FACTLOG
www.factlog.eu

ENERGY-AWARE FACTORY ANALYTICS FOR PROCESS INDUSTRIES

Deliverable D1.2

Cognitive Factory Framework

Version Version 1.0	**Lead Partner** NISSA
Date 31/10/2020	**Project Name** FACTLOG – Energy-aware Factory Analytics for Process Industries

FACTLOG has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 869951

Call Identifier
H2020-NMBP-SPIRE-2019

Topic
DT-SPIRE-06-2019 - Digital technologies for improved performance in cognitive production plants

Project Reference
869951

Start date
November 1st, 2019

Type of Action
IA – Innovation Action

Duration
42 Months

Dissemination Level

X	**PU**	Public
	CO	Confidential, restricted under conditions set out in the Grant Agreement
	CI	Classified, information as referred in the Commission Decision 2001/844/EC

Disclaimer

This document reflects the opinion of the authors only.

While the information contained herein is believed to be accurate, neither the FACTLOG consortium as a whole, nor any of its members, their officers, employees or agents make no warranty that this material is capable of use, or that use of the information is free from risk and accept no liability for loss or damage suffered by any person in respect of any inaccuracy or omission.

Executive Summary

The main goal of this deliverable is to provide the foundation for the realization of the cognition-driven solutions in pilots. Since this type of systems is a novel one, this deliverable also explains some of basic concepts, such as cognition and cognition process. It appears that a metaphor of human cognition, as taken from psychology, can be a very suitable basis for developing the concept for an efficient design of cognition-driven industrial systems, which we consider as Cognitive Factory Framework (CFF).

The main advantage of CFF is that it brings together two perspectives. One is the way in which human cognition deals with new information/situations, esp. in the case of unknowns (it is not known how to react based on existing models and past data). Another perspective is the industry-process oriented one: how a process behaves under variations (internal, external), i.e. how (un)stable are the process performances (KPIs) in such situations. The main goal of CFF is to make industry processes able to deal with variations efficiently, based on the analogy to the human cognition. More precisely, we define cognition process consisting of four basic phases:

1. Detect variations
2. Understand root causes of variations
3. Understand the impact of variations
4. Find optimal reaction

We also define the roles of four basic technologies, data analytics, knowledge graph, process modelling and simulations and optimization in these phases, illustrating that the envisioned cognition-driven processing is feasible.

In addition, we provide a deep analysis of all pilots regarding the realization of Cognitive Factory Framework, demonstrating that CFF is general enough to be applied in various use cases / scenarios.

This deliverable will serve as a kind of guideline for the realization of the specific components in other WPs, which combined together (based on the architecture described in D1.3) will provide desired cognition-driven solutions in all pilots.

Revision History

Revision	Date	Description	Organisation
0.1	01/09/2020	Table of content	NISSA
0.2	15/09/2020	Initial input: technologies	Technology partners
0.3	29/09/2020	Initial input: pilots	Technology support partners
0.4	02/10/2020	Refined input: technologies	Technology partners
0.5	09/10/2020	Refined input: pilots	Technology support partners
0.6	13/10/2020	Cognition part	NISSA
0.7	17/10/2020	Refinement	NISSA
0.8	23/10/2020	Prepared for the review	NISSA
0.9	28/10/2020	Internal peer review	Qlector, C2K
1.0	31/10/2020	Ready for submission	NISSA

Contributors

Organisation	Author	E-Mail
NISSA	Nenad Stojanovic	nenad.stojanovic@nissatech.com
AUEB	Yiannis Mourtos	mourtos@aueb.gr
AUEB	Georgios Zois	georzois@aueb.gr
AEUB	Panos Repousis	prepousi@aueb.gr
AEUB	Gregory Kasapidis	gkasapidis@aueb.gr
AUEB	Stavros Lounis	slounis@aueb.gr
UNIPI	Pavlos Eirinakis	pavlose@unipi.gr
UNPI	Konstantinos Kaparis	k.kaparis@uom.edu.gr
UNIPI	Gregory Koronakos	gregkoron@gmail.com
MAG	Kostas Kalaboukas	kostas.kalaboukas@maggioli.it
EPFL	Jinzhi Lu	jinzhi.lu@epfl.ch
EPFL	Dimitrios Kyritsis	dimitris.kiritsis@epfl.ch
BRC	Alexander Adams	alexander.adams@brc.ltd.uk
C2K	Kevin Greening	kgreening@control2k.co.uk
JSI	Aljaž Košmerlj	aljaz.kosmerlj@ijs.si
Qlector	Jože Rožanec	joze.rozanec@qlector.com
Qlector	Klemen Kenda	klemen.kenda@qlector.com
TUC	George Arampatzis	garampatzis@pem.tuc.gr
TUC	George Tsinarakis	tsinar@dpem.tuc.gr
DOMINA	Caterina Calefato	caterina.calefato@domina-biella.it
DOMINA	Marco Vallini	marco.vallini@domina-biella.it
PIA	Alessandro Canepa	alessandro.canepa@piacenza1733.it
CONT	Alin Popa	alin.3.popa@continental-corporation.com
SIMAVI	Andreea Paunescu	andreea.paunescu@siveco.ro

Organisation	Author	E-Mail
SIMAVI	Radu Popescu	radu.popescu@siveco.ro

Table of Contents

List of Figures

1 Introduction

This deliverable represents the foundation for the usage of the cognition, as a decision-making method/process, in resolving challenges defined in pilots, which will result in developing cognition-driven solutions.

There are many definitions of cognition, mainly inspired by the human cognition and its role in human reaction on the changes happening around her/him. Usually, this reaction starts with the perception (with all senses) of signals from the environment and continues with their interpretation on different levels, with the main goal to properly understand new situation and react correspondingly.

This process is used as an inspiration for the cognition process presented in this deliverable, showing that complex problems can be solved by mimicking human behaviour and reasoning. The main point is that cognition process starts with some changes in the environment (internal, external) and a computing system should be able to perform the monitoring and the detection of changes in a very efficient way.

We argue that one of the most important contributions of this work is related to that challenge. Namely, we structured the cognition process in phases which focus on particular challenges, like detecting variations, discovering the causes, understanding the impact, optimizing the reaction, which can be resolved with selected technologies: data analytics, knowledge graphs, process modelling and simulation, optimization. In that way we developed a very powerful framework, we call Cognitive Factory Framework, which is the one of the central concepts of this deliverable.

In order to validate this rather theoretical approach, we applied the described cognition process in pilots, with the goal to demonstrate that all cognition-phases can be mapped on selected processes in pilots. Moreover, each phase can be realized with one or more technologies, making our approach very robust and efficient.

However, the work in this deliverable has served as a conceptualization phase in realizing envisioned cognition-driven solutions. We argue that this foundation is very sound and hope that in the scope of WP2-WP5 it will be completely implemented and in WP7 fully applied in all pilots.

1.1 Purpose and Scope

The main purpose of this deliverable is to report the work related to the development of the Cognitive Factory Framework, done in the scope of the task T1.4 Cognitive architecture.

This work is the foundation for the work in other work packages (WP2-WP5) on the further development of the components required for the full realization of the cognition process and its application in pilots.

This deliverable doesn’t pretend to define a new theory about cognition, but rather to define a practical view on the usage of advanced processing that mimics the way in which humans resolve problems (esp. those dealing with uncertainties) on the resolution of the problems in the industry.

1.2 Relation with other Deliverables

There are two deliverables relevant to this deliverable:

- D2.1 – which provides more details about the role of data analytics (including the cognition process)
- D1.3 – which provides a technical view on the interaction between particular technologies (components), which is very important for the realization of the cognition process

As mentioned above, this deliverable will influence the deliverables in WP2-WP5.

IMPORTANT: some parts related to the descriptions of pilots and technologies (e.g. introductions in pilots) and are taken from other deliverables in order to make this deliverable self-contained.

1.3 Structure of the Document

This deliverable is structured in the following way:

Section 2 introduces the interpretation of the cognition process from the non-technical (psychology) point of view and its implication on IT-based decision-making processes

Section 3 explains the Cognitive Factory Framework and details the role of each technology in it

Section 4 provides detailed view on the usage of Cognitive Factory Framework in pilots

Section 5 provides concluding remarks

2 Cognition

2.1 Foundation

In this section we provide some basic information about cognition, as considered in psychology[1], since one of our main goals is to be inspired by the principles of human cognition for the resolution of the challenging scenarios provided in the pilot cases.

Cognition derives from the Latin verb *cognoscere*, which means “get to know”. This means that cognition focuses on knowledge, albeit not as a static substance or “thing”, but as a process. More generally, when we speak about cognition we are focusing on the mind as an information processor, i.e. a system that acquires, uses and transforms information.

The human nervous system is capable of handling endless streams of information. The senses serve as the interface between the mind and the external environment, receiving stimuli and translating it into nervous impulses that are transmitted to the brain. The brain then processes this information and uses the relevant pieces to create thoughts, which can then be expressed through language or stored in memory for future use. To make this process more complex, the brain does not gather information from external environments only. When thoughts are formed, the brain also pulls information from emotions and memories (cf. Figure 1). Emotion and memory are powerful influences on both our thoughts and behaviors.

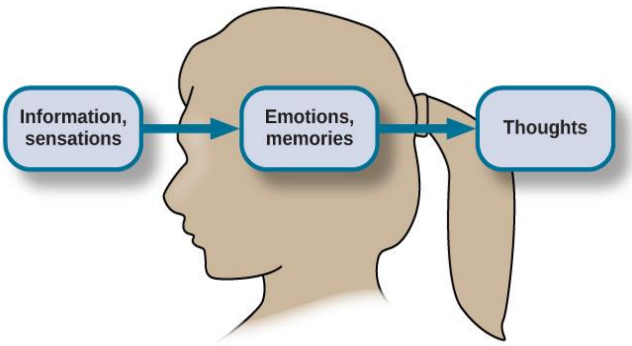


Figure 1: Human cognition as a decision support mechanism

Therefore, the process consists of following steps:

1. The senses serve as the interface between the mind and the external environment, receiving stimuli and translating it into nervous impulses that are transmitted to the brain.
2. The brain then processes this information and uses the relevant pieces which are held in working memory, and later expressed through language or stored in memory for future use.
3. To make this process more complex, the brain does not gather information from external environments only.

[1] https://louis.oercommons.org/courseware/module/134/student/?task=3#:~:text=The%20senses%20serve%20as%20the,are%20transmitted%20to%20the%20brain.&text=When%20thoughts%20are%20formed%2C%20the,emotions%20and%20memories%20(Figure).

4. When thoughts are formed, the brain also pulls information from emotions and memories. Emotion and memory are powerful influences on both our thoughts and behaviors.

In the following figures (2 – 4), we provide the illustrations of the role of cognition in understanding real-world inputs (new situations).

It is important to mention that these layers correspond to the modern, well accepted, theory from cognitive psychology[2] about human cognition, optimized for dealing with complex structures, enabling deep understanding of the real-time situation (monitoring).

Figure 2 depicts the situations that no changes are happening in the real world. In that case, the cognition is in standby mode since there is no need for any interpretation. Consequently, there are no actions (triggers) in the decision-making system.

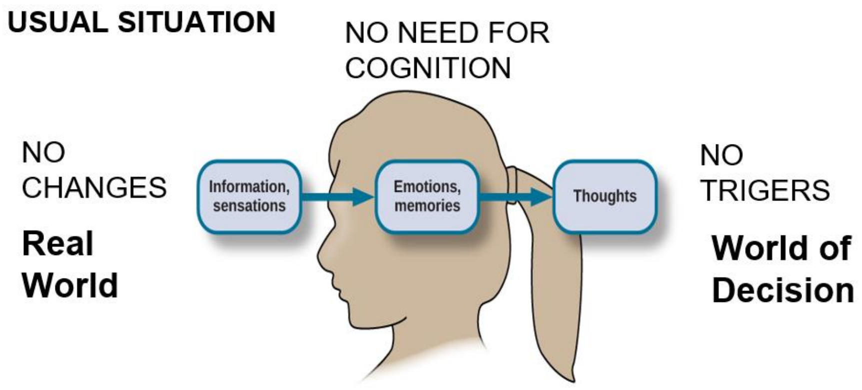


Figure 2: Human cognition in standby mode

Figure 3 illustrates the situations when changes in the environment are happening, but they are known in the real world. In that case, there is no need for understanding the changes, but only for their recognition which can be done without activating the entire cognition process.

Regarding human psychology, this processing corresponds to System 1, so called **Fast Thinking**[3] which operates automatically and quickly, with little or no effort and no sense of voluntary control.

[2] https://www.scientificamerican.com/article/kahneman-excerpt-thinking-fast-and-slow/
[3] https://scottbarrykaufman.com/wp-content/uploads/2014/04/dual-process-theory-Evans_Stanovich_PoPS13.pdf

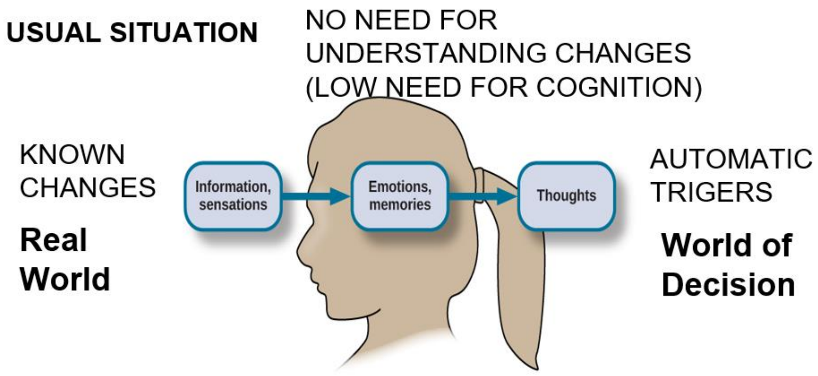


Figure 3: Human cognition in action – processing usual situation

Figure 4 depicts the situation when an unknown change is happening, requiring additional processing for understanding causes and impacts, i.e. going through the cognition process.

Regarding human psychology, this processing corresponds to System 2 (see footnote for System 1), so called **Slow Thinking**, which allocates attention to the effortful mental activities that demand it, including complex computations.

Moreover, it enables complex and efficient processing of complex situations, by creating digital models of their behaviour from sensed data, supporting timely and precise decision making.

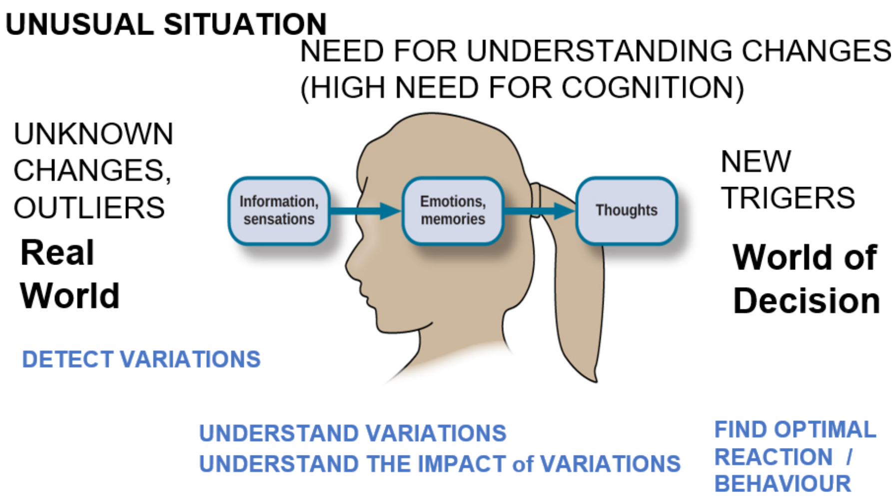


Figure 4: Human cognition in action – processing unknown situation

It is clear that cognition requires a set of activities (we call phases) which enables the rather complex processing. As illustrated in Figure 4, there are four basic phases:

1. Detect variations
2. Understand root causes of variations
3. Understand the impact of variations
4. Find optimal reaction

These phases are elaborated in the Section 3.

2.2 Cognition as a (Decision Support) Process

In this section we provide an interpretation of cognition as a problem-solving methodology which can support decision making processes (in the industry) efficient.

