MOdel based coNtrol framework for Site-wide OptimizatiON of data-intensive processes

D2.1 - MONSOON Platform Usage Scenarios

Deliverable ID	**D2.1**
Deliverable Title	**MONSOON Platform Usage Scenarios**
Work Package	**WP2 – Requirements Engineering and Reference Architecture**
Dissemination Level	**PUBLIC**
Version	**1.0**
Date	**2016-11-30**
Status	**Final Version**
Lead Editor	**Fraunhofer**
Main Contributors	Alexander Schneider (**Fraunhofer**) Sophie Fosson, Rosaria Rossini (**ISMB**) Peter Bednár (**TUK**) Dimos Ioannidis (**CERTH**)

Published by the MONSOON Consortium

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723650.

Document History

Version	Date	Author(s)	Description
0.1	2016-11-03	Fraunhofer	First Draft with TOC and Vision Scenario
0.2	2016-11-23	Fraunhofer	Added extended vision scenarios and common scenarios. Added specific scenarios for Aluminium and Plastics sector
0.9	2016-11-25	Fraunhofer, CAP, CERTH, ISMB, KIMW, TUK	Added chapters from CAP, CERTH, ISMB, KIMW and TUK with scenarios detailing specific aspects of using the MONSOON platform. Deliverable now ready for internal review.
1.0	2016-11-30	Fraunhofer	Final version of deliverable.

Internal Review History

Version	Review Date	Reviewed by	Summary of comments
0.9	2016-11-29	Sophie Fosson, Rosaria Rossini (ISMB)	Approved with minor comments and one change to the structure
0.9	2016-11-29	Nathalie Praizelin (CAP)	Approved with minor comments
0.9	2016-11-29	Muhammad Aamir (KIMW)	Approved with minor comments

Table of Contents

Executive Summary

The purpose of this deliverable is to document and describe a set of plausible usage scenarios in the year 2020 and beyond for the MONSOON platform in the two domains Aluminium and Plastics but should represent scenarios that are also common or at least applicable to other process industry sectors as well.

The process of Scenario Thinking (or Scenario Planning as it is sometimes called) is widely recognised as a tool for creating user requirements specifications under uncertainty. Chapter 2 explains the concepts of scenario thinking as it was used in the creation of this deliverable.

The vision scenario in Chapter 3 aims to provide a coherent, comprehensive, internally consistent description of plausible futures built on the imagined interaction of key trends for the process industry with respect to the model-based optimizations of data-intensive processes with the help of the MONSOON platform.

Furthermore, some scenarios that should be common to various process industry sectors that represent the MONSOON vision are also included in Chapter 3.

Chapter 4 contains several scenarios that are specific to either the Aluminium or Plastics domain and were developed based on user workshops and interviews.

Chapters 5-7 contain scenarios regarding different specific aspects of the MONSOON platform and give an overview that will be further extended during the project runtime.

For the future, we will continuously update and extend the scenarios also with the help of the Extended Stakeholder Group (ESG). The ESG consists of representatives of companies from the process industry as well as experts which have expertise in one or more technology fields related to the MONSOON platform.

The scenarios will be continuously updated and extended based on the needs and the evolution of the project.

1 Introduction

The purpose of this deliverable is to document and describe a set of plausible usage for the MONSOON platform in the two domains Aluminium and Plastics as well as scenarios that can be envisaged by other process industry sectors.

Creating scenarios of end user behaviour and interaction with platform functionality is a very useful instrument for identifying key technological, security and business drivers for future end user requirements. The scenarios will provide the framework for subsequent iterative requirements engineering phases.

Based on the scenarios and storylines, a systematic formalisation of all relevant user requirements and subsystems requirements will be derived.

The basis will be user-centric requirements originating from the ecosystems of the MONSOON industry partners. These include functional requirements, energy requirements and business requirements. The non-functional requirements will include requirements related to ethics, inclusion and data protection, quality of use, legal, etc.

