshold (but still lower than what is expected at the end of the lifecycle of the engine). The MessageSamplerFilter is set to only forward messages once per day so the client gets a daily reminder, but can wait for a good offer from the engine manufacturers without being bombarded with notifications.

Notification Service

Idea

The goal of digitalization is the efficient and rapid exchange of information. Lots of valuable information/data are generated using MindSphere like analytics, events, etc. These need to be transported quickly to your customers so that they can react quickly based on the information type. The notification service provides various interfaces (channels) to communicate and share information among the users of MindSphere via

- E-Mail
- Push notification
- SMS

Access

- The Notification Service API is only available in developer or operator tenants.
- Third Party Applications can implement Notification Services via a Technical User only. Notification Service specific roles/groups will not be explicitly available in Developer Cockpit Application for Third Party App implementation.
- Kindly contact the support team in order to generate a new Technical User.

For accessing this service you need to have the respective roles listed in Notification Service roles and scopes.

Basics

The notification service uses different controllers for the communication and information transmission. Each controller offers various tasks for message processing. The following list describes the content of each individual controller:

- Communication Channel: this resource gives information about available communication channels. It gets a list of all active communication types with the channel ID and the channel name such as e-mail and push notifications.
- Address Type: this resource lists all the address types supported by recipient service such as personal mail, office mail or push notification.
- Recipient: with this API controller, you can create a new recipient using the e-mail and phone details and manage the accounts by using an API call. You can execute a search by recipient name and get a recipient based on the recipient ID. Thymeleaf is used as an HTML template engine. Hence, templates are expected to use thymeleaf tags.

- Certificate Store: the Certificate Store controller manages the recipient's certification system. You can update, check, retrieve and delete a recipient's certificate.

- Template Param: this resource gives information about template parameters. You can use this API to get available template parameters for a requested template set.

- Template Manager: the Template Manager allows you to use templates for sending messages such as e-mail and push notification. You can also merge template parameters in an existing template. Via the API you can get a list of available templates and view the template details and the contents. Each template has an ID.

- Communication Category: the Communication Category controller manages the communication categories. Every category has an ID. You can create and delete the categories. The communication category controller allows you to manage your recipients into different categories. You can also unsubscribe the recipients from a category. For example, a technical support can create a category using recipients and template to define one communication category.

- Encryption Service: the encryption service controller encrypts the CcMail, e-mail and plain text. If any one of the recipients for a triggered message does not have the respective certificate available, the email notification is sent as unencrypted to all the recipients, irrespective of the availability of their certificates. PGP or S/MIME encrypted e-mails are signed with noreply@mindsphere.io. Users can download the public key of the PGP certificate from https://pgp.mit.edu/ for installing it in their e-mail client.

- Message Publisher: the Message Publisher controller is the basic component of the notification service. It publishes the messages to the queue for further processing and routing to the appropriate channel.

- Communication Service Audit: the Communication Service Audit controller saves the message in a database and logs the message information in an audit file. You can use the API to search for the stored audits in the database. The messages are stored with an audit log which would be available up to three months.

Features

The Notification Service exposes its API for realizing the following tasks:

- Digital Certificate Management: Upload/update/delete public certificates for the e-mail encryption.

- HTML template management: Use pre-defined HTML templates for the e-mail notifications. Upload different HTML templates and reuse them for the different notification channels.

- Configuration: Configure HTML template and recipient or recipient group. Use a unique configuration name for the target audience and reuse the configuration.

- Security: Use different e-mail encryption mechanisms, e.g. PGP or S/MIME. Encrypt the messages using different public certificate type of the intended recipients.

- Audit logs: Trace the history of the sent notifications via audit log.

Example Scenario

The manager of a wind farm wants to trigger a push notification every time the wind speed exceeds a certain level. The manager uses the API and connects the notification service with the aspect data of the wind turbine. At a wind speed of 8 km/h, the administrator of the wind turbine receives a notification. At a wind speed of 9 km/h, an additional notification is sent to a pre-defined service user list for further action.

APPENDIX E: Operations of Signal Calculation Service

One-argument mathematics

- Absolute: Absolute value of a numeric value
- Lg: Logarithm in base 10 value of a numeric value
- Ln: Logarithm in base e value of a numeric value
- Exp: Exponential value of a numeric value
- Cos: Cosine value of a numeric value
- ArcCos: Arc cosine value of a numeric value
- Cosh: Cosine hyperbolic value of a numeric value
- Sin: Sine value of a numeric value
- ArcSin: Arc sine value of a numeric value
- Sinh: Sine hyperbolic value of a numeric value
- Tan: Tangent value of a numeric value
- ArcTan: Arc tangent value of a numeric value
- Tanh: Tangent hyperbolic value of a numeric value
- Sqrt: Square root value of a numeric value
- Not: Negation value of a numeric value
- IsNull: Checks if null
- SignNegation: Changes the sign of a numeric value
- Two-arguments mathematics
- Add: Adds two numeric values
- Subtract: Subtracts two numeric values
- Multiply: Multiplies two numeric values
- Divide: Divides two numeric values