As already described, ccognition is related to the possibility for **explaining the behavior of a system** (with the goal to optimize it). It is **based on the various models** learned from past data and on numerical models, which describe the **usual/expected behavior** of the system, that means an industrial system is operating in a **stable** way (with less variations).

However, the systems operate in heterogenous and sometimes harsh environments, so that one of the main challenges is to detect variations. Variation can reflect unusual behavior (various ad-hoc events, outliers) or known anomalies.

The next step is to **understand the variations**, esp. what is the root cause of variations and how the variations can be explained (why did they happen).

After having clarified the causes of variations, the next challenge is to **estimate/predict the impact** of the detected variations in order to understand how critical they are.

Finally, cognition supports the **optimization of the behavior** of the system.

In order to realize this process, a corresponding technology stack is required. We have selected four technologies which are:

- Data analytics, mainly focused on detecting variations and predicting/estimating their impact.
- Knowledge graph, focusing on the explanation of the causes of the variations and providing domain knowledge for understanding the root causes.
- Process modeling and simulation, mainly focused on modeling the process knowledge required for understanding the impact of variations, as well as the simulation for doing various what-if analyses related to the situation of interest
- Optimization, focusing on having optimal reactions on the situations the system should react to, based on collected/available information

2.3 Cognition vs AI

In order to clarify the position of cognition regarding the AI, in this section we provide a short discussion about both technologies.

Cognition (represented in IT as the concept of Cognitive computing) and AI are technologies that rely on data to make decisions. The main difference is in the way in which the human is interacting with the technologies. In the following table, we present a comparison taken from literature[4].

Cognitive Computing	Artificial Intelligence
Cognitive Computing focuses on **mimicking human behavior** and reasoning to solve complex problems.	AI **augments human thinking** to solve complex problems. It focuses on providing accurate results.
It **simulates** human thought processes to find solutions to complex problems.	AI **finds patterns** to learn or reveal hidden information and find solutions.
They simply **supplement information** for humans to make decisions.	AI is responsible for **making decisions** on their own minimizing the role of humans.
It is mostly used in sectors like **customer service, health care, industries**, etc.	It is mostly used in **finance, security, healthcare, retail, manufacturing**, etc.

Thus, we can say, cognitive computing helps us make smarter decisions on our own leveraging the machines. Whereas, AI is rooted in the idea that machines can make better decisions on our behalf.

[4] https://laptrinhx.com/what-is-cognitive-ai-is-it-the-future-2900285987/

3 Cognitive Factory Framework

3.1 Introduction

In this section we provide a general view on the Cognitive Factory Framework (CFF). Its main role is to provide a setting for the development of various use cases organized in different industrial environments. Therefore, CFF can be seen as a template/framework to be followed when realizing use cases by following the concept of **cognition** (so called cognition-driven use cases).

Cognitive Factory Framework is based on the results from D1.1, esp. related to Operational scenario which is illustrated in following figure (source: D1.1, Section 2).

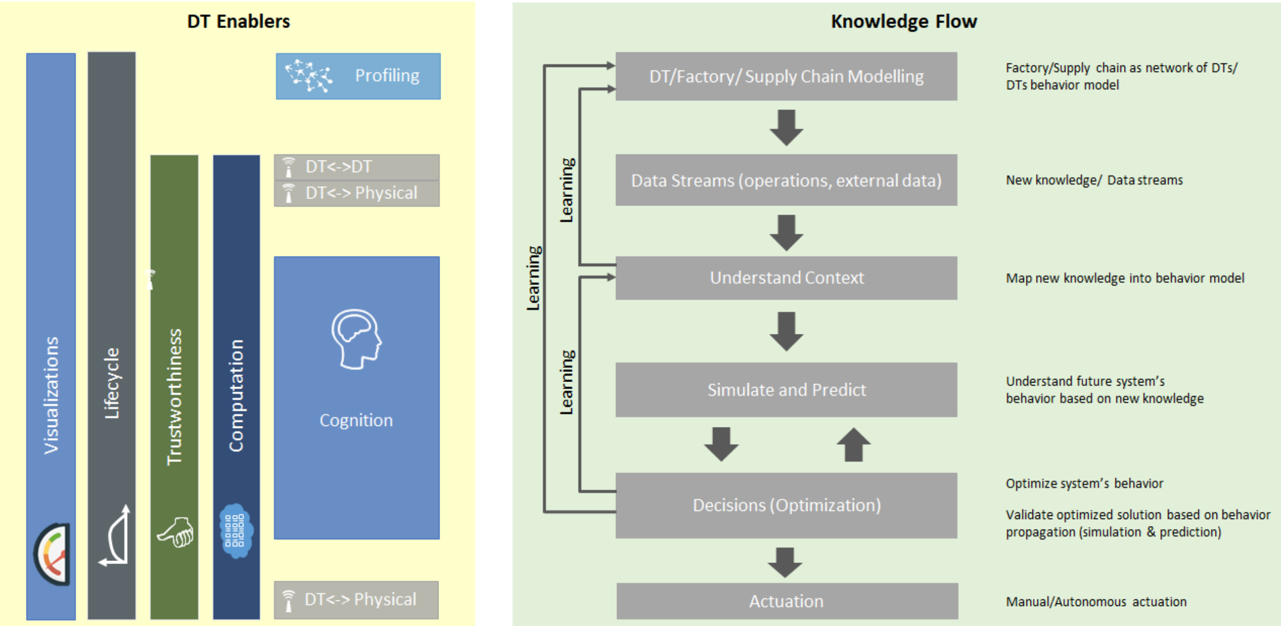


Figure 5: FACTLOG Operational Scenario (source: D1.1, Section 2)

The aim of the Cognitive Factory Framework is to enable realization of the presented knowledge flow in the operational scenario through the **cognition process**, as we structured in previous section.

One of the interpretations of the presented knowledge flow is that it follows the line of actions required for supporting processing of new knowledge, including the situations where the new knowledge/data cannot be "easily" understood (e.g. the root-cause, or the impact of a new situation) and some advanced processing, like simulations of future situations are needed. Therefore, in a nutshell, CFF is a cognition process used as a general processing pipeline for realizing cognition-driven use cases / scenarios, based on the Knowledge flow from the Operational Scenario.

This mapping is the basis for describing the Cognitive Factory Framework, as presented in the following section.

3.2 The Framework

Based on the general elaboration in Section 2, we see the cognition as an enabler for the challenging use cases, where the new data leads to unusual situations. More precisely, cognition supports:

- detection of (NEW) problems/anomalies,
    - sensed through observing real-time data
    - incl. previously unknown problems
        - Not in existing models
        - Not in past data
- their full understanding
    - root-cause
    - impact/consequence
- optimal reaction to resolve the potential issues

It is clear that these steps should be supported by corresponding technologies which realize some of the challenges and can be combined for the implementation of the whole use case.

The following figure depicts the mapping between the knowledge flow and cognition process, illustrating the steps which should be performed in each of the pilots.

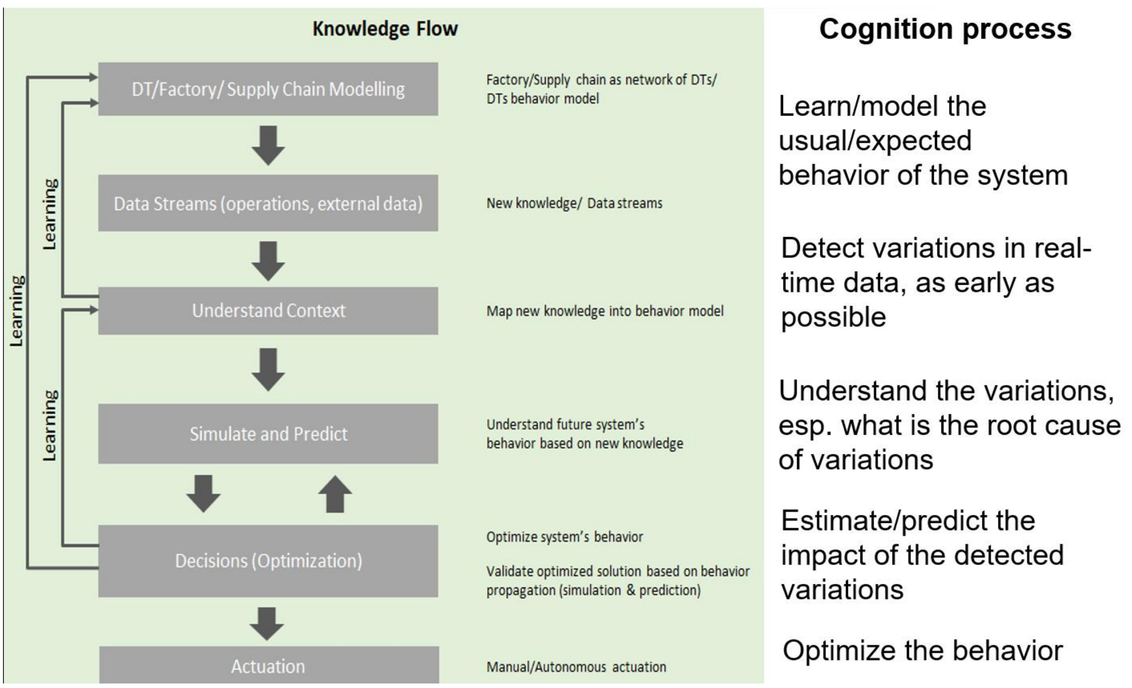


Figure 6: Mapping Knowledge Flow (Operational Scenario) to the Cognition Process

In the following figure we present the full mapping between the Cognition, Technologies and the Process-oriented activities, which is the basis for the realization of the CFF.

Cognition	Technologies	Process
Learn/model the usual/expected behavior of the system	Data analytics	Learn the anomaly-free process model (STABLE – voice of data)
Observe variations - detect variations in real-time data, as early as possible	Data analytics	Monitor the process Detect INSTABILITIES in the process (real-time)
Understand the variations, esp. what is root cause of variations	Knowledge graphs	Understand INSTABILITIES in the process
Estimate/predict the impact of the detected variations	Simulations	Propagate instabilities
Optimize the behavior	Optimization	

Figure 7: Total Mapping between Cognition, Technologies and the Process

Therefore, there are four technologies, grouped around three paradigms (Data-driven analysis, Model-driven analysis, Optimization) which are driving CFF, as presented in the following figure.

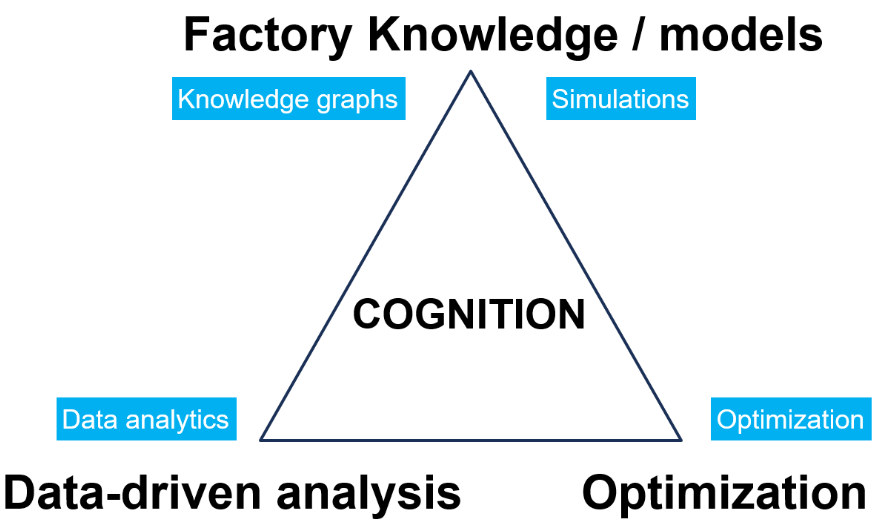


Figure 8: Technological background for Cognitive Factory Framework

3.3 The role of technologies in CFF

We define cognition process through 4 phases:

Phase 1: Detect variations in the behaviour, as early as possible

Phase 2: Understand the root-cause of variations

Phase 3: Estimate the impact of variations

Phase 4: Support the optimization of the behaviour

In the following text we provide a structured description of each of these phases, by emphasizing the roles of the technologies, which will be detailed in the rest of this section.

1. DETECT VARIATIONS

- **Goal**: Find variations by observing data
    - Known: already defined situations
        - Given by expert, e.g. described as CEP (Complex Event Processing) patterns
        - Learned from data
    - Unknown: unusual situations
- **Role in Cognition**: trigger the Cognition process
- **Methods**:
    - Validate data against a model
        - Data-driven models (**data analytics learns models from data**)
        - CEP (models are predefined usually by an expert)
        - Statistical models: (Statistical) Process Control (**variation detection methods**)

2. UNDERSTAND VARIATIONS

- **Goal**: provide additional context/knowledge for problem (variation) analysis and support root cause analysis
- **Role in Cognition**: contextualization of the problem

Methods:

- Root-cause analysis
- Using process models
- Knowledge graphs

3. UNDERSTANDING THE IMPACT of PROBLEM

- **Goal**: support understanding what is the impact of the problem
- **Role in Cognition**: understanding when (if) to react on detected problem
- **Methods**:
  - Analytics, Simulations, Optimization
  - Through the **simulation and prediction services** (root-cause analysis), we can propagate (predict) the behavior of the production line DT (Digital Twin), considering the particular risk of failure of the machine DT.

4. OPTIMIZE BEHAVIOUR

- **Goal**: support the analysis for which changes are required
- **Role in Cognition**: understanding how to react
- **Methods**:
  - Analytics, Simulations, Optimization
  - **Optimization services** will make an initial suggestion for the best point at which to make the change and the time estimations required. **Optimization services** get the input from the **process models** on how the process works and other metrics

3.4 Data Analytics

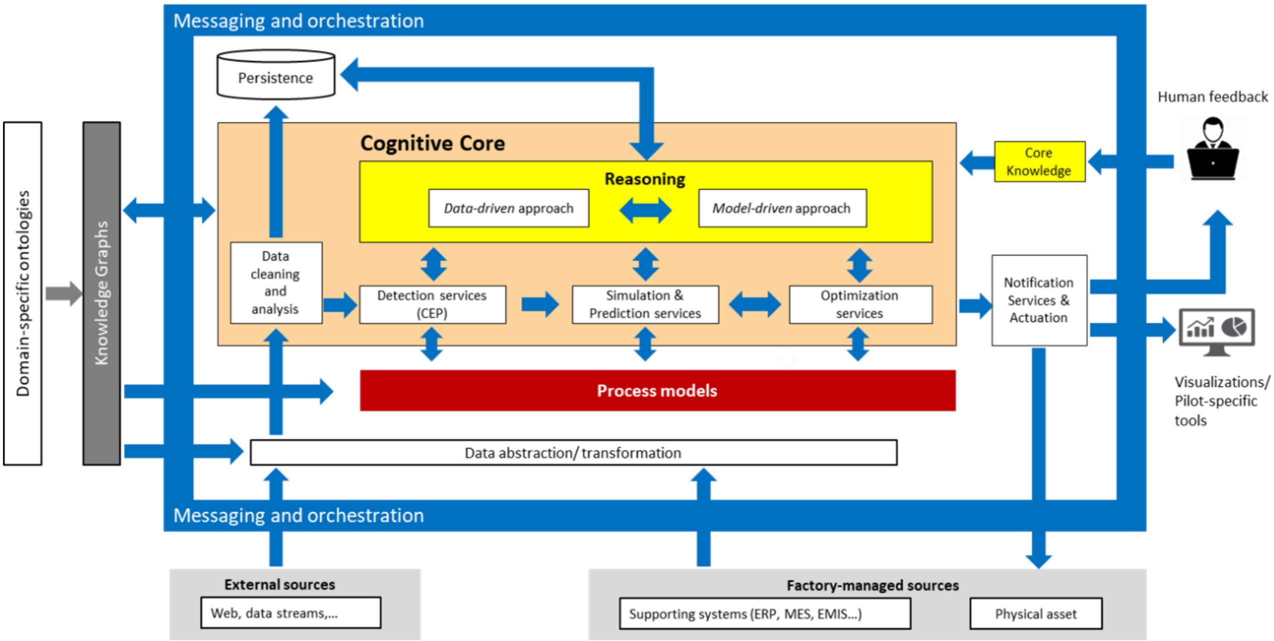


Figure 9: FACTLOG cognitive framework

The analytics system plays a fundamental role, providing insights into incoming and existing data. These insights may be useful to detect and identify patterns of pieces of data relevant to the use cases we operate within the framework, and provide information that may drive new understandings of the use cases. Data analytics supports the process of building theoretical models, models based on expert knowledge and historical data. Data analytics is therefore considered a fundamental building block of digital twins, on top of which we build digital twin cognitive processes. Analytics serve the purpose of understanding typical operation scenarios to detect and understand variations, their impact, and how given settings can be optimized based on observed changes.