The deliverable documents the work undertaken in task T2.1 Scenario Thinking and provides top-level user requirements in the form of a vision scenario of the future use of the MONSOON platform as well as more detailed scenarios either common or specific to the Aluminium and Plastics domains. The next step will be to produce the initial set of requirements based on user-centred workshops with representatives from the industry partners in MONSOON. This work will be conducted in task T2.3 Evolutionary Requirement Elicitation and Innovations.

2 Scenario Thinking Methodology

2.1 Introduction

The vision scenario aims to provide a coherent, comprehensive, internally consistent description of plausible futures built on the imagined interaction of key trends for the process industry with respect to the model-based optimizations of data-intensive processes with the help of the MONSOON platform.

Making reliable predictions about future user requirements calls for a great deal of certainty - an adequate level of knowledge and confidence in our assumptions about that knowledge. But defining user requirements today is extremely complex, as it is taking place in a fast-changing, information- and technology-driven environment. On their own, familiar planning and forecasting practices that have served us well in the past cannot deliver the insights and answers necessary in today's world of shifting values and policies, changing social structures and behaviour, which increasingly challenge predictions of how the future will look.

The process of Scenario Thinking (or Scenario Planning as it is sometimes called) is widely recognised as a tool for creating user requirements specifications under uncertainty.

Scenario Thinking is not about predicting the future; neither is it about choosing the best way forward, though it is indeed a powerful and invaluable tool to this end. Its primary value lies in the development of new skills for improving the definition and planning of user requirements.

Developing and deploying these skills enables us to transcend the specific or narrowly defined solution, to go beyond short-term or one-off successes and acquire a consistency and robustness in coherent long-term user scenarios. We come to know the right questions to ask and where to look for answers to open issues; how to recognise unique opportunities and choose the best way to go.

The first step in Scenario Thinking requires us to anchor ourselves securely in the present. When thinking about the future, we always do so within a context, and a starting place provides an opening array of ideas or facts, which in turn are related to some perception of a desired goal or objective for future user interaction.

As we convert this information into well-defined stories of possible future situations and our options for action in these, we unveil the inherent uncertainties that must be dealt with or overcome. An obvious fact often forgotten is that these uncertainties have been initiated by our original thinking, assumptions, omissions and commissions.

The quality and disposition of original input will strongly influence the flow of thought, handling of material and quality of output. To make the best use of scenarios, our intentions must be explicit and the issues or areas to test must be clearly identified.

The purpose of Scenario Thinking is to challenge the preconceived notions we may have of the future, allowing us to revise or revisit our accustomed approach. The process is intended to open the way one thinks about the future. Scenarios help to identify threats, recognise opportunities and make choices about strategically important issues. A scenario illuminates the possible, what might be. It prods one to do something slightly counterintuitive; to go beyond the known into the unknown, outside one´s area of expertise.

While reading the scenarios one should think about answers to such questions as:

- Is this even remotely possible?
- Would the world be a better place in this scenario?
- If one were a stakeholder in this scenario, what would we be doing differently?
- If one knew for sure that this scenario was to come true, what would we do now?

The Scenario Thinking process is designed to arrive at several parallel, co-existing hypotheses about the future. These variant hypotheses are given a concrete form, and stakeholders can visualise them because they are embedded in a story or a scenario. In turn this means that the same person can look at the scenario through different sets of glasses and see things from different perspectives.

2.2 Development of the scenarios

The scenarios were developed in a two-step process – first the research partners drafted the overall vision scenario and the initial context scenarios from their knowledge of the main areas of MONSOON. Those draft scenarios were then refined after having conducted user workshops and interviews with several industry partners from both domains.

3 MONSOON Vision Scenario and Common Scenarios

3.1 Main Vision Scenario

Michelle is a data scientist working for the company FDF from the process industry, an industry aiming at improvements in the efficient use of raw resources and energy. The company established goals to continuously improve process efficiency in a rapid and cost effective way. One key aspect to achieve these goals, is to have access to relevant data to achieve a clear understanding of the overall performance of the plant.

Michelle is responsible for the optimization of the plant's performance. Michelle's work is supported by a new platform: the Data Lab. This new tool consists of a "Data Storage and Analytics Platform", which is fed with data from different sources e.g. the sensors and actuators form the production units, process data and management data. Thus, the Data Lab provides analytics through which insights and trends of various operations of the plant can be uncovered.