- Power: Raises the first Operand to the power of the second
- LessThan: Compares two numeric values
- GreaterThan: Compares two numeric values
- GreaterThanOrEquals: Compares two numeric values
- LessThanOrEquals: Compares two numeric values
- Equals: Checks if two values are equal
- NotEquals: Checks if two values are not equal
- And: Applies boolean logic on two values
- Or: Applies boolean logic on two values
- Multiple arguments mathematics
- Min: Compares multiple numeric values and generates minimum
- Max: Compares multiple numeric values and generates maximum
- Mean: Generates mean of multiple numeric values
- Median: Generates median of multiple numeric values
- Sum: Generates sum of multiple numeric values
- Filter
- Removes all entries of a time series where the selected boolean property is set to true.
- Merge
- Merges two or more time series. The output time series contains existing time stamps plus new "interpolated" time stamps, depending on the selected mode:
- Last-Non-Null
- If a non-synchronized time stamp is found, for each time series the last non-null value is taken for the new time stamp.
- Linear-Interpolation

- If a non-synchronized time stamp is found, for each time series the interpolated value between the latest two consecutive values is taken for the new time stamp.
- Resample
- Obtain interpolated values from existing time series data at specific and equidistant, user defined timestamps
- Constants
- Uses constants for applying other operations
- Constants
- Converts a features type into another type
- One-argument String
- Capitalize: Makes the first letter of a string uppercase
- LowerCase: Makes all letters of a string lowercase
- UpperCase: Makes all letters of a string uppercase
- Md5Checksum: Generates the MD5 hash of a string
- RemoveDiacritic: Deletes the diacritic characters from a string
- RemoveDuplicates: Deletes the duplicated characters from a string
- Reverse: Reverses the order of characters from a string
- ToEmpty: Converts the string to empty string
- Length: Returns the number of characters of a string.
- Two-arguments String
- CapitalizeBy: Converts the second operand to uppercase on the first operand
- RemoveChars: Removes the second parameter from the first one
- Concat: Concatenates two strings
- Compare: Compares two strings
- RegexMatcher: Checks if a string (first argument) respects a given regex (second parameters) and return a Boolean value.

- Count: Returns the number of appearances of the second parameter into the first operand
- IndexOf: Returns the index of the second argument into the first one
- LastIndexOfChar: Returns the last index of the second argument into the first one
- SubString: Returns a substring of first argument starting from the given index (second argument)
- Three-arguments String
- Replace: Searches on a string (first operand) a given string (second operand) and replaces it with a given string (third operand).
- RegexReplace: Searches on a string (first operand) a given regex (second operand) and replaces it with a given string (third operand).
- SubStringInterval: Returns a sub-string of a string (first operand) starting from the given start index (second operand) and end index (third operand).
- One-argument DateTime
- IsWeekend: Check if a DateTime is during a weekend.
- IsWorkDay: Check if a DateTime is on a work day.
- GetYears: Returns the number of years from a given DateTime.
- GetMonths: Returns the number of months from a given DateTime.
- GetDays: Returns the number of days from a given DateTime.
- GetHours: Returns the number of hours from a given DateTime.
- GetMinutes: Returns the number of minutes from a given DateTime.
- GetSeconds: Returns the number of seconds from a given DateTime.
- GetMilliseconds: Returns the number of milliseconds from a given DateTime.
- Two-arguments DateTime
- AddDateTime: The operation adds two DateTime values
- AddYears: Returns a new DateTime that adds the specified number of years to the value of the instance.

- AddMonths: Returns a new DateTime that adds the specified number of months to the value of the instance.

- AddDays: Returns a new DateTime that adds the specified number of days to the value of the instance.

- AddHours: Returns a new DateTime that adds the specified number of hours to the value of the instance.

- AddMinutes: Returns a new DateTime that adds the specified number of minutes to the value of the instance.

- AddSeconds: Returns a new DateTime that adds the specified number of seconds to the value of the instance.

- AddMilliseconds: Returns a new DateTime that adds the specified number of milliseconds to the value of the instance.

- SubtractDateTime: The operation subtracts two DateTime values and return milliseconds interval.

- SubtractYears: Returns a new DateTime that subtracts the specified number of years to the value of this instance.

- SubtractMonths: Returns a new DateTime that subtracts the specified number of months to the value of this instance.

- SubtractDays: Returns a new DateTime that subtracts the specified number of days to the value of this instance.

- SubtractHours: Returns a new DateTime that subtracts the specified number of hours to the value of this instance.

- SubtractMinutes: Returns a new DateTime that subtracts the specified number of minutes to the value of this instance.

- SubtractSeconds: Returns a new DateTime that subtracts the specified number of seconds to the value of this instance.

- SubtractMilliseconds: Returns a new DateTime that subtracts the specified number of milliseconds to the value of this instance.

- LessThanDateTime: The operations compares two DateTime values.

- GreaterThanDateTime: The operations compares two DateTime values.

- GreaterThanOrEqualsDateTime: The operations compares two DateTime values.
- EqualsDateTime: The operation checks if two DateTime values are or not equal
- NotEqualsDateTime: The operation checks if two DateTime values are or not equal
- Observation Windows and Aggregators
- Time window: Collects a consecutive sequence of entries from on a time interval and applies a list of aggregate functions
- Length window: Collects a consecutive sequence having a fix number of entries and applies a list of aggregate functions
- Delay
- Events based: Shifts to the right the events from the input with the specified number
- Time period based: Delays the time of the entries with the specified milliseconds received in the request
- Generator
- Continuous: Generates every given period of time a simple entry containing only the timestamp
- Temporal: Generates copies of the input entries based on the period parameter and number of events received as input

All operation inputs and outputs use one of the data types integer64, double or boolean. If one of the operands is Null, the result is Null.