To have a sense of everyday operation scenarios, we must know the data velocity, data lifecycle, typical values, and data distributions at different levels. Different granularity levels are required to model the manufacturing processes at individual, aggregate, or high-level stages.

In standard operation scenarios, models predict possible outcomes. If historical data is available, machine learning models can be built, providing them a set of examples from which they can learn to predict the target variable based on the values of variables of interest. Examples of context variables in a demand forecasting setting can be the prices of raw materials required to manufacture the product, economic variables such as GDP growth, or employment rates, which influence demand. We make predictions on a novel set of variable values, which we may not have seen before. The model issues a prediction based on patterns it learned from past data. Some algorithms also provide a measure of uncertainty along with the prediction for each case. Predicted values can be categorical (e.g., "the machine will break down") or numerical (e.g., "we expect the machine to break down in half a day").

Anomalies detection serves multiple purposes. It helps, in the first place, understanding of events that do not comply with typical operational scenarios. As such, we must address them to understand the root causes of the detected anomaly and assess potential impact in the manufacturing process. Second, if the anomalies persist and affect a meaningful time window for some prediction model, it can affect prediction outcomes. In such a case, proper alerts need to be issued to the users, to avoid eroding confidence in model predictions under typical operational scenarios.

We can build digital twins for manufacturing processes aggregating digital twins that model underlying processes. This aggregation reflects into analytics and machine learning models as well, where models from bottom stages provide useful inputs to the model at an aggregate level. An example of such aggregation is the aggregation of different production line stages into a production line digital twin. Such an aggregated digital twin allows us to simulate the production process and predict how many products will be manufactured, as well as to identify issues such as scrap above the expected level or production delays.

By following the Monte Carlo principle and running several such simulations by adding small perturbations in the inputs/outputs and settings to consider possible errors and noise, we can determine the process's most likely outcomes. Having this prediction computed in advance, we can compare it to the actual plant's readings and raise an anomaly alert when they deviate significantly from the predicted values. By following the deviations back through the system, we can highlight the most likely root-causes of the anomaly.

Anomalies can also be detected by observing the historical data and identifying the typical states of the observed manufacturing system. Using clustering algorithms on the historical sensor values, we can identify the system's typical states and its transitions between them. When the system deviates from these states or the transitions between them, the system should raise an anomaly alert.

Cognitive twins (WP3)

Analytics services are one of the building blocks for cognitive twins. Cognitive twins aim to relate current and potential events that occur at a given point in the factory to other components to obtain a comprehensive understanding of their impact on the manufacturing process. An example could be how the system reacts to a forecast informing on potential delays on stock replenishment. The event should be related to manufacturing processes that use these raw materials to understand how particular products manufactured will be affected. Then it will re-schedule production processes and reorganise employees, as well as recommend what actions to take regarding deliveries: a different transport may be required to deliver in time to the client, in order to compensate for manufacturing delays, and still honour original contract times.

Knowledge graphs and process modelling (WP4)

Algorithm results require context for a correct interpretation. We store data related to this context into a knowledge graph. The knowledge graph is not only a data store: it models and provides information regarding semantic relationships between pieces of data. As such, and given some algorithm results, it allows for the retrieval of semantically meaningful context for that result, to provide a correct interpretation. An example is that given some demand forecast, the model may inform that GDP growth over the last year and crude oil prices over the last three months influenced demand. Such information can be retrieved from the knowledge graph and presented to the user. This system can complement this insight retrieving information regarding GDP growth and crude oil prices for the months of interest. It may also provide a broader context, e.g., the evolution of GDP growth over the last years and oil prices over the last months, and expected future values obtained from projections published by some trustworthy organism.

Optimization (WP5)

Optimization algorithms attempt to find the best solution given specific criteria, searching for the problem with the areas they address. In the context of manufacturing, the algorithms can be used to find the best production schedule given specific live information from the shopfloor and the availability and status of machines. The analytics module can provide the required values from a given setting and historical data when they are otherwise not clear in such a context. The optimization algorithms may use analytics data during operation to make decisions and thus ensure they achieve the best results efficiently.

Technical Design

A detailed description of the technical design is presented in deliverable 2.1, to which we refer the reader for an in-depth understanding of the analytics module. In this section, we summarize the most important aspects of the analytics module.

FACTLOG technical design contemplates a loosely coupled architecture. In order to avoid tight coupling, we use a messaging bus and define a REST API interface. Both provide a uniform communication interface to underlying services and enable other users and services to consume them based on required resources or expected functionalities. This way, all services are abstracted from the specific underlying architecture, services arrangement, and infrastructure required to scale them. It also provides the advantage of allowing a smooth and transparent interaction with external services. This aspect is of great importance for FACTLOG architecture adoption since supporting a hybrid configuration allows integrating other service providers' technology stacks transparently. Such cases would be for interested clients or companies, who may incorporate FACTLOG as a part of their product while still developing their services that complement or enhance FACTLOG functionalities.

The messaging bus allows exchanging data among multiple modules, following a particular message convention. Messages are published to the messaging bus, reaching interested subscribers, such as the persistence and other services. The persistence service stores consumed data into appropriate databases based on the type of data obtained. This data can be later accessed for analytics purposes and replay events when required.

The message bus allows multiple services to process the same data, even simultaneously, if required, for different purposes. When doing so, streaming and batch processing can be applied, depending on the use case requirements. This abstraction also enables proper decoupling from user interfaces, allowing not only exposure of functionality through a web application but also to build multiple tools, such as command-line interfaces (CLI).

FACTLOG services use a comprehensive set of tools to achieve its purpose. We use three kinds of models from the machine learning perspective: batch and streaming machine learning models, and probabilistic models.

We train batch machine learning algorithms on historical data which are then deployed into production to make predictions using new data that becomes available over time. These algorithms may grow stale over time. Thus, it is essential to monitor their performance and changes over data distribution regarding the original data we trained them on. When staleness is detected, a new model shall be trained and deployed. Among commonly used batch algorithms we can mention logistic regression, SVM, RNN, and LSTM, which are implemented in libraries such as QMiner and scikit-learn.

We train Streaming machine learning algorithms over a stream of data, such as sensor information. They are designed to see the data only once and adapt to seen information one sample at-a-time. Theoretically, their performance shall be close to batch machine learning algorithms, while in practice, they usually have a slightly lower performance. On the other hand, they use a limited amount of memory and computational resources, and we may train them in real-time on incoming data. Some streaming algorithms can detect changes in data distribution and adapt accordingly to it. Among commonly used streaming algorithms we can mention Hoeffding trees and FIMT-DD, which are implemented in libraries such as QMiner and scikit-multiflow.

Probabilistic models consider data distributions and simulate possible outcomes given those distributions. By doing so, they provide accurate estimates of potential outcomes and associated uncertainties as well. One such widely used algorithms is the Monte Carlo simulation. Another tool we will use to this purpose is the Probabilistic Soft Logic, , which using first-order logic and probabilistic graphical models, allows us to model probabilistic and relational domains.

Qlector LEAP provides an intelligent platform capable of ingesting data, running different machine learning and probabilistic models, as well as correlate outcomes to events based on encoded knowledge.

For a detailed description of tools and algorithms, we refer the reader to deliverable 2.1.

Ideally, the role of the Analytics component in the cognition process is as follows:

1. Gain an understanding of the use case processes and build data-driven models of their dynamics.
2. For each model identify the data features with informational value and perform appropriate data cleaning and preparation.
3. Analyse variations in the stream of data from the sensors in the use case and detect anomalies critical to the use case operation.
4. Analyse the possible sources of the anomalies to identify the most likely root cause querying the knowledge graph and process models for guidance.
5. Build generative models of use case elements and machine components from historical data for use in simulation.

3.5 Process Modelling / Simulations

Process & Simulation Model (PSM) denotes a generic model with all related methods, algorithms, mechanisms, services and tools it directly uses, integrated into an overall modeling application or platform. In any specialized (e.g. FACTLOG use cases) model, these methods, algorithms and mechanisms do not change. It is just the process model (digital) representation per se that changes. PSM interconnects and interoperates with external AI tools, Optimization tools, Analytics tools, etc. These may also change from subclass to subclass etc. according to specific application needs, however PSM itself does not change from application case to case.

This digital model core may have two components: (a) a knowledge engineering component (typically a knowledge base, graph or network of some type) that represents formal system knowledge and data in the form of rules, relations, associations and predicates; (b) a dynamic operational model of the production system, usually (but not always) in the form of some dynamic Petri-net or process-flow model. This will also require complementary services from a number of other components built either on top of this Digital Core (i.e. taking advantage of its services) or as information gateways into and out of it (providing services to it or from it to man-machine and machine-machine interfaces). Such essential components are: (a) a real-time monitoring and analytics platform, (b) analytics tools, (c) optimization tools, (d) machine learning and inference tools, (e) production system management decision support.

Ideally, the role of the central (production system) modeling component, can be defined in relation to the cognition process, as follows:

1. Create a **dynamic digital shadow** of the physical system, providing its accurate image of its operation at any time, updated by getting real-time monitoring data from the analytics platform; The dynamic image can be also used by DSSs and man-machine interfaces for system monitoring, control and management.
2. Accurately model all production **processes** as well as all entities along the product's value chain and all inputs and outputs, providing a comprehensive **systemic view** and **operational statistics**;
3. Use KPIs, dynamic-LCA and business rules and benchmarks to **assess** the performance of the entire system and each component; also compare estimated and real-time data for **self-assessment**;
4. Provide **systemic knowledge** (entities, relations, material flows, process states, performance) and operational data to the knowledge engineering component (e.g. Knowledge Graph) and the machine-learning tools; while using the stored knowledge to make the model more intelligent;
5. Provide support for root-cause, risk analysis and hypothesis testing to **AI problem inference** tools;
6. Provide use case models and data to support **optimization** algorithms and tools;
7. Provide base model and support for demand-supply and other **prediction** tools or in-build routines;
8. Support system **adaptability** by building, running and assessing system adaptation scenarios.

3.6 Optimization

Cognition as a process that includes the detection of variations in a given ongoing behaviour, followed by the ability to understand these variations, as well as their impact can hugely benefit from optimization, as can the overall end-response to issues that have occurred. The overall FACTLOG Cognitive Framework is designed and developed in that direction. As industry progresses and we are directed towards the Cognitive Factory, optimization as part of cognition (or following cognition) can be seen as the mechanism that is utilized as the response deriving mechanism to the identified (or predicted) problem(s). In particular optimization in the scope of the cognitive factory can be seen as a means to improve the operations of the factory including the production optimization (e.g. increase throughput and/or eliminate wastage) or to enable asset optimization (e.g. predictive maintenance through the prediction of an upcoming asset failure or the prescription of maintenance strategies) among others [1]. The aforementioned are two of the business values brought forth initially by the introduction of cognition in an IoT setting followed by the optimization, similar to the scope of the Cognitive Framework in the FACTLOG project. Additionally, and taking under consideration the ability for cognition through the Cognitive Twins (of machines, processes etc. in the project) as enablers that materialize Cognitive Manufacturing, the 2018 WMF Report [2] mentions four lines of work needed, two of which are captured by Cognitive Twins: 'Hyper-Connected Intelligent Machines' and 'AI-Driven Cognitive Operations'. The third one, named 'Smart Optimisation of Resources', is not currently treated, at least not explicitly and on that account the Cognitive Framework of FactLog is introducing optimization (giving rise to the Enhanced Cognitive Twin) in order to facilitate decision-making at a narrow or a broad scope [3].

This introduction of optimization in the core of the Cognitive Framework is conducted through the Optimization toolkit. The Optimization Toolkit, in the context of the FACTLOG cognitive framework denotes the tools, methods, services and algorithms utilized to derive optimal output per use case relevant to the needs of the pilots. As in the Cognitive Framework the overall goal is to identify and understand situations that are (or will be) problematic, in most cases upon identifying and understanding of problematic situations in the production process, optimization is triggered to create a new schedule that caters for the observed problematic situations. Additionally, to the aforementioned, when there is normal operation and no problematic situation is presented, optimization is utilized in the context of the Cognitive Framework to produce the respective optimal solutions under normal ongoing situations. Briefly discussed optimization in the case of Tupras aims to optimize the return to on-specification production of LPG following an identified problematic situation, optimization in the case of BRC aims to provide schedules of production that take under consideration the machines involved in the process and the same stands for the PIA case where scheduling of production of new fabrics occurs taking under consideration the involved machines (e.g looms) all whilst receiving information from all other modules pertinent to cognition.

In order to do so Optimization has different components: (a) a problem specification component that represents the part of the Optimization toolkit where the problems are formulated relevant to the use cases, (b) a model initializator where the selected problem is now fed with required data stemming from within and outside the Cognitive Framework, (c) a solver component that utilizes different algorithms and case-specific solution methods that are executed after being configured (and utilizing data from other modules) respectively for all cases and (d) a results component that interfaces the results to the FACTLOG project. Lastly (e) a component responsible for the access control and interconnection to all other components within the cognitive framework.

Ideally, the role of the Optimization toolkit component can be defined in relation to the cognition process as follows:

1. Understand the optimization problem per use case.
2. Utilize the systemic view, operational statistics, use case models and data provided by the process models as input to the optimization problem.
3. Utilize information relevant to operational status and data provided by the CEP Services, Simulation and Prediction Services as input to the optimization problem
4. Provide per case the optimized solution to the given problems (e.g. Production Scheduling, Settings for In-specs production of LPG).

3.7 FACTLOG Standardization and Ontology

FACTLOG standardization and ontology are the basic specification for the FACTLOG architecture. They are used to define the unified terminologies and conceptual entities across the pilots, domains, digital twins, tools and platforms in the FACTLOG project. The ontology is the basis to construct the knowledge graph models which are used to support cognition implementation. The ontology is designed for system and digital twin aspects based on systems engineering standards including ISO 42010 and ISO 23247. Moreover, the ontology entities are also defined according to the existing ontology specifications including IOF and BFO.

Systems thinking will be used to identify the FACTLOG terminology and conceptual entities for cognition whose workflow is demonstrated as follow:

- Identify requirements for cognitions in the FACTLOG pilots: Cognition entities and their interrelationships are identified to capture the key concepts for ontology formalisms based on ISO 42010 and ISO 23247.
- Define use cases: the cognition is defined as general concepts to support advanced decision-makings in FACTLOG. However, the scope of manufacturing is wide including different use cases and scenarios. This step is designed to identify the ontology domain concepts and their interrelationships for specific use cases.
- Analyze the ontology using reference specifications: based on the system boundary of ACTs and use cases, IOF and BFO concepts are analyzed to provide an initial ontology development plan.
- Formalize ontology concepts using reference specifications: based on IOF and BFO, the ontologies are defined.
- Implement KG: based on the ontology identified from the previous step, KGs can be developed.
- Evaluation of ontology logicality and KG completeness: to evaluate the ontology and the KG implementation, we verify them for completeness and logicality.
- Develop cognition tool prototype: development of cognition tool prototypes based on proposed ontology and implemented KG.
- Prototype evaluation: after a thorough evaluation of the cognition tool prototype, a new round of ontology design is started.