Michelle develops models that result of the analysis and identification of patterns in complex data sets and correlations.

When Andrew, the floor manager, finds inefficiencies for example related to the levels of energy consumption in one area of the plant, he reports it back to Michelle, who integrates this information to his analysis. At this point, Michelle can create a new model to identify the cause of this abnormality. She accesses the Function Repository, where knowledge on similar problems and solutions from the past or similar events is stored (algorithms and control functions) to check for further support. In addition, Michelle takes advantage of the visualization features of the new tool to ease the interpretation of correlated historical and real-time data, which are useful to identify deviations from normal conditions, detect malfunctions and failures and identify areas for optimization.

Afterwards, Michelle simulates the model in iterative cycles and evaluates its performance and impact until she finds an optimal solution. This allows her to test the proposed solution before an actual deployment.

When Michelle considers the model appropriate for implementation based on the simulation results, the model is transferred to the "Real-time Operations Platform". Here, controls are made operational, integrated to the plant control infrastructure and all automatic system managing processes are updated.

Thus, when Andrew is back to his floor in the plant, he can already see that the parameters of the motors have changed in accordance with Michelle's model and the energy consumption of the machines he is responsible for is back to normal.

Michelle knows that it is important to feed knowledge back to the platform to support the search for an appropriate solution the next time a similar scenario occurs, or to make it available for cross-sectorial use cases. Therefore, the new algorithm, which she employed for the model is saved in the Data Lab Functions Repository and it is made available for use in future cases.

3.2 Common Context Scenarios

3.2.1 Advanced Analytics

The integration of the Data Lab to Michelle´s workflow has definitely had a positive impact in the quality of the results that she delivers and in the way she achieves her goals at work. The advanced data processing and machine learning capabilities that the tool provides allow Michelle to correlate enormous amounts of information and create complex algorithms.

Michelle remembers how it used to be in the past when the tools she employed to analyse data were extremely limiting. Back then, she had to rely only on some spreadsheets and a few functions to create her reports. This limited enormously the type of analysis she could make. Thanks to the new platform, Michelle feels empowered and has gained high confidence in the value of the results that she can deliver.

Before it was hard to identify models that were used in the past to apply them to new similar scenarios. Now, things are different, the job of the data scientist has become faster and more efficient thanks to the collection of information found in the functions repository fed by her colleagues and other experts in the field. Michelle can now reuse models and functions which gives her more time to focus on bringing her analysis to a new level by exploring new areas for optimisation.

Furthermore, ever since Michelle has gotten access to the Data Lab, she can develop models and functions that are more flexible. This is mainly since the Data Lab provides her access to enormous amounts of high quality data coming not only from different areas of the plant but also from different domains.

Michelle is now confident enough that the models she creates can adapt easier to the needs of the several layers of the process and the organization.

She is also quite satisfied with the intuitive collaboration features of the Data Lab. Thanks to them, she is now part of a collaborative work environment, where colleagues and other experts in the field, can share knowledge and work together towards the common objective of process optimisation. Thanks to this cross-domain collaboration, her analyses now are more rich, meaningful and informed.

3.2.2 Best practices in the plant

Andrew, floor manager, has always been aware that understanding all the different elements that influence the quality and efficiency of the production process in his domain is not an easy task. With all the number of factors that influence the production: from parameters in the machine and its correlation, origin of raw materials, temperature in the plant, etc. it is necessary to have a rather high level of expertise in the area in order identify the core of the problem and act in moments of crisis.

In the past, Andrew and his team used to struggle to take decisions on how to tackle malfunctions on the spot. Identifying the right combination of values in the parameters of the machines or what measures to take to solve an issue was not simple task. The workers used to make mistakes in moment of crisis due to the complexity of the scenarios since they did usually not have the resources needed to interpret the data given to take the precise measures.

Since the new platform has been installed in the plant, things have changed. The new tool can make clear and comprehensive recommendations in order to tackle problems on the spot. Andrew can now see alerts or recommendations on how to tackle issues in the monitor and forward them to the corresponding member of his team.

This has made the work in the plant way more efficient. The workers are not anymore expected to spend time figuring out how to solve a problem. On the contrary, they just must follow the set of instructions forwarded by their supervisor. Now the team is able to solve issues that are not too complex, right on the spot without the need to escalate the case and ask for advance technical support. Andrew is now confident that his team can react fast in an informed way. They can now put all their focus on achieving the targets needed and increasing the production according to the established goals.

4 MONSOON Context Scenarios – Aluminium Domain

4.1 Aluminium sector and the Data Lab

The company InPro is operating in the aluminium domain. They have established goals to continuously improve process efficiency in a collaborative and fast manner.

One of the actions they have taken, is the creation of the core expert team. There, a group of technology experts are in charge of solving specific problems in targeted site processes using high quality data and advanced analytical tools and models. The main tasks of the new team are:

1) Support direct requests from the site in order to help with crisis management
2) Look for long term optimization options by modelling studies
3) Develop predictive functions to be able to identify problems before they occur

In order to achieve their objectives, they make use of the Data Lab. Thanks to the Data Lab, they are able to:

- Analyse data on high levels to go cross-domain and cross area of the different smelters.
- Create models that are adapting to the needs of the several layers of the process and the organization.

4.2 Optimization by modelling studies

Joe, is a technology expert and a member of the core expert team. On a regular basis, he and his team perform site wide process performance reviews in order to identify areas of potential progress. They analyse historical data and Key Performance Indicators (KPI) and correlate these with real time data. This allows them to give recommendations for improvement and prioritize technology deployments in each one of the smelters through modelling studies.

During their last meeting, Joe and his team analysed the possibility of deploying operational changes in the reduction area of the plant to increase the overall productivity.

They usually discuss their ideas and recommendations with process coordinators in regular meetings and through the communication functionalities of the tool: close relationships with the sites is critical since the operational changes have to be analysed carefully by the experts in order to consider all factors that could be affected after the deployment.

Making use of the functions repository and the development container Joe and his team can analyse and test different algorithms in different ways:

- by benchmarking against similar production processes uploaded to the collaborative section by competitors or other industries.
- by comparing the current performance to previous data from the reduction site
- by correlating the real-time data with predictions from the simulation developed when they set up the new algorithm.

After exploring the results and interlinking different data sets, Joe has proved his assumption right. He has identified a model that could increase the overall productivity of the plant by changing only a few parameters of the motors.

Once the feasible model has been developed the changes take place through the support of the process manager and his team in the smelter. After a few weeks, the deployment proves already to be successful and the stakeholders in the plant already look forward to be advised again by Joe´s team on similar optimization options in the future.

4.3 Crisis management and direct demands from the site

Tom is a plant process manager and his main task is to support overall production process across the different areas of the plant.

Today, he attended an urgent meeting with the process coordinator of the carbon area.

Tom and his team have identified malfunctions in the motors of the paste plant, according to the analysis of the current situation, they realized that the actions needed are complex and ambitious. They consider that now they do not have the sufficient resources to tackle the issue.

For this reason, they decide to trigger a request for support from the core expert team which will help to make sure the smelter is meeting the targets on moments of crisis.

Joe, data analyst expert working as part of the core expert team responds immediately to Tom´s request. He makes use of the Data Lab, where he has access to big amounts of correlated information and knowledge gathered from the different production sites. Based on this data, he can propose an informed recommendation to tackle the crisis Tom and his team are encountering.

Joe, simulates possible solutions through different cycles of development and evaluation and compares results until the most suitable model is developed. Thanks to the simulation capabilities of the Data Lab, it is possible to perform an evaluation of the impact of the operational changes in the motors across different areas in the smelter. Joe can now be confident that the model he has created will not have a negative impact in any stage of the production process and, on the contrary, it will help to solve the issue the smelter is facing.

Once the development of the model has been completed, the dedicated new function is deployed in the paste plant motors in real time operation with the help of the process manager and his team. Tom and his team are glad to know that the crisis is gone and they can continue working in their daily activities.