3.8 Knowledge Graph Modelling

Knowledge Graph Model (KGM) denotes a generic ontology representation and description with all related product and equipment elements. Moreover, in order to construct cognitive enhanced twins, AI methods, algorithms, mechanisms, services and tools, are also described directly and integrated into an overall modeling application or platform. In any specialized FACTLOG use cases, the knowledge graph models require a unified and high-level abstract ontology definition in order to describe the related products, methods, algorithms and mechanisms in a standardized way. In the whole FACTLOG platform, the knowledge graph models are used to describe the domain specific knowledge, platform services and their interrelationships. KGM interconnects and interoperates with external AI tools, Optimization tools, Analytics tools, data visualization tools, etc. These tools are required to develop the data interfaces based on the developed ontology for importing and exporting knowledge graph models. According to specific cognition needs, the knowledge graph models can be developed using the unified ontology case by case.

The FACTLOG cognitive model includes three components: (a) knowledge graph models to describe domain specific knowledge and data with their interrelationships within the form of subjectives, objectives and predicates; (b) a dynamic process model of the production system, based on dynamic Petri-net or process-flow model; (c) data-driven AI models based on historical data. This will also require complementary cognition APIs for creating and implementing the cognitive models and provide cognition services. Thus, the essential elements and components in the FACTLOG platform related to knowledge graph models includes: 1) Ontology definition standardizations; 2) ontology modeling tool; 3) ontology verification tool; 4) KGM interfaces to FACTLOG platform; 5) KGM transformer in FACTLOG platform; 6) Reasoning component for KGM; 7) KGM interrupter.

Ideally, the role of the KGM, can be defined in relation to the cognition process, as follows:

1. Define the unified ontologies, including physical systems, FACTLOG services and FACTLOG platform, providing its meta-definitions for the nature of the products and cognitions.
2. Create a topological description of the physical systems and construct the basic for cognitions.
3. Describe operational process model structures and the interrelationships between them and simulation results.
4. Provide domain specific knowledge (entities, relations, material flows, process states, performance) and operational data in the knowledge graph models for the machine-learning tools;
5. Provide support for reasoning, risk analysis and hypothesis testing to optimization algorithms and tools;
6. Provide basic information and support for demand-supply and other prediction tools or in-built routines;
7. Support system formalisms by analyzing and defining system use case scenarios, historical data, real-time data and simulations.

3.9 Digital Twin Context

Digital Twins (DTs) in the underlying modelling entity where all the cognition/ services operate. More particularly, we consider that all information collection, analysis and optimization refer to the digital twin of the process, workstation of factory itself. In FactLog, the system starts with the modelling of process units, workstations, process lines and other factory assets either as atomic DTs (e.g. a process unit) or a network of DTs (e.g. process line).

In line with the operational framework we can say that the DTs will have the following cognition capabilities:

a) Able to **self-learn**, and thus to effectively detect and react to anomalies and disruptions but also to opportunities that may arise,

b) enjoy a local or global view of operations and (**aggregation/ disaggregation**) – thus being part of different hierarchical levels (see Figure 2).

c) are capable for short-, mid- and long-term **reasoning and optimization.**

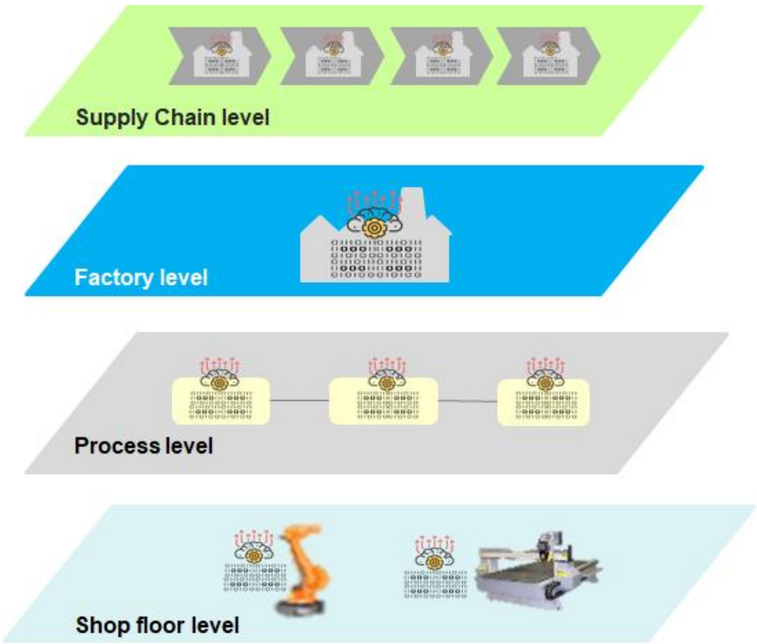


Figure 10: Different hierarchical levels of DTs

All the above capabilities are realized through a Digital Twin Platform (to be implemented) which – in line with the Standard ISO 23247 will monitor the DTs with the following information:

- Static (profile information)
- Status information (availability, real-time operation performance….)
- Context info (2d/3d models, specs, …)
- Communication interfaces (messaging and interfaces: DT <-> Physical, DT<->DTs)
- Relationships (parent, child, in a process, …)
- Operational logic (any s/w function that characterizes the behavior of the physical asset, including cognition services).

This means that the DTs as S/W components will model all assets and be able to connect with different services supporting the cognition process.

4 Cognitive Factory Framework in Pilots

4.1 Pilot BRC

4.1.1 Brief introduction

BRC manufactures bespoke products for the construction industry with a lead-time of 5-7 days where each batch is unique and can be up to 2 tons of steel in one product batch. Under BS8666 these can be in the form of simple straight bar, “U” shaped bars to complicated 99 shape codes where it could be 3D shapes. The process is to cut and shape from stock lengths of straight or coiled rebar and go through the flow process, which is shown below:

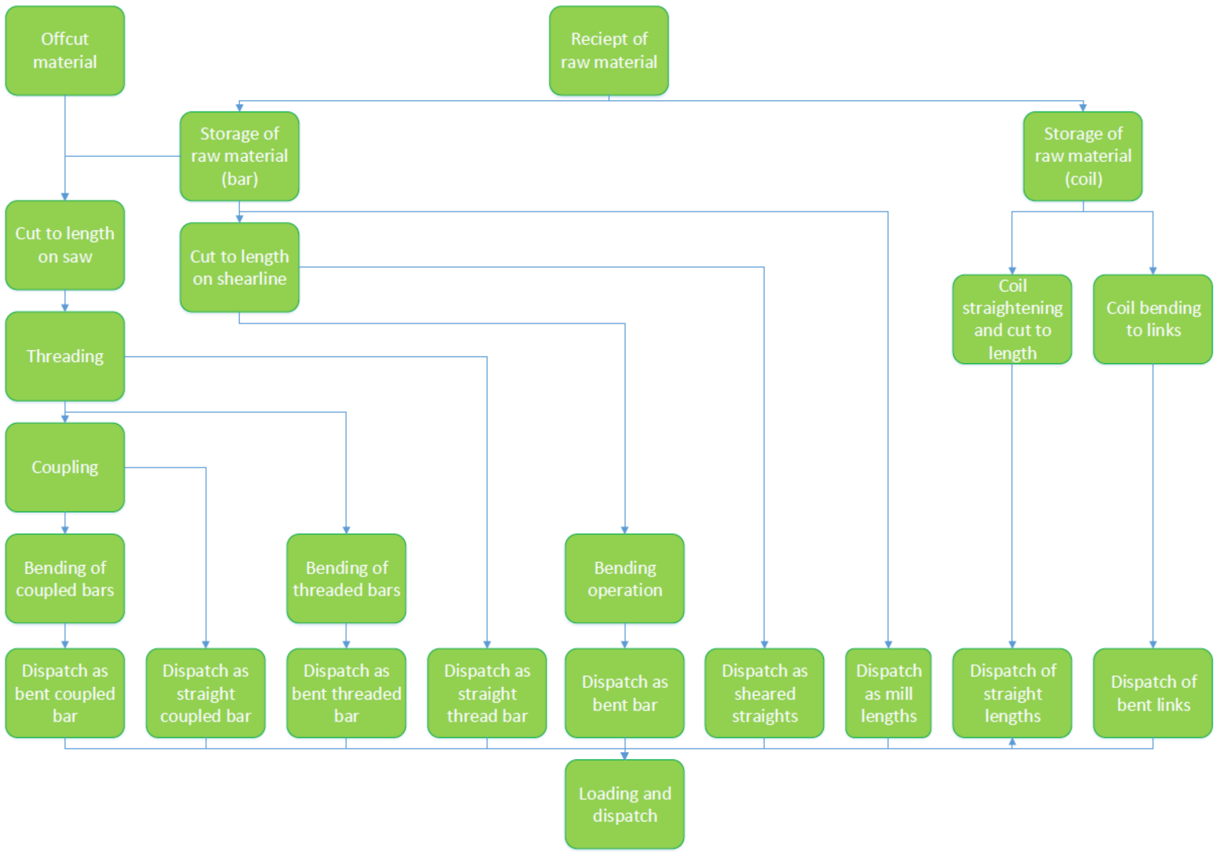


Figure 11: BRC Flowchart

The main KPI’s in the business are TPMH (Tonnes per man hour), Tonnage, operational costs and energy consumption. TPMH and tonnage go hand in hand with TPMH being efficiency targets and tonnage being the output of the business. The Newport business unit has put out 1686 tonnes over the last 6 months at an average of 0.76 Tonnes per man hour. At peaks it can reach up to around 2000 tonnes and hit a TPMH of 0.9, we look to hit this consistently through efficiency improvements but due to legacy machinery and current technology better overview is needed. Through efficiency improvement comes overall operational cost improvement and energy efficiency hence the advancement in this would greatly benefit in main business KPI’s and allow the business to improve environmentally. The key factors currently causing challenges for the business is the use of historical data and no live view status of the shopfloor, Personnel informing maintenance of breakdowns and storage of bar within the factory. These affect main processes areas logistics (crane movements), production process (Shearing, Manual bending and De-coil) and Maintenance.

4.1.2 Framework description

4.1.3 Cognition cycle

The cognition cycle for the project relies on a large number of data sets most of which are collected manually. The overall goal of the cognitive framework when applied to the BRC case is to increase production efficiency and hence productivity and reduce energy usage per tonne of product produced. To achieve this the main aim of the FACTLOG project is to improve the production process using a cognitive framework to provide best practice production scheduling. To implement this most effectively other systems have to be taken into account which will also use cognition to streamline their operation and provide data for the higher order cognitive systems to allow changes in system states to dynamically alter the optimisation processes.

4.1.3.1 Process cognition cycle

The goal of the cognitive framework in the BRC pilot case is to optimize and increase the production by managing the following aspects:

- Orders will be prioritized on receipt depending on delivery requirements, raw material and machine availability
- To optimize the production flow and loading schedules to trailers
- Schedule maintenance based on a stochastic model of machine performance to predict failures so they can be dynamically integrated into production schedules.

A provisional factory workflow for a new order is currently generated by the MES system based upon the following order requirements

1) Bar diameter
2) Bar length
3) Bar quantity
4) Bending segments
5) Vertical bends
6) Shape code
7) Machine speed for straights and bends

The provisional production workflow is then manually adjusted to take into account factors such as machine, crane and operator availability. The resulting production schedule is then passed to the production department for processing.

4.1.3.2 Machine Cognition cycle

The goal of the machine cognitive framework is to provide a predictive mechanism for machine failures based on learned behaviours for normal process operations. To achieve this a complex interrelation data set will need to be collected for machine variables while in production and at rest. This will in essence provide a digital twin for the machine which will build a model of the acceptable variations in the variables whilst the machine is in different modes of operation. From these models', perturbations in variables outside normal operating conditions can be analysed using CEP (Complex event processing) or other techniques to provide an output to the user for possible causality and a time related prediction of outcomes.

Proposed sensor data from EVG pilot machine

- Step time - variations in time per unit distance (require shape code info)
- Mains Current – variations in current demand / feed speed in operation and at rest
- Mains power – variations in power demand
- Hydraulic temperature – variation in temperature outside operating parameters
- Hydraulic pressure – variations in pressure at rest and operating
- Step power – power required to perform feed, bending and cutting operations (require shape code info)
- Step time – time taken to perform feed, bending and cutting operations
- Feed speed command – feed command value for bar to be used with step power/time variables
- Bend angle command – command for bend operation to be used with step power /time variables

4.1.3.3 Detecting Variations

To be able to optimize behavior we will first need to understand what the variations are likely to be.

- For instance, if the hydraulic pressure drops and can't be maintained by the pump there will be a list of likely root causes, e.g. the pump has broken, has become worn or we have an oil leak.

- The same is true for combinatorial events where multiple signal variations are recorded such as the power drops and the pump pressure drops and the hydraulic temperature rises, all of which could be the result of a faulty valve.

So, each signal being measured has to be considered in detail, and what variations in those signals could indicate with respect to machine operational efficiency.

4.1.3.3.1 Sensor signals

- **Step time - variations in time per unit distance (require shape code info)**

Step time when consider in conjunction with other measured variables and operational data for product code and dimensions can be used to produce a table of operating parameters which take into account the size of the product being produced and will allow for variations of the current operating characteristics to be analyzed against historical results and possibly highlight any discrepancies.

- **Mains Current/Power – variations in current demand / feed speed in operation and at rest**

The mains current used at rest has a direct bearing on the machine health to retain hydraulic pressure and therefore the efficiency of the hydraulic pump, would be used in conjunction with hydraulic pressure and temperature.

- **Hydraulic temperature – variation in temperature outside operating parameters**

Hydraulic temperature is proportional to the work being done by the machine so when analyzed in relation to the machine operation time gives a good indication of the hydraulic component's health and operating efficiency.

- **Hydraulic pressure – variations in pressure at rest and operating**

Hydraulic pressure should be a constant when the machine is running as it will be controlled by a pressure switch, if the pressure drops during operation this could be indicative of a faulty pump. Variations of the pressure when the machine is idle will also indicate the pump is struggling to maintain pressure due to a pump problem or possibly a leak

- **Step power/time – power and time required to perform feed, bending and cutting operations (require shape code info)**

Step power will be measured and stored for individual cycle operations; this will be indicative of the process being performed so will be dependent on the product being produced so the shape code for current production will need to be considered. Variations for similar products could indicate changes in the machine performance.

- **Feed speed command – feed command value for bar to be used with step power/time variables**

This is the feed speed for the bar being fed into the machine, the variation in this could indicate bar slippage caused by worn drive rollers.

- **Bend angle command – command for bend operation to be used with step power /time variables**

This is the command to the bending mechanism to bend the bar. This when measured against the bending time could be used to indicate a problem with the bending mechanism or hydraulic pump.

4.1.3.4 Understanding Variations

The impact of variations and step changes in signals from the process have the possibility of a range of outcomes from immediate breakdowns to a gradual degradation in the operation efficiency of the machine. Without any historical data it is difficult to predict the outcomes of a set of anomalous signal changes but we could give a best guess scenario for possible outcomes. To be able to give a more detailed study of these outcomes ideally, we need to look at historical data for the machines for previous breakdown events and see if these could have been preconceived from data analysis. To understand what variation of the measured variables looks like, we first need to have a set of base line "normal" data set for each variable being measured for all of the machine operating modes and product codes. However, an initial analysis of the measured variables can suggest possible causal events which will then need to be confirmed by empirical results from the machine when connected.

To this end a matrix has been developed (Appendix 1) which gives the expected sensor data being captured along with predicted failure modes for the signal when different types of signal variations are experienced.

4.1.3.5 Optimization in the BRC case

Having considered the overall processes and interactions of the Cognitive Framework the optimization aspect of the framework relevant to the BRC case is responsible for the creation of the production schedule based on the input from the remaining modules. Therefore, the optimization module in the overall framework and respective case creates a robust production schedule that may take under consideration the different identified bottlenecks, anomalies etc. of machines and cranes and always based on the availability of data.