5 MONSOON Context Scenarios – Plastics Domain

5.1 Problem identification with support of MONSOON

Georg is plant manager in a company, which produces plastic goods. When a new product is produced (which happens quite often, now that so many things can be individualized), the process is carefully planned by him and his colleagues. During the planning, they aim to have not only an efficient process but also a robust one with a big process window and a high-quality outcome.

Today Georg's day does not start very well because he got negative feedback from quality control referring to a new piece of plastic with a difficult form they just started to produce. The quality inspection needs up to 16 hours. That means that the company is producing not-sellable parts for quite some time without noticing. Georg reminds himself to talk to his boss again, about how much quicker they could react to problems if they had more sensors installed. Heat and especially pressure sensors are expensive, but on days like today, Georg is sure that the costs would amortize very quickly.

Even though there are not as many sensors in place as Georg considers that are needed, he is still very happy about the fact that they implemented the MONSOON platform last year. He knows that detecting the problem in the process is a lot easier now and so he goes straight to figure it out and starts the Data Lab software on his tablet computer.

He got information from quality control that not all the parallel produced pieces have the same, required weight. From experience, Georg knows that the most likely explanation is that the amount of plastic injected to the molds is variable. What Georg and his colleagues have to do now is to find the parameter(s) responsible for this variability. They will do empirical tests, where they change parameters and relate the outcome of the quality control of the so-produced pieces to the according parameter set of the process. Potential parameters are e.g. how the mold is moved by the machine, or the viscosity of the plastic (which e.g. is influenced by the room temperature during drying). Because there are a lot of parameters and potential interactions between them, which could lead to the bad result Georg learned about this morning, he enters all the setups and results from quality control into MONSOON's Data Lab. MONSOON does all the calculations and provides him with the best parameter set. In addition, MONSOON makes suggestions on which parameter changes could be done in a new test, in order to get results that are even more reliable and solutions that are more efficient.

While Georg lets the Data Lab do its calculations, he remembers the old times with a smile, when they still looked up potential solutions in static paper documentation which was limited in the amount of insights it could provide. The first improvement to the booklet was a software, which was able to correlate parameter settings with the outcomes.[1]

Still, this software was not able to handle multi-variate models. Since Georg works with MONSOON, the solutions really get the most out of the process by using more complex tests to identify the best parameter sets.

After the empirical testing support by the MONSOON platform is finished some days later (they always have to wait for the results of quality control after each test), Georg can inform the workers about the new parameters set which needs to be employed. Knowing that he has not to worry about this anymore, Georg calls his boss to get an appointment to remind him of the need of installing more sensors again.

5.2 History and Communication support with MONSOON

Klaus works at the same company as Georg as a plant worker. Today Klaus restarts the production of a new large batch. The customer from the automotive industry wants to be supplied with 1000 pieces every quarter. However, the restart of the process does not run well. Even though Klaus and his colleagues have mounted all the molds and rely on the planned process parameters like last time, they already know by visible inspection of the outcome that something goes wrong.

[1] e.g. https://www.kistler.com/de/de/produkte/produkte-nach-anwendung/spritzgiessen-prozessueberwachung-produkte/#como_neo_5887_a

Klaus opens MONSOON's Data Lab on the touch screen of his machine. He enters the key for the production piece in the process documentation database and after some filtering he sees, that last time they had similar problems. That puts Klaus in a good mood as he has good reasons to assume that this solution may work today as well and would keep them from a larger testing series to find the problem.

Before putting the solution recommended by MONSOON based on historical data into action, Klaus quickly checks that the required parameter changes are within the range he can take responsibility for independently in his position and does not have to ask for approval. To Klaus's delight, the change can quickly be employed and the outcome looks much more promising, though they still have to wait for the final feedback from quality inspection. Klaus uses the reporting field in the Data Lab view to keep the database up to data about today's deviations from the planned process.

The MONSOON platform will automatically decide whom to send this report (e.g. the planning team to get this feedback). This is a significant improvement compared to how it used to be before, when reporting issues and communicating changes was always done on paper which would generally get either lost or forgotten in a drawer.

Klaus is quite happy to work with MONSOON now. First, he and the other workers had strong objections because they used to handle most problems in their own way and were proud of that. Each worker used to bring his own parameter set on a USB stick, and plug it into the machines before his shift started. When management a) decided that the variability in the outcome was just not reasonable this way and b) wanted to increase efficiency by a learning platform, which kept track of changes in order to benefit from previous solutions, the workers were afraid that they would not be able to change anything on their machines without asking permission first. The idea, that all their mistakes would be tracked, seems to signify that their lives would become harder with the new software solution.

However, the workers are now more in line with the results from the MONSOON platform as they notice that solutions are faster at hand and that they still have a range for each parameter in which they can act based on their experience. Overall, the fact that the process planner and management are informed about important changes, gives Klaus the freedom to focus on his job and not worry about all this communication issues.

6 MONSOON Cross-Sectorial Use-Cases

This section presents some general scenarios that can be envisaged by different industrial domains. Indeed, the MONSOON platform combines known best practice methodologies for model based site-wide control and integrates them within the collaborative Cross-sectorial Data Lab methodology. The Data Lab supports the development of new dynamic model based multi-scale controls e.g. dynamically combining and chaining controls using data from different systems. Such approach enables innovative business cases solutions to other cross-sectorial domains beyond plastics and aluminium.

MONSOON cross-sectorial perspective is ensured during the project by the involvement of a group of external experts chosen in different domains, called here “External Stakeholder Group”.

The ESG as an external and essential part of the MONSOON cross-sectorial concept and consists of representatives of companies from the process industry as well as experts in different domains. The MONSOON ESG will be composed by representatives from multiple domains actively involved in co-design workshops and evaluation, in order to overcome traditional barriers between sectors and to stimulate cross-fertilization of experiences.

6.1 Modelling work-flows and semantics scenario

Ann is a data scientist working for the process industry company. She is now working on the new project for improving of production processes in her company using the data analytics methods. In order to better understand particular industry domain and collected data, she is using Semantic modelling tool from Data Lab platform. This tool provides web-based user interface, so she opened the browser and logged into the tool. The work in the tool is organized in projects and she can open a new project from the list of existing templates or start from scratch. She created a new blank project. Project is divided into four parts – Process, Data, KPIs and Predictive Functions. At first, she created a semantic process model using the graphical editor.

Semantic process model notation specifies how the overall production process and its workflow is divided to phases. For each phase, she can specify data produced or consumed in the particular phase. Input (consumed) data are representing input parameters (such as material properties, quantities etc.) and control actions/signals. Output (produced) data are representing various monitored measurements or observed diagnostics signals.

All specified data elements can be further described in the Data part of the modelling project. Here she can display data elements in two views – logical view and physical view. Logical view provides list of data elements with the type and short description. It was possible to simply filter-out elements of particular type or related to the particular process phase. In physical view, Ann can open Data Storage file browser where she can specify mapping between the logical data elements and already uploaded data files and physical data columns.

Additionally, to Data and Process model, Ann can specify list of general Key Performance Indicators such as energy consumption or environmental impact and map these indicators to the measured/observed data elements. From the semantic model, it was now possible to identify which data elements are the most critical for KPIs optimization. In the last part of the semantic model, Ann selected some of the critical data elements as the targets for the new predictive functions.

Similarly, to Data modelling, modelling of Predictive functions is divided to logical and physical view. In logical view, Ann specified for each new function logical input and target data attributes and analytical method (e.g. classification, regression, clustering, anomaly detection etc.). Later, when the data will be processed and she will build predictive function in the Data Lab Analytics Platform, she will map logical description to the physical predictive model exported and stored in the Data Storage file. Part of the export will be also the model quality report, which will provide performance statistics (e.g. model accuracy, number of false/true positive/negative signals etc.). After the importing of this report, Semantic Modelling tool will map performance statistics to KPIs, so it will be possible to directly evaluate how the performance of the predictive model influence overall quality of the production process.