At any given point cognition may lead to the identification of events affecting optimization (a) a machine is malfunctioning / brakes down or should be stopped for maintenance, (b) a crane is not available to load/unload a machine or transfer a finished product, (c) a finished product cannot be tracked down in the laydown area and (d) an order has an updated delivery time (priority has changed). In the case these events / anomalies are identified or predicted, Optimization following the respective feedback from the other modules and the relevant data is able to schedule the production sequences with respect to the machine capabilities and availability. The same may stand at a later stage with crane availability and capabilities. Lastly, as the events in production can be stochastic the production plan will be created in a robust sense. In this multistage flowshop problem and within the cognitive framework the optimization module is fed (from other components) with indications with respect to, productions times, anomalies detection relevant to the machines' availability and schedules of maintenance. More precisely, other components provide inputs with respect to operation and set up times for each product type in every production step, different detected anomalies in the involved machines informing about a potential problem (e.g. underperformance based on currently produced batch and oil pressure levels and speed). Lastly and in relation to the cranes detected anomalies in operation (e.g. availability, movement based on production schedule etc.) will also have to be identified in order to also inform the optimizer so as to be able to derive to a schedule.

Lastly to elaborate on the application of the Optimization module to the Cognitive Framework for the BRC case instantiation as the overall goal post cognition is to provide a robust production schedule for products created from Bars, the following interconnections and data exchange is foreseen with the other modules of the Cognitive Framework.

1. Optimization Module Receives input on:

    - Orders, Raw Material, from (FACTLOG persistence data pool stemming from BRC systems)
    - Availability (operational, under repair, blocked or starved), feasibility of assignment of jobs to machines, processing and setup times of machines per product
    - Operational availability, transportation times of cranes
    - Transformation process (Raw material -> Product from machine) from BRC Flowcharts

- Crane movement (Pick up / Loading / Times taken - potential scenario)

2. Registers for acceptance to FACTLOG, an optimal production schedule of jobs of the given orders with respect to the specified optimization criterion.
3. Upon Acceptance, the schedule is implemented at the shopfloor and redirected to the Message Bus to be stored in FACTLOG persistence and utilized for further consumption (e.g. to Process Models in order to examine feasibility of the schedule, overall duration and to compare its efficiency with alternative possible schedules)

When the realized scenarios deviate substantially from the input from #1 corrective actions might be needed to restore feasibility.

4.2 Pilot PIACENZA

4.2.1 Brief introduction

The research and development work for Piacenza pilot in FACTLOG starts with the study of the optimization problem in the weaving department. The study of the finishing department has been postponed in order to decompose the cognitive and optimization problem of the textile process (namely the weaving and the finishing) in sub-problems, by following the sequence order of the manufacturing (weaving first, finishing second).

The weaving is organized and planned on looms. The textile worker (i.e. weaver) organizes jobs on looms according to:

- respecting the deadline with the customer (model constraint, rigid and not changeable)

- optimizing queues (human decision, currently not optimal and not considering energy savings)

Deadlines

The most common unexpected event is a new and more urgent order to be managed. Orders might come from 3 regular categories: production, sampling or prototyping. Because the strategy of Piacenza is based on its capacity of develop new designs, it offers more than 1,000 new fabric design per season (F/W and S/S - women's, men's, accessories), >30% customised for the customers, and grants a first prototype delivery in less than one week. Consequently, the internal orders from these last two categories are launched daily. This type of event changes the planning of the jobs on the machines. Each order has a priority (that can be modelled on a scale from 0 to 100). If an unexpected event comes and it has an earlier deadline, the scheduling will be adjusted according to current deadline overview.

In case the deadline is the same, the decision is taken upon the priority value. Higher priority will be processed earlier, considering the risk to get some delays in the queued orders. Delays of production causes the request of discounts by the customers, which have to modify their production planning accordingly.

Due to COVID 19, the planning of production is becoming even less predictable: in the past some orders of fabrics where placed in advance by customers on the basis of their sales forecast for the next season. Now, any forecast has become impossible and these orders are not placed anymore. Therefore, it is expected an increase of the season packs of production alternated with periods of layoff of workers due to lack of orders. Also, in relation with this situation, the rapid optimisation of production will become even more critical for Piacenza, and in general for textile and fashion sector.

The deadline in the dataset refers to delivery to the customer, not to the internal department deadline. Therefore, it encompasses both weaving and finishing processing time spans. In order to determine the internal deadline for the weaving department for finishing the job, it is necessary to subtract 18 working days from the expected deadline. The remaining days are the expected time span for the finishing process. In case of the time span being negative, a delivery delay will occur.

Optimizing queues

Considering the weaving, Piacenza has a fixed number of Uniform Parallel Machines. Jobs are managed in meters of production. Same jobs can be split on different looms (physically cutting the chain) and parallelized, in order to address a deadline. Currently the process is not optimized toward energy savings.

On the contrary, the finishing is a process organized by a sequence of steps. The sequence changes according to the type of fabric. It is not decided automatically, but by the finishing required. Weaved fabrics which face the same sequences are put together to form the finishing LOAD. The load is created in order to respect deadline which is the primary target. To save energy is a secondary one, when possible. Currently, the process is not optimized toward energy savings. At the time of writing this document, Piacenza is working to find the most effective way to collect and update this type of data which are not fully digitalized, in order to address the FACTLOG aims and scope.

It must be also highlighted that looms are machines using only electricity (an on/off process – directly proportional with the running time of the machines) while finishing process exploits also warm water and heating. In this last case all the sources of energy must be taken into account, considering that some of them are generated by the general supporting services of the company and that, in some cases, also the machine start-up consumption from cold must be taken into consideration. A typical case is the warm water, which is generated by 3 very large steamers for the whole company. The starting at least of one of them is necessary even in the case of the finishing of small orders and therefore the planning of the finishing process must carefully consider the warm water consumption.

4.2.1.1 Job assignment for weaving

The job assignment is performed according to the following (human) decision making process. The weaver has a list of orders (jobs) to be assigned to the looms. To assign each job to a most suitable looms (creating the queue) the weaver takes into considerations the following facts:

- **orders deadlines**
- **CC/CA codes** of the order (job): if they are compatible it is possible to chain (literally) the wefts, or to make knots (knotting warps), saving time for the loom setup (loom setup and its rules will be presented later on).
- **the ongoing jobs** on the looms, the remaining time to finish such jobs.

Considering that the weaver decides where to allocate the new order, according to the exanimated context. It could possible that the weaver decides to setup an empty loom, even if the CA/CC is compatible with an ongoing work. This decision is due to the fact that the ongoing job will finish too late to put the other one in queue. Hence the setup of an empty loom is the most suitable choice to avoid to miss the deadline.

It must be highlighted that the fixed time to prepare and load a new warp on a loom is from 4 to 8 hours, independently from the length of the warp (explained in detail in par.4.3.1.4.). Therefore, the decision not to exploit the CC/CA chance causes a significant increase of preparation time and the consequent reduction of production capacity.

4.2.1.2 Weaving current dataset

The current dataset of weaving department contains the following information:

- **Status**: there are three types of status: LANC, i.e. the order will be processed; however, no warp neither weaving steps are performed; ORDI, i.e. the chain is available and the warp is ready; ATEL, i.e. the order is scheduled for the loom for weaving.
- **Chain**: it is the code of the job that can be ongoing (status=LANC) or that can be arriving (status=ORDI).
- **CA & CC codes**: are the codes used for the setup rules
- **Type**: is the type of production. P = Produzione (regular production); C = Campione (sample); F = Fazzoletto (pre-sample)
- **Quantity** (m): is the meters of fabric to be produced
- **Delivery date**: it is the deadline for the delivery to the customer
- **Fabric**: it is the code that with design code determines CA & CC
- **Design code**: it is the code that with fabric code determines CA & CC
- **Strokes per meter**: it is the speed of the loom
- **Yarns number**: it is used in setup rules
- **Comb code**: it is used in setup rules

The chain is ready to be weaved when the warping job is finished. In the case of the weaving department, we need to consider just these two different cases: status = ORDI (the chain is still in the warping department, it is arriving); status = LANC (the chain is already on a loom). When the status is ATEL, the order is scheduled to start at start date.

More precisely, the finishing time of each ORDI is the time of warping finishing, that does not correspond to the setup time automatically. Hence, we can consider them a unique queue. It is not relevant when the warping finishes. We can plan the assignment of ORDI orders to the most suitable loom, according to deadlines and remaining job on looms.

If the status is LANC, the only choice that can be done to optimize the loom jobs is to make knots (if there is compatibility). If the status is ORDI the textile worker, according to the length of the CHAIN (ref. to quantity) and the CA and CC codes can decide to make the job on a unique loom, or to split the chain to different looms, in order to parallelize the job (currently it is a human decision). It means that the quantity of the job may be split in parts that can be executed in parallel in multiple machines. Hence, the CHAIN is not referred to machine (looms), but the same chain code can be worked on two different looms (we have two records with the same chain code).

Currently, **the priority parameter is not present** into the dataset. This because it is a human decision that is taken at the moment, when a new urgent order occurs. Whether it happens, the loom is setup and the order is inserted into the system. We have no direct evidence of this, in the dataset. For this reason, considering the FACTLOG model, the optimization algorithm can assign a priority. It is not important to stick to the reality, because priority is changed manually according also to external factors, usually related with the commercial priority of the customer.

4.2.1.3 Weaving improved dataset

In order to support the improvements aimed for in the project, the collected data will be extended with new data sources, in particular with regards to the quality of input materials (e.g. yarn for weaving) and from inside sources, including incremental output (e.g. fabric quality) and performance data (e.g. machine speed, inferred energy consumption for looms). Namely the FACTLOG DB will include:

- **complexity index**: it is related to the fabric quality and affects the performance of the loom.
- **inferred energy consumption for looms**: we will calculate the theoretical energy consumption for a subset of looms, relying on the nominal value of the machine consumption and the current energy consumption of the weaving department

The complexity index shall be considered in this way:

- from 1 to 6: easy
- from 7 to 14: medium
- > 14: difficult

About yarn quality, at the time of writing this document Domina is collaborating with Piacenza to identify if the yarn lot is tracked outside the warehouse, in order to be monitored also in the weaving department. Whether or not, it will not be possible, for current model, to include further yarn information.

4.2.1.4 Loom setup rules

Setup of the looms has different level of difficulty that affects the time we need to setup a loom (hence the decisions for the optimization, the job splitting, etc.)

Loom setup can be:

- **Easy**: 4h of works of the setup team
- **Medium**: 6h of works of the setup team
- **Difficult**: 8h of works of the setup team

SETUP TYPE

1. If the CA code of the outcoming chain (close to be finished) is the same of the incoming chain, the setup type is KNOTTING.

1.1. If the CA codes are different the following check shall be performed: if the yarns number (“FILI”) and the comb code (CODICE_PETTINE) are the same, it is still considered a KNOTTING
2. According to the check done at point 1, it is established that the KNOTTING is not feasible, we need a FULL SETUP

SETUP DIFFICULTY LEVEL

1. KNOTTING: if yarns are > 5000: hard knotting. If yarns are between 3000 and 5000: medium knotting. If yarns are less than 3000: easy knotting
2. FULL SETUP: if the comb height (ALT_PETTINE) is the same or less than 10 cm: easy setup. If the difference between the two comb heights is more than 10 cm: difficult change.

The setup teams have a total amount of hours in a day that can be used for setup (independently from the number of workers) = 50 h/day. Each setup shall be subtracted from this amount. If the hours available for the setup are all used (hour = 0) any other needed setup is postponed to the following working day.

The setup follows the priority according to the loom number, in a FIFO (First In First Out) scheduling. The first loom that has finished the work is the first one that is setup for the next work.

Efficiency of the loom: from 0,6 = easy fabric; 0,5 = medium fabric; 0,4 = difficult item. The difficulty of the item is related to the performance index. This information is not available into the historical data set, but it could be considered for the planning.

The efficiency of the loom shall be applied to the loom speed.

4.2.2 Cognition cycle

4.2.2.1 Detect variations in the behaviour (variations / anomalies / outliers)

Understand the variations

The working calendar is organized as follows (5 working days in the week): two work shifts on January, February, August (but it is closed 3 weeks for vacation) and September. In the other periods there are three shifts. One worker is responsible to monitor 3 looms, and recover the related broken yarns, whenever it happens.

The goal of the use of cognitive framework in Piacenza is to optimize and increase the production by managing the following aspects:

- New urgent orders which shall be managed with the others
- Prevent looms malfunctions, which can cause delay in orders
- To optimize the creation of the finishing loads, to save energy

At this concern, the relationship among looms malfunctions, related for example, to component wearing, has never been addressed in Piacenza. Hence the company interest is very high. Currently, Piacenza strategy to prevent machine malfunction is to perform a careful maintenance of all the plant machines during the month of August, when the production is stopped for the summer break.

We have one main variation, that encompasses both the weaving and finishing departments (we can consider the whole plant), that is hardly predictable: incoming orders with maximum priority. The real challenge, for the cognitive optimization system, is to manage the variation, without causing delays for the other orders. Any delay is an economic loss for the company.

Hence, once a new urgent order come from the commercial office, ongoing jobs shall be re-scheduled in weaving department (manual intervention firstly to assign priority and deadline, then the scheduler assigns the jobs to the machines). Then ongoing jobs shall be re-scheduled in finishing department too. In finishing, jobs with the same sequences are carried out together, to save energy, creating the so-called “finishing loads”. Whenever an urgent order comes, currently it is not be possible for the human operator to optimize both energy savings and deadlines.

About **weaving anomalies,** it can be considered as anomaly in weaving process the breakage of yarns during weaving. It is a sudden event. If it happens too frequently the most likely causes are two: the loom speed is too high; the yarn lot has quality problem. We have said that a single worker has to monitor 3 looms, and s/he needs about 15 minutes to fix broken yarns.

About unusualities, within FACTLOG, Piacenza has opened a new industrial research project in its plant, to monitor spikes in loom energy consumption, in order to determine whether they are linked to the wearing of the components. In this case, a better maintenance planning could be feasible. For Domina, these results represent an important experience to include in its ERP/MES software preventive maintenance functions.

About **finishing anomalies and unusualities**, at the time of writing this document, Piacenza is working on the modelling of finishing process, for cognitive research purposes.

Summing up, we can define the following variations:

- New order inserted into the scheduler of the weaving department
- Spikes in looms energy profiles
- Yarn breakage (loom stopped)
- Changed composition of the finishing loads

4.2.2.2 Understand variations

The goal of this part of the analysis is to provide additional context/knowledge for problem (variation) analysis and support root cause analysis.

Useful information/knowledge to perform this study will be provided by Domina and Piacenza, by means of an anonymized and access-restricted database for project partners. This database is provided under the NDA project agreement and the dataset can be used only for the project purposes.

The database will make available for project partners, by means of API, the following information:

- **Historical data set of scheduled weaving orders** (as explained in par. 4.2.1.2 and 4.2.1.3): an initial data set has been provided as Excel file

- **Finishing processes definition**: to be provided
- **Finishing sequences target for optimization**: to be provided
- **Historical energy consumptions**: to be provided
- **Current energy consumptions**: to be provided
- **Theoretical looms energy consumption**: to be provided (computation formula already identified)
- **Real looms energy consumption**: to be provided, once the installation of the meter will be confirmed.
- **Knowledge about weaving process/product**: provided into D1.1 and D1.2
- **Knowledge about finishing process/product**: partially provided into D1.1 and D1.2
- **Maintenance scheduling**: to be provided, if needed

4.2.3 Optimization in the PIA Case

Having considered the overall processes and interactions of the Cognitive Framework the optimization aspect of the framework relevant to the PIA case is responsible for the creation of the production schedule based on the input from the remaining modules. Therefore, the optimization module in the overall framework and the respective subcases creates a robust production schedule that may take under consideration the different identified bottlenecks, anomalies etc. of machines (e.g. looms) and always based on the availability of data. In the sequential fabric production process of PIA, the different process steps that include different sets of machines will have different potentially interesting anomalies detected that can be relevant to: (a) a loom / finishing machine is underperforming /malfunctioning/ breaks down, (b) a (batch of) yarn is broken and requires temporarily halting the machine and repairing the breakage or re-setting the speed under which the fabric or yarn is processed. Additional events may include a high priority order arrives that needs to be handled. On that account, Optimization needs to examine the whole process to support the production scheduling process taking under consideration the current and upcoming orders by proposing efficient algorithmic methods, that indicate which orders should be processed in what sequence and on which machine, so as the products to be ready on time for shipping.