Building of predictive functions in Data Lab Analytics Platform is the iterative and interactive processes consisting of multiple steps such as data pre-processing, application of data mining algorithms and statistical validation of functions. Ann can use Semantic modelling tool to generate project plan for the whole data mining process. For example, from the semantic mappings between the logical and physical data elements, Semantic modelling tool can automatically infer and suggest which data transformations and pre-processing steps is required for the building of the particular predictive function. Similarly, it can suggest additional data elements relevant for the particular KPIs, which could be included as the attributes of the predictive model.

After couple of revisions, Ann exported the logical views of the semantic model as the new template for the company's sector for the future projects.

6.2 Predictive functions scenario

Matthew is a data scientist working at a company from the process industry. His company utilizes data-driven techniques and methods from the field of predictive analysis and applies them in several parts of their production processes. Their strategy targets to exploit patterns and to identify opportunities to make the production processes more efficient on distribution of raw materials and the energy consumption. Also aims to improve of the overall quality of the final product and to decrease production losses.

Matthew is responsible for analysing a vast amount of available data gathered from multiple sensors that monitor the production in the plant. He is in close collaboration with Peter, process coordinator in one area of the company, who is responsible of making sure the performance of the machines in his area is maintained to an optimum level.

Matthew applies statistical functions to visualize the linear trend profile of the time series from different production parameters. All available data, historical and real-time, that he needs in order to develop his models is acquired from the "Data Lab". Matthew uses feature extraction methods on all the available features and parameters aiming to find those that provide the most relevant information from the original data, so as to feed them in his predictive models. He knows that existing data can be used to create predictive algorithms, meaning that a ground truth is available (similar problems and solutions that happened in the past). He also applies anomaly detection techniques (i.e. local outlier factor), machine learning and deep learning techniques aiming to provide early warnings or even predictive signals for global or individual anomalies throughout the production process.

Based on the data analysis, he generates and incorporates into the Data Lab robust alert functions that trigger several notices, such as notification, warnings and alerts based on situations that happened each time. The predictive functions developed by Mathew, offer Peter the ability to face and handle potential issues as he is provided with predictive warnings in his computer or in the monitors across the plant. Along with the notice (Warning/Alert), the system provides context information about the features of the problem such as information of the characteristics of the machine used, a brief analysis of internal and external parameters performance curves and raw materials supply curves. The implementation of these statistical functions helps Peter to identify deviations from normal trends in an early stage in order to achieve an optimum site-wide scheduling and further more to prevent malfunctions which could lead to a possible future downtime of the whole system. Peter is aware that identifying trends in an early stage is crucial in order to prevent issues that that could lead to a big crisis later on.

Thanks to the warning, Peter can manage in a satisfactory way these undesirable conditions and avoid crisis effects for the company, such as financial costs.

6.3 Training Scenario

Mathew and Ann are data scientists and have limited knowledge of production process and related variables. To start with a new project, data scientists need to understand the production process, the data acquisition and signal conditioning performed by sensors on production machines and tools. Their training starts with the physical phenomenon or physical property of an object that must be measured depending on the type of production process, machines and tools. This physical property could be, e.g. temperature change, pressure, force applied, displacement, and so on. A sensor is a device capable of converting a physical property into a measurable electrical signal. This electrical signal can be, e.g., voltage, current, a variation in resistance or capacity. The ability of a DACS system to measure the distinct phenomena depends of the transductors to convert the signals coming from the physical variables to hardware data. Signal conditioning is usually necessary, because the transductors do not normally provide the adequate signals for the hardware used for acquisition. The signal could require the following operations:

- Amplification: It is the most common type of signal conditioning. To get the most precision available the entrance signal must be amplified so that its maximum level coincides with the maximum entrance value the ADC can read.
- Excitation: The conditioning of the signal sometimes generates excitation in some transductors, such as extensiometric gauges, thermistors or RTDs, in some cases because the transductors need it or because its connection configuration (Wheatstone bridge).
- Filtering: To eliminate noise or unwanted signals from the one being measured, for example the 50/06HZ of the electrical network. High frequency signals, such as vibration signals, need an antialiasing filtering, to eliminate the signals with higher frequency than the maximum frequency measurable, to avoid superimposing signals generating errors.
- Multiplexing: Commutation of the entrances of the converter, so that with only one converter it is possible to measure different entrance channels.
- Isolation: Electrical isolation between transductor and the PC acquiring signals, to protect the acquisition system from current peaks that can damage it.
- Linearization: Many transductors, such as thermocouples, present a non-lineal response to lineal changes in the parameters being measured. Although linearization can be performed by software using numeric methods, it is a better idea to correct this using extern circuitry.