In order for Optimization to be able to enable the dynamic scheduling of the fabric production process, different modules of the cognitive framework should be able to provide indications relevant to:

1. Detected anomalies in the active machines (e.g. looms, finishing machine)
2. Detected anomalies in the intermediate steps of production (e.g. breakage of yarn)
3. Scheduled maintenance activities
4. Relation and similarities among orders under production and on queue
5. Energy consumption within every production step.

Therefore, optimization can create a reactive scheduling based on input received from the different modules and with respect to detected anomalies and understood the meaning behind them. The following interconnections and data exchange are foreseen between Optimization and the other modules of the Cognitive Framework.

1. Receives input on:

    - Identified (or predicted) anomalies /maintenance activities/performance degradation of active machines (looms, finishing machines)

- Order related information (e.g., quantity, deadline, release time, fabric properties including the splittability of each order/product) from
- Machine related information (speed, availability) and workers' information (shifts, availability)
- Energy consumption information of different units

  (All aforementioned will derive from FACTLOG persistence data pool stemming from PIA systems)

2. Provides a Production Schedule with respect to the orders that exist and the ones that come in to FACTLOG for acceptance (via operator)
3. Upon Acceptance, and introduction of the order to a queue, the schedule is initiated at respective machines and then redirected to the Message Bus to be stored in FACTLOG persistence and utilized for further consumption.

4. When the realized scenarios deviate substantially from the input from #1 corrective actions might be needed to restore feasibility

4.3 Pilot JEMS

4.3.1 Brief introduction

JEMS is developing and selling waste-to-fuel transformer plants. These plants are transforming any hydrocarbon-based waste into high-quality synthetic diesel fuel. In these plants, JEMS uses a chemical-catalytical de-polymerization process that runs on a low temperature and low-pressure. Due to the low temperature, no harmful or carcinogen gasses are produced as by-products. The transformer plants can process biochemical components such as wood, paper, agricultural residues, and food leftovers. The resulting synthetic diesel fuel complies with the highest quality standards. It can be used in any modern diesel engine or electricity generator without any negative technical or mechanical impact. The synthetic diesel has a low clouding point and can be used as an additive for low-temperature use.

The transformer plants have multiple pipelines, and downtimes usually result from clogs that take place in them. In this context, cognition is required to understand what may lead to clogs, predict them, identify anomalies and their root causes, and understand what parameters result in the best production setting. These models can be built based on data obtained from sensors used to monitor the plant's different stages. Such data may help to understand factors and parameters that influence overall efficiency, identify potential failures to avoid downtimes, and reduce the number of operators required to operate the plant efficiently.

The complexity of the problem is related to the wide range of factors that can influence diesel production, from mixtures of ingested waste to different configurations required to process these mixtures efficiently, minimize downtimes, and produce the highest possible amount of diesel per hour.

The main KPIs of interest for this use case is:

- **efficiency**: the amount of fuel per hour produced by the transformer plant
- **downtime**: reduce the amount of working time affected by process failures, preventing regular operation and fuel production
- **number of plant operators**: reduce the number of operators required in each plant

4.3.2 Framework description

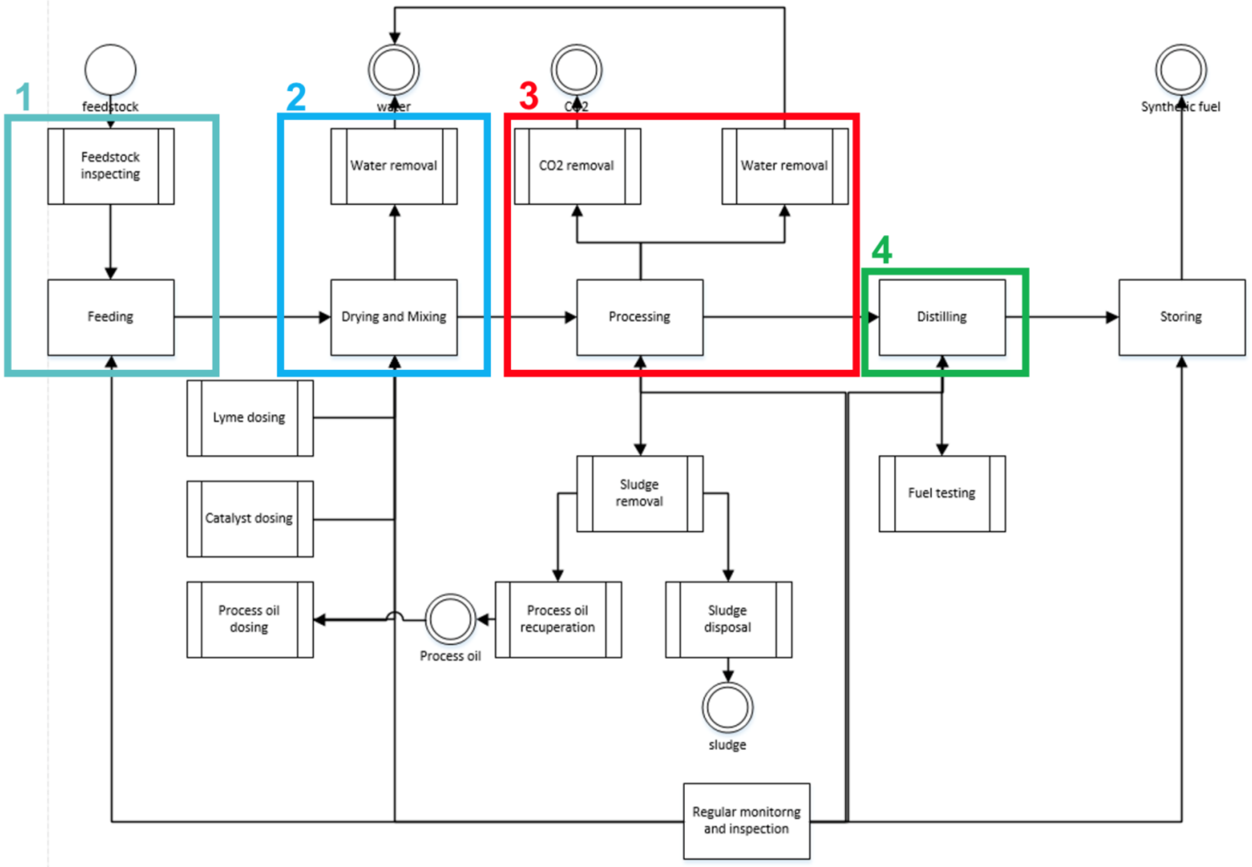


Figure 12: JEMS waste to fuel plant diagram

There are four separate stages in the JEMS waste-to-fuel transformer plants, where models can be applied to solve different challenges.

In the first stage, the feedstock is prepared, cutting, and mixing different feedstock types before feeding them into the system. We are interested in finding the best feedstock mixture possible to maximize the amount of fuel produced per amount of feedstock.

The second stage is related to water removal, and the addition of different chemical components required in the next stage to start the chemical transformation process. In the second stage, there were no anomalies detected regarding temperature and pressure, which are controlled through a sensor feedback loop, to maintain required conditions.

In the third stage, the mixture is heated and mixed to facilitate the chemical process. In this stage, we continue to remove water through evaporation and remove CO2 and leftover sludge.

The fourth stage corresponds to distillation, where the original mix is heated producing vapours that condense at different levels to produce different kind of fuels, which can be tested and stored.

There are two challenges that are required to be solved for this use case: avoid pipes clogging and enhance the feedstock composition to increase fuel production efficiency. The waste-to-fuel transformer plants have multiple sensors, which can be used to these purposes. Most relevant are sensor readings which inform:

- Ingested material flow speed
- Mixing power
- Temperature
- Pressure
- Turbine flow
- Pump speed: pump speed when pumping from B100 to turbine

Regarding pipe clogging, it is important to monitor mixture density, which can be influenced by its original composition as well as the water removal process. Excess of water removal increases the mixture density and viscosity. This can be indirectly measured with information regarding power required by machines to mix, variations in pump speed when pumping from B100 to the turbine as well as turbine flow velocity. This may help us find patterns on how increasing density and viscosity correlate with clogging events. Such information is useful to alert the user to take preventive measures, such as add more process oil to ensure the whole mixture can flow appropriately through the pipes.

Anomalies in pressure and temperature are unlikely to occur but require careful monitoring as well. In the second and third stage it is crucial to monitor operation and identify possible anomalous conditions, such as sudden increases in temperature above 160°C and 300°C respectively. Such a situation can indicate that something went wrong with the process - a situation that we must prevent and avoid in a future. If these temperature variations are short-lasting (less than thirty minutes), there may be two possible scenarios: that there was some issue with the sensors and temporarily did not correctly sense the temperature; or that the temperature was temporarily increased on purpose. If temperature changes last for more than 30 minutes, there is certainty that something went wrong with the process. Such a variation may require stop the whole fuel generation process, affecting downtime metrics as well as to affect fuel generation efficiencies. These anomalies require to understand the chemical process, and if the feedstock mixture requires some adaptation in order to avoid such a situation in the future. The method used to this purpose may be a model developed by TUC, which based on mixture parameters shall help predict expected fuel generation efficiencies. If the mixture does not behave as expected, we may find the mixture was not properly prepared or that the model may not properly illustrate some cases and requires some adjustment. We may react to such a situation making use of a knowledge graph model, which suggests which actions may be taken by the user based on the anomaly severity.

FACTLOG framework can help in both aforementioned scenarios: prevent pipe clogging and optimize input materials mixture (feedstock). Pipe clogging can cause downtimes of nearly 7hs to the whole process, before getting fixed. Feedstock mixtures directly affect fuel efficiencies, but currently there are no records regarding different input mixtures and achieved efficiencies.

In this context, the FACTLOG framework aims to provide three services.

First, develop mathematical models that give specific feedstock options and compute the best mixture that maximizes the amount of fuel produced by the waste-to-fuel plant. We can develop such a model given known feedstock parameters, and later adjusted to new components as well, given new feedstock material properties.

Second, over the whole process, it is necessary to monitor for anomalies and understand the root causes that drive them and how do affect it. This process can be done two-fold: using a general anomaly detection framework to identify events or work on past anomalous events to identify similar future conditions.

Third, to relate the anomalies to events that cause it, and to understand the impact it may have over the waste-to-fuel transformation process, a knowledge graph is required to semantically model and link the different waste-to-fuel plant components. We base such a knowledge graph on an ontology developed for the FACTLOG project. The knowledge graph can be instantiated based on real data and model outcomes used to solve different use cases. We can mine the knowledge graph to identify patterns, which help us to understand normal operations, or identify corner cases, understand how frequently they may happen, conditions that lead to them, advice how the users shall react to such situations, and prevent them in the future.

4.3.3 Cognition cycle

4.3.3.1 Detect variations in the behaviour (variations / anomalies / outliers)

Understand the variations

The working calendar has seven working days in the week, with three working shifts and the process does not stop on vacations.

The goal of the use of cognitive framework in JEMS is to avoid pipe clogging or anomalous situations that may require to stop the process, and to increase fuel production efficiency by managing the following aspects:

- Identify factors that may lead to pipe clogging, such as increased mixture density
- Identify factors that may lead to anomalous situations such as increased temperature and pressure
- Optimize the feedstock mixture, to increase waste-to-fuel efficiency

Currently JEMS supports a single type of feedstock mixture, and prevents clogging by performing human monitoring of multiple sensor values and adding process oil when estimated is necessary.

We have two main variations, that need to considered. First, mixture density shall be measured through the process, since increases in density also increase the likelihood of clogging and thus the risk of a process downtime. Second, we have to monitor for anomalies in temperature and pressure, which may signal some intervention is required, incurring in a downtime. Each downtime has an important impact on production efficiency and produces an important economic loss for the company.

About process anomalies, it can be considered an anomaly in the process the significant increase of power required by machines to mix or pump the mixture to the turbine, which signals increased density. It is also considered an anomaly an increase of temperature above 160°C and 300°C on second or third stages respectively.

About unusualities, within FACTLOG, JEMS opens a new industrial research issue in its plant, waste to fuel efficiencies related to feedstock composition as well as likelihood of feedstock composition to be associated with some kind of anomaly (increased probability of clogging or a higher frequency of anomalous events related to out of expected range temperature and pressure values). Such capabilities would be a valuable addition to the software regarding plant preventive maintenance functions.

Summing up, we can define the following variations:

- Changed composition of feedstock mixture for the waste-to-fuel process
- Levels of power required by mixers to mix and pump the mixture
- Pipe clogging (requires stopping the process)
- Spikes of pressure or temperature (requires stopping the process)

4.3.3.2 Understand variations

The goal of this part of the analysis is to provide additional context/knowledge for problem (variation) analysis and support root cause analysis.

Useful information/knowledge to perform this study will be provided by JEMS, by means of an anonymized and access-restricted database for project partners. This database is provided under the NDA agreement with each interested party, and the dataset can be used only for the project purposes.

The database will make available the following information:

- **Sensors description:** provides an ID and description of type of sensor, units in which measure data
- **Values measured by sensors**: provides time series with values measured by an array of sensors at different stages of the process
- **Anomalies were not tagged in data**: such data is not available, and other approaches are required.

4.3.3.3 Impact of Variations

We find the following impacts of variations:

- Changed composition of feedstock mixture for the waste-to-fuel process:
  - may increase waste-to-fuel efficiency
  - may influence the likelihood of process anomalies, such as increased temperature or higher density of mixture and potential pipes clogging
- Levels of power required by mixers to mix and pump the mixture:
  - high levels of power signal increased density of mixture, which may lead to clogging or machine shutdown to prevent the machine from breaking down
- Pipe clogging (requires stopping the process)
  - when a pipe clogging event takes place, the process must be stopped in order to perform required maintenance
- Spikes of pressure or temperature
  - spikes in temperature or pressure may signal there is something wrong with the process, and may require to stop the waste-to-fuel plant.

4.3.3.4 Optimization
In the JEMS use case there is opportunity to apply optimization models, which model feedstock properties and how those translate to different levels of fuel production. Such models shall help to make decisions on the best possible mix that shall be fed into the waste-to-fuel plant.

4.4 Pilot TUPRAS

4.4.1 Brief introduction
TUPRAS has a complex network of process lines which work to produce LPG. The main scenarios are the following:

a) To identify a potential anomaly in a particular phase of the process
b) To understand the impact of this anomaly in the whole context of the LPG flow in an attempt to find the optimal interventions.

The above is a full representation of the cognition process. The need for cognition comes from the fact that anomalies come from different sources:

- Off spec production
- Sub optimal production

In both of the above cases, we need to detect the types of anomalies, understand their nature (is it just a threshold? It is an outlier? Other?) and decide whether this requires some action for improvement/ correction. Therefore, analytics regarding the anomaly detection and understanding variations are needed.

When such an anomaly is detected, we need to understand how this impacts the process (in which the particular DT where the anomaly found belongs to). This is not obvious because the final LPG (stored in the final tank) is a mixture of outputs from 10 different processes and in some cases those processes have common units/steps.

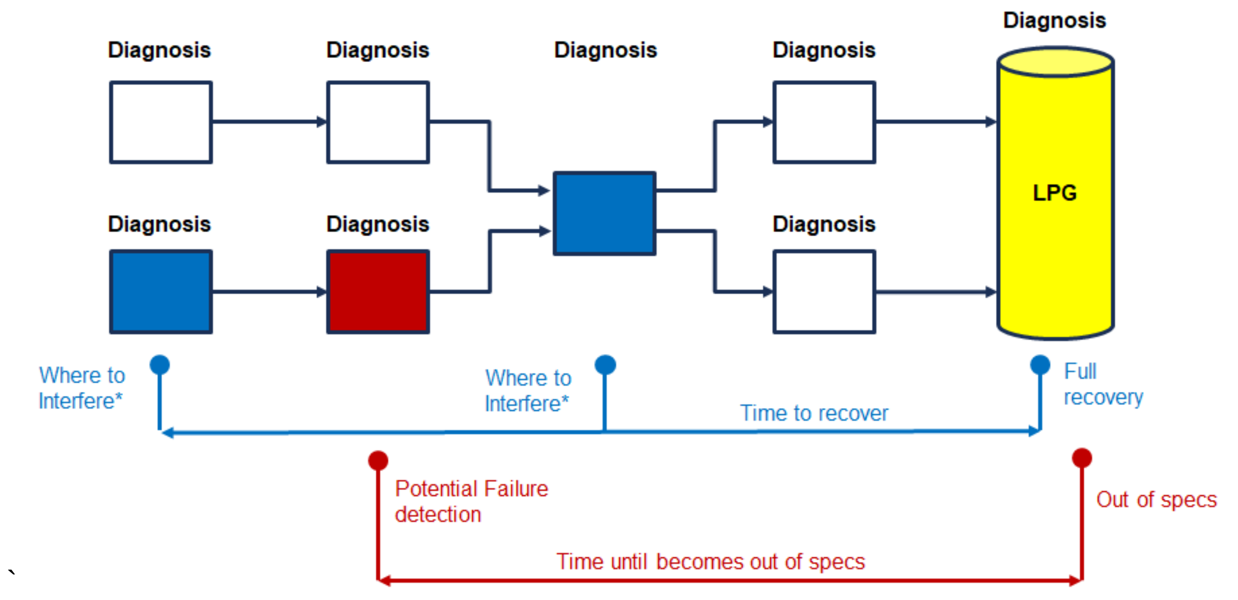


Figure 13: Anomaly detection, impact assessment and understanding intervention points (indicative schema)

4.4.2 Cognition cycle

Based on discussion in the Cognitive Factory Framework teleconferences

4.4.2.1 Detecting Variations

The goal of the use of cognitive framework in TUPRAS is to detect, in the early phases on the route from LPG raw streams towards LPG refined streams, possible trends and anomalies of the ingredient constitution threatening to compromise the quality in the final output tank, and to subsequently perform energy-based optimization of the on-specs recovery process. The above can be achieved by managing the following aspects:

- Identify factors that may lead to off-spec or suboptimal production, mainly in terms of the impurities being within legally set limits (chief among them being the sulphur content)
- Identify factors that may lead to related anomalous situations, such as problematic material flow rate, pressure, temperature, etc., including performance indicators, at unit (e.g., DEA), process and tank level.
- Optimize the operational parameters of involved process units (e.g., increase temperature at the top of the unit by a specified number of degrees), in order for the LPG within the final tank to recover to on-specs production within a given time frame.

Currently, the relationship between the production and storage tank qualities are controlled manually by the planning and production engineers. If some anomaly occurs throughout the production line, its root cause is investigated by trying to detect anomalies at each step of the production separately. Further, if there is off-spec production anywhere in the product line, which cannot be seen in the process parameters, then it cannot be noticed until the lab results arrive. And even if an anomaly is identified in the process parameters of a specific process unit or in the product pool, the responsible engineers make decisions about what corrective actions should be taken again with restricted information about the rest of the process line. The above can result in time delays, suboptimal resource exploitation and product quality, and even unsellable (due to off-specs) blend.

We have two main types of variations, that need to be taken into account. The first concerns off-spec deviations of LPG impurities based on online and laboratory analyzers. The second is about monitoring for anomalies in indicators like temperature, flow and pressure, in order to recognize performance trends impacting on the quality of the LPG in the final tank.

About unusualities, within FACTLOG, TUPRAS opens a new industrial research issue in its plant, by monitoring process variables and associated anomalies and enforcing data-driven methods, in order to determine whether they are linked to the level of impurities eventually measured. Such capabilities would be a valuable addition to the software regarding plant preventive maintenance functions.

Summing up, we can define the following variations:

- Chemical properties of the LPG (including levels of various impurities) measured by online analysers at different points in the process and in laboratory setting.
- Sensor readings concerning temperature, flow, pressure and capacity level.

4.4.2.2 Understand variations

The goal of this part of the analysis is to provide additional context/knowledge for problem (variation) analysis and support root cause analysis.

Useful information/knowledge to perform this study will be provided by TUPRAS by means of an anonymized and access-restricted database for project partners. This database is provided under the NDA agreement with each interested party, and the dataset can be used only for the project purposes.

The database will make available the following information:

- Sensor data
    - Sensors description: provides an ID and description of type of sensor, units in which it measures data, range, etc.
    - Values measured by sensors: provides time series with values measured by an array of sensors at different stages of the process
- Readings coming from online analysers at different points in the process and in a laboratory at the collection tank. The ranges of these readings differ depending on what impurity they are testing for, as well as in frequency. At least two years of historic data are available at the frequency of up to 1 data point/sec (for non-laboratory values). All values are real-valued numbers which are well suited for machine learning algorithms.

On top of the provided pilot data and respective analytics results, process models and knowledge graph background knowledge can support root-cause analysis in order to figure out what is causing the off-spec LPG readings.

4.4.2.3 Impact of Variations

We find the following impacts of variations:

- Level of impurities at intermediate points, as well as the final tank:
    - May influence the likelihood of going off-spec, since the levels of impurities from the individual process lines practically govern the specs in the final LPG pool
    - If not addressed, they might mean unsellable product
    - Might affect energy consumption
    - May influence the likelihood of process anomalies, such as increased temperature
- Deviations on flow, temperature, pressure, capacity level and other indicators:
    - May affect/reflect energy consumption across refining operations
    - May signal failures, response time to which is significant in order to diminish the amount of off-spec production, as well as achieve overall production optimisation.
    - May themselves lead to/reflect off-spec product
- They can indicate the highest-impact machine units/process lines and, hence, where to interfere.

4.4.2.4 Optimization in the Tupras case

Having considered the overall processes and interactions of the Cognitive Framework relevant to the TUPRAS case, the optimization aspect of the framework is responsible for the provision of a global optimization solution relevant to the recovery from off-specs LPG production within a given time-frame, while minimizing energy consumption. As the LPG production process is a complex process with different process units at any given point, different anomalies can be detected based on the ongoing production of LPG in any given unit. Additionally, it may be identified that the LPG within the final production tank has gone off-specs, i.e., the LPG in the final tank does not meet the desired quality specifications. On that account upon identification of an anomaly and its understanding of the cause of the anomaly, Optimization is triggered to examine the whole process, given the different input support recovery decisions by indicating which process units need to be utilized and under which operational scenario (e.g., increase of temperature at the top of the unit by a specified number of degrees), in order for the LPG within the final tank to recover to on-specs production within a given time-frame. Additionally, to the aforementioned, as the process to restore in-specs LPG creation takes time, several rounds of interconnection of optimization module to other modules may be required to receive info and input relevant to the ongoing values and projected outcomes.

Within the context of the cognitive framework and for optimization to be enabled other modules of the cognitive framework should be able to provide:

- **Indications with respect to anomaly detection** for the different units involved in the production, as a whole and per unit involved.
- **Modelling of the transformation process of each process unit that participates within the LPG production process.**

Each process unit transforms input into output. In that regard, the different settings that can be applied to each different unit to restore on-spec production (e.g., higher/lower temperature at the top/bottom with different levels of pressure applied, etc.) will result in different outcomes with respect to the amount/percentage of impurities removed but they also correspond to different energy consumption/cost levels. Hence, Optimization assumes that for each process unit there exists a model (which can be a physical, regression or machine learning model) of how it transforms input into output. Overall, the Optimization within the Cognitive Framework interacts with the remaining modules of the Cognitive Framework in the following manner:

1. Receives input on:

    - Identified (or predicted) off–specs situation on different units, intermediate steps and final pool from analytics

    - Identified (or predicted) anomalies in the units, intermediate steps and final pool from analytics

    - Operational scenarios for each unit, corresponding to a wide range of operational conditions. Each scenario is fully specified by the (i) flow rates and composition of inputs to the unit (feeds), (ii) operating parameters of the unit (pressure, temperatures, etc.), (iii) flow rates and composition (including impurities) of outputs from the unit (products) and (iv) energy (heat and/or electricity) consumption. The mapping from (i) and (ii) to (iii) and (iv) is achieved through unit specific models based on first principle or data-driven models based on regression or machine learning approaches from process modelling and analytics.

- Quantification of estimated performance reduction caused by anomalies from analytics and simulation.

- Real-time production data (e.g. Quantity of LPG in the LPG pool, Capacity of LPG pool currently connected, Flow rate of raw feed towards all debutanizers, Type of raw feed etc.) and real-time economic data (e.g., Price of energy consumed, Price of Natural Gas, Price of LPG) from the FACTLOG persistence and manufacturing entities

2. Provides a Production Schedule with respective settings for units that aims to drive on-specs recovery in a given timeframe to FACTLOG for acceptance (via operator)
3. Upon Acceptance, and introduction of settings, the schedule is initiated at respective units and settings and then redirected to the Message Bus to be stored in FACTLOG persistence and utilized for further consumption

4. During execution and after the given time-frame, in case of additional received off-spec situations from analytics then a new Optimization round is initiated, making the optimization execution continuous.

4.5 Pilot CONTINENTAL

4.5.1 Brief introduction

The manufacturing process from Continental has been described in detail in deliverables “D1.1 Reference Scenarios, KPIs and Datasets”, “D2.1 Analytics System Requirements and Design Specification”.

Short summary of the presented info:

- The Continental plan in Timisoara, Romania produces electronic products like airbag control units, seats controllers, hand brake controllers etc.
- These products have a high complexity degree, but their manufacturing process can be described as follows:

  - **SMT (Surface Mount Technology) lines:** highly automated lines where electronic components are placed on the PCB boards.
  - **PCBA (Printed Circuit Board Area):** PCB area where the electronics built in SMT will be separated into smaller parts (PCB’s) and tested electrically (In Circuit Test).
  - **FA (Final Assembly) and Test Area:** This is the step of production where the electronics are connected to the mechanical part and finally tested and labelled. The processes in this area connect the mechanical parts: Screwing, Press Fit, Gluing, Riveting, Snap In. The testing area consists of tests like Functional test of the product, Automatic Optical Inspection, Force monitoring for the snap in, air leakage test.
  - **Packaging and delivery operation:** the products are packed in customer specific boxes and all the information needed by customer is linked to the unique number of each box.

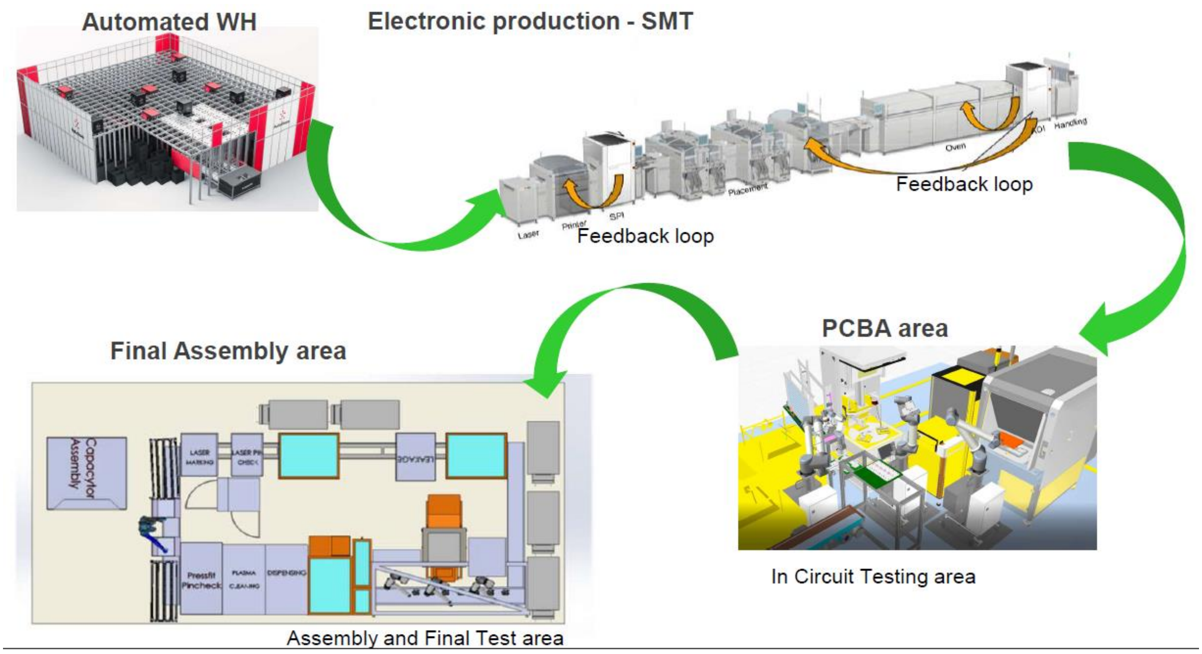


Figure 14: Production Flow in Continental Timisoara Plant

In the figure above there is a representation of the manufacturing process in Continental Plant.

In the context of the FACTLOG project, Continental focuses on the Final Assembly area. The diagram for FA area is presented below (details of the steps can be found in deliverable “D1.1 Reference Scenarios, KPIs and Datasets”):

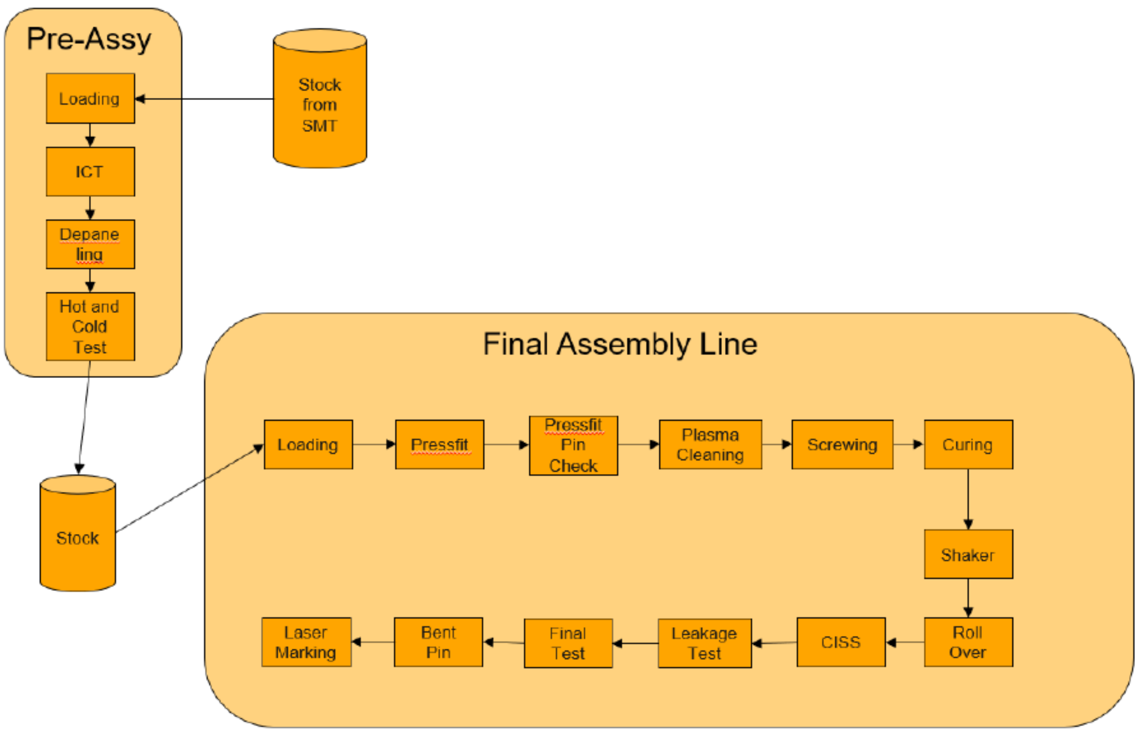


Figure 15: Final Assembly area representation

Every machine on the Final Assembly line is equipped with sensors and their measurements are sent to traceability system via the network. For each process, Continental stores specific data into traceability database.