Depending on the complexity of the DACS system, Ann and Mathew needs the information that must be extracted and the type of control to perform over the process, to implement different approaches further for data processing to develop complex algorithms and predictive modelling of the production process.

6.4 System interoperability and mapping scenario

Mario works in an industry and is in charge of controlling the process phases. He manually collects all the data from the machines. After a while, looking at the historical data, Mario realizes that something has gone wrong in one specific phase of the process. Then, he needs to go deeper into details as well as analyse all historical and real-time data in order to understand what causes such an anomaly.

Mario knows the MONSOON platform and the peculiarity of supporting the use of harmonized data models and common pattern for data integration, data mapping, inter-linking field systems with site-wide models.

Integrating the MONSOON platform into his plant is easy. In fact, the IT infrastructure of MONSOON provides a lower level called 'Real-time Plant Operations Platform' that implements techniques to ensure dependable communications under data-intensive conditions, including methodologies for failure awareness and self-recovery. Furthermore, the mapping is ensuring by the cross-sectorial actions engaged by MOONSON. In particular, data models used to drive the MONSOON analysis can be adjusted to different domains. In fact, the different heterogeneous data that are coming from the shop floor are processing and labelled in this stage in order to allow a better storage and analysis.

The MONSOON Architecture includes a dedicated environment, namely the Cross Sectorial Data Lab. The Data Lab supports the development of new dynamic model based multi-scale controls e.g. dynamically combining and chaining controls using data from different systems. It includes a Big Data Storage and Analytics Platform, supporting a scalable storage solution where data from the local plant/site and possibly a multitude of production sites is stored as well as enabling big data analytics.

Once the platform is installed and connected, Mario can easily manage the life cycle through the MONSOON system and operate on the plant level in order to find and likely correct possible errors.

6.5 Development environment Scenario

Andrew, the floor manager of the FDF Company, is facing a big challenge. The company has decided to launch a new type of product, requiring new production equipment. Since he has the MONSOON Platform, he is using predictive functions for product quality on all equipment of his workshop, and this brings real value, as he has made significant savings. Andrew has now to integrate these new equipment and process into the MONSOON model. Andrew exposes his issue to Ann, the Data Scientist, and to Tom and Michael, the business experts. They communicate with the supplier of the new equipment, in order to identify the available data from this equipment, and work closely together to define how to exploit them. Using the MONSOON platform, and the Cross Sectorial Data Lab structure, they have identified that another company, GMH industry, using the same type of equipment, has already built a predictive function for product quality, quite similar to the one he is looking at, that could be a good starting point. This function is available in the Functions repository of the Data Lab Environment. After analysis, Ann communicates to Robert, the IT expert, the production data required from the new equipment to build the new predictive function. Robert has now to put in place the interfaces to retrieve the data into the MONSOON Platform. He remembers how tricky this type of task was before the arrival of the MONSOON platform.

Now, using the Real Time communication framework and the Virtual process industries resources adaptation provided by MONSOON, he has access to powerful tools, allowing him to make new connectors in a few hours, and providing all the communication and QoS monitoring tools. Very quickly, the Big Data storage and analytics platform is populated with the new production data, allowing Ann to test the initial predictive function provided by GMH industry, and to customize it for FDF by machine learning based on their own production data. After validation by Andrew on Data Lab environment, Robert has the green light to deploy the new function on site. Using the Run time container provided by the real-time plant operations platform of MONSOON, the deployment is done in one click. Although he has been using MONSOON for several months, Andrew is always surprised to see how quickly new functions are made available in production.

7 Conclusions

In this deliverable, a set of scenarios have been collected and described, based on the first round of interviews and user workshops. In the future, we will continuously update and extend the scenarios, based on the evolution of the project and with the help of the Extended Stakeholder Group (ESG).