An example of measurements from the Pressfit stations is presented in the following image:

Id	Timestamp	SerialNumber	Station	StationType	StationNum	Material	TestDescription	TestValue	TestResult	USL	LSL	Format
47853	7/3/2020 7:57:35 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730702*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	49	Pass	9999	0	R6.2
47854	7/3/2020 7:57:35 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730702*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
47855	7/3/2020 7:57:35 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730702*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.243	Pass	56.14	54.33	R6.3
50705	7/3/2020 7:57:57 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720701*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	18	Pass	9999	0	R6.2
50706	7/3/2020 7:57:57 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720701*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
50707	7/3/2020 7:57:57 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720701*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.23	Pass	56.14	54.33	R6.3
54226	7/3/2020 7:51:37 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730704*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	10	Pass	9999	0	R6.2
54227	7/3/2020 7:51:37 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730704*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
54228	7/3/2020 7:51:37 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730704*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.209	Pass	56.14	54.33	R6.3
57078	7/3/2020 7:59:12 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720703*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	37	Pass	9999	0	R6.2
57079	7/3/2020 7:59:12 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720703*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
57080	7/3/2020 7:59:12 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720703*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.201	Pass	56.14	54.33	R6.3
59930	7/3/2020 7:59:34 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730706*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	39	Pass	9999	0	R6.2
59931	7/3/2020 7:59:34 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730706*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
59932	7/3/2020 7:59:34 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015730706*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.224	Pass	56.14	54.33	R6.3
62782	7/3/2020 7:59:55 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720705*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	13	Pass	9999	0	R6.2
62783	7/3/2020 7:59:55 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720705*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
62784	7/3/2020 7:59:55 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720705*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.295	Pass	56.14	54.33	R6.3
65634	7/3/2020 7:58:42 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720707*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	11	Pass	9999	0	R6.2
65635	7/3/2020 7:58:42 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720707*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2
65636	7/3/2020 7:58:42 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720707*=	PRSFIT	Final Assembly	6	A2C15768304	Distance value	55.253	Pass	56.14	54.33	R6.3
68486	7/3/2020 8:01:05 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720709*=	PRSFIT	Final Assembly	6	A2C15768304	Carrier no	2	Pass	9999	0	R6.2
68487	7/3/2020 8:01:05 AM	3#5WA959655A #H007S0523#*C0DKTSR-TSR03.07.2015720709*=	PRSFIT	Final Assembly	6	A2C15768304	Force value	4.132	Pass	4.95	3.3	R6.2

Figure 16: Pressfit station example measurements

Main process characteristics that are tracked are:

a) process parameters e.g. torque, pressing forces;
b) measured values of the electronic components;
c) product’s reaction to different temperatures;
d) wearing of the tools during production phase;
e) deviations from accuracies machining / processing.

Also, workers from the shop floor play an important role through the information provided about equipment’s behavior in operations. After each intervention performed on machines, the workers record a report detailing the symptoms and the cause of the malfunction.

Examples of such reports are presented as follows:

Example 1:

Line	VW 40
Equipment	Housing Preparation
Product	VW20
Problem	The stopper nut on the rotating cylinder keeps loosening and no longer takes the pins ok.
Action Taken	I changed the stoppers between them and put the lock inside and the nut on the stopper on the outside to be easier to tighten and walk to adjustment. Made mechanical adjustments.
Time Spent	90
Intervention Type	Maintenance
Date Time	01.08.2020 22:03
Type Of Causes	Mechanical
Escalation	No
Solved	Yes

Example 2:

id	61718
Team	4
Shift	2
Line	VW 40
Equipment	Flipping & Screwing
Product	VW20
Problem	I had an error on the spot error panel communication with the controller.
Action Taken	I tried to reset the equipment but without success, I stopped and restarted the controller and it worked.
Time Spent	20
Intervention Type	Maintenance
Date Time	4/30/20 20:19
Type Of Causes	Mechanical
Escalation	No
Solved	Yes

To maximize life-time of equipment involved in the production processes (Process equipment but also Test equipment), Continental's maintenance & repair departments perform different maintenance techniques. Inspection, maintenance and repair activities are performed daily to annually, in conjunction with the technical prescriptions and machine age. In this regard, the sensors' measurements and intervention reports presented above are valuable information that is interpreted in order to detect early possible failures or defects.

Through FACTLOG project, Continental Pilot is pursuing the improvement of monitoring the Pre - Assembly lines and Final Assembly lines and move from only parameter monitoring to automatic reporting, automatic preventive maintenance in order to reduce/limit the number of down times caused by breakdown.

The main issues for Continental are:
1. Self-diagnosis and predictive maintenance at each machine
2. Aligning Predictive Maintenance with production plan
3. Optimized operational mode per machine
4. Energy and performance monitoring using dashboards.

In the FACTLOG process these issues are going to be approached using DT modelling and cognition as described in deliverable "D1.1 Reference Scenarios, KPIs and Datasets".

By solving these issues, the objective is to improve the values of the KPI as presented in the following table:

KPI	**Today**	**With FACTLOG**
Machine downtime because of breakdowns	>8%	< 5%
Total maintenance costs as a percentage of total operational costs	18%	<12%
Energy consumption of idling machines as a percentage of total energy	>11%	<7%
Overall Equipment Efficiency (OEE)	80%	>87%

4.5.2 Framework description

Related to the cognitive process for Continental use case, we can start from the understanding of the production process (as presented in Figure 15).

The production on this line starts on the Pre-Assembly area where the units from the previous level of production need to be available. Once there is a stock of components, the production can start with the Loading station where we just "Load" the unit to the line. On the first station from Pre-Assy line an in-circuit test is performed where we detect if all the electrical circuits are ok and there is electrical connection between some predefined test points from the unit. This type of test is done on PANEL (multiple single PCB on the same board) level, but the testing is done individually on each single unit in parallel. The next station is a de-panelling machine where the split of the PCB is done. After a Hot and Cold temperature test, the units are ready for the next level of assembly, Final Assembly Line.

On the Final Assembly line, we are defining the final product that will be delivered to the customer, but only if it was a PASS on all the stations from the line. On each process/station from the production line, it is followed the same process in regards with MES connections. Each unit, before entering into a station, is checked according to the status from MES system if the unit is ready to be produced on that equipment. If everything is in order, the unit goes inside the machine where the process starts. The machine is doing the process defined for that workplace and when everything is ready the process is stopped and a checkout is sent to the MES system together with the status:

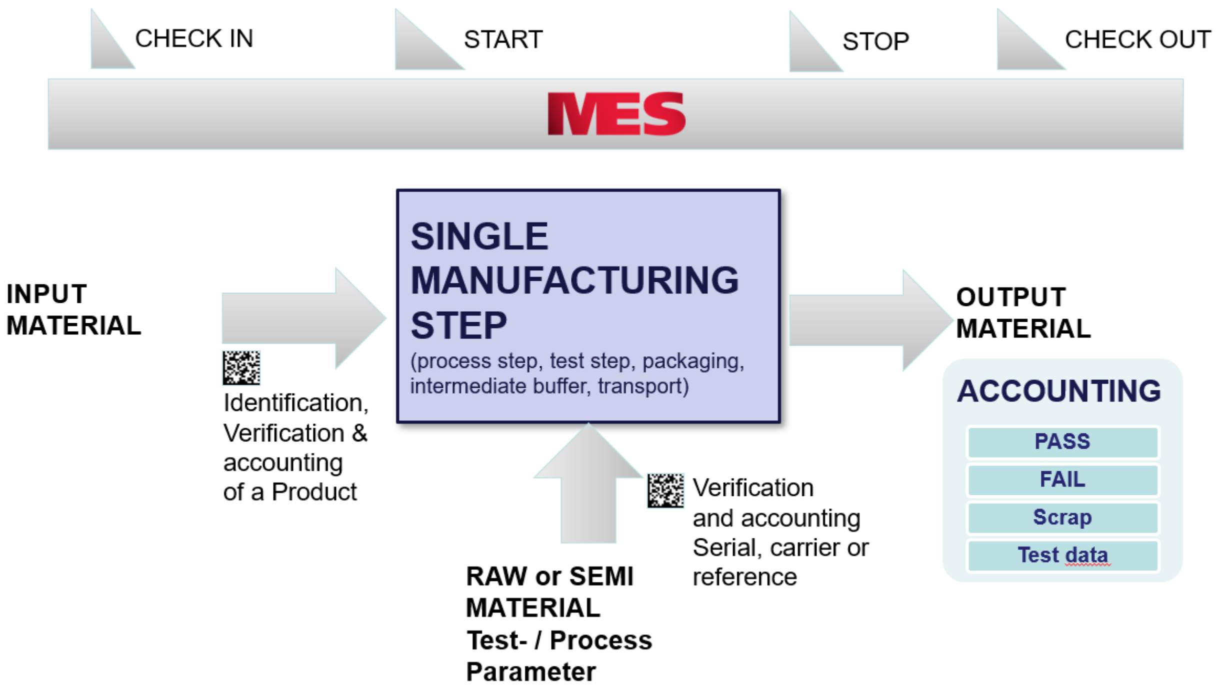


Figure 17: Process description

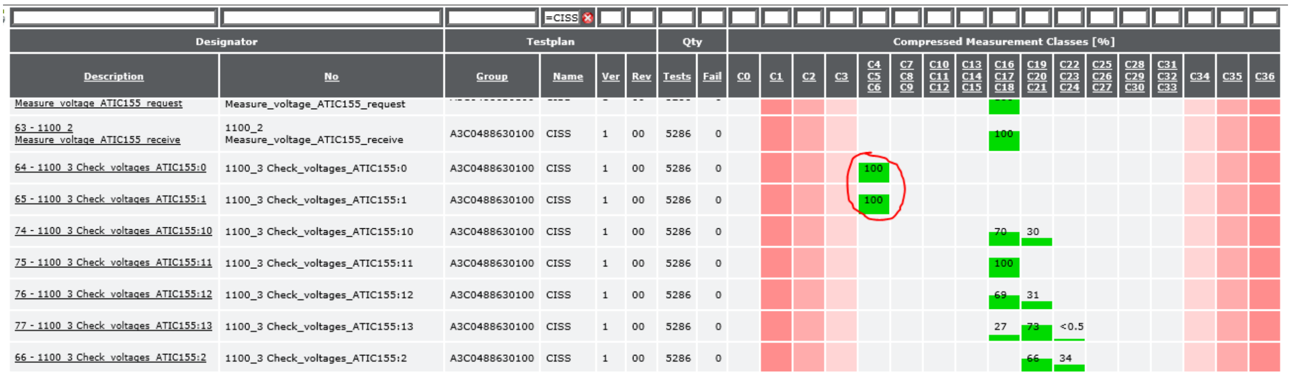
After the information is sent to MES, all the data is also stored in the DB for each parameter. All values are compared with a predefined template and in case the values are out of the limits, then the unit will have the status "FAIL" by the MES system.

TestName	Measured Value	Result	LSL	USL	Format
Force	5.2	P	-9999	9999	R6.2
Distance	55.3	P	-9999	9999	R6.2
Camera2	1	P	1	1	R6.2
Quantity	131	P	-9999	9999	R6.2
RPM	3200	P	-9999	9999	R6.2
Tension	100	P	-9999	9999	R6.2
Frequency	13	P	-9999	9999	R6.2
Intensity	10	P	-55	9999	R6.2
Height1	5	P	-9999	9999	R6.2
Height2	3	P	-9999	9999	R6.2
Height3	6	P	-9999	9999	R6.2
Height4	5	P	-9999	9999	R6.2
Height5	2	P	-9999	9999	R6.2
Height6	5	P	-9999	9999	R6.2
Height7	9	P	-9999	9999	R6.2
Torque1	132	P	-9999	9999	R6.2

Figure 18: Example data

If all the values are between the limits this is not affecting the main KPIs monitored by the production: OEE (Overall Equipment Efficiency), FPY (First Pass Yield). But this does not mean that there are no anomalies with the process. For instance, if all the statuses are "PASS", but the values have a big dispersion or are near to one of the limits, this could be a good sign that something is wrong with the process.

Designator		Testplan				Qty		Compressed Measurement Classes [%]																
			=CISS																					
Description	No	Group	Name	Ver	Rev	Tests	Fail	C0	C1	C2	C3	C4 C5 C6	C7 C8 C9	C10 C11 C12	C13 C14 C15	C16 C17 C18	C19 C20 C21	C22 C23 C24	C25 C26 C27	C28 C29 C30	C31 C32 C33	C34	C35	C36
58 - 0255 27 ATIC255 Init4 receive	0255_27 ATIC255_Init4_receive	A3C0488630100	CISS	1	00	5286	0									100								
59 - 0255 28 ATIC255 Init5 request	0255_28 ATIC255_Init5_request	A3C0488630100	CISS	1	00	5286	0									100								
60 - 0255 29 ATIC255 Init5 receive	0255_29 ATIC255_Init5_receive	A3C0488630100	CISS	1	00	5286	0									100								
129 - 068 1 Check CISS Max Knock1	068_1 Check_CISS_Max_Knock1	A3C0488630100	CISS	1	00	5286	1					<0.5	4	14	26	29	17	7	2	<0.5	<0.5		<0.5	
130 - 068 1 Check CISS Max Knock2	068_1 Check_CISS_Max_Knock2	A3C0488630100	CISS	1	00	5285	1					1	4	14	27	28	17	7	2	<0.5	<0.5		<0.5	
131 - 068 3 Check CISS Max Knock3	068_3 Check_CISS_Max_Knock3	A3C0488630100	CISS	1	00	5284	1		<0.5			<0.5	5	14	27	28	17	7	2	<0.5				
133 - 070 2 calculate area deviation Knock 1	070_2 calculate area deviation Knock 1	A3C0488630100	CISS	1	00	5283	0								50	50								
134 - 070 3 calculate area deviation Knock 2	070_3 calculate area deviation Knock 2	A3C0488630100	CISS	1	00	5283	0								50			50						
135 - 070 4 calculate area deviation Knock 3	070_4 calculate area deviation Knock 3	A3C0488630100	CISS	1	00	5283	0								50			50						
1 - 0996 DTP Check and Read	0996 DTP_Check_and_Read	A3C0488630100	CISS	1	00	18	0									100								
2 - 0996 01 Report DTP Version	0996_01 Report_DTP_Version	A3C0488630100	CISS	1	00	18	0									100								
3 - 0997 DTP Read Variables	0997 DTP_Read_Variables	A3C0488630100	CISS	1	00	18	0									100								



Designator		Testplan				Qty		Compressed Measurement Classes [%]																
Description	No	Group	Name	Ver	Rev	Tests	Fail	C0	C1	C2	C3	C4 C5 C6	C7 C8 C9	C10 C11 C12	C13 C14 C15	C16 C17 C18	C19 C20 C21	C22 C23 C24	C25 C26 C27	C28 C29 C30	C31 C32 C33	C34	C35	C36
Measure_voltage_ATIC155_request	Measure_voltage_ATIC155_request																							
63 - 1100_2 Measure_voltage_ATIC155_receive	1100_2 Measure_voltage_ATIC155_receive	A3C0488630100	CISS	1	00	5286	0									100								
64 - 1100_3 Check_voltages_ATIC155:0	1100_3 Check_voltages_ATIC155:0	A3C0488630100	CISS	1	00	5286	0					100												
65 - 1100_3 Check_voltages_ATIC155:1	1100_3 Check_voltages_ATIC155:1	A3C0488630100	CISS	1	00	5286	0					100												
74 - 1100_3 Check_voltages_ATIC155:10	1100_3 Check_voltages_ATIC155:10	A3C0488630100	CISS	1	00	5286	0									70	30							
75 - 1100_3 Check_voltages_ATIC155:11	1100_3 Check_voltages_ATIC155:11	A3C0488630100	CISS	1	00	5286	0									100								
76 - 1100_3 Check_voltages_ATIC155:12	1100_3 Check_voltages_ATIC155:12	A3C0488630100	CISS	1	00	5286	0									69	31							
77 - 1100_3 Check_voltages_ATIC155:13	1100_3 Check_voltages_ATIC155:13	A3C0488630100	CISS	1	00	5286	0									27	73	<0.5						
66 - 1100_3 Check_voltages_ATIC155:2	1100_3 Check_voltages_ATIC155:2	A3C0488630100	CISS	1	00	5286	0										66	34						

Figure 19: Data dispersion example

In case something like this happens or if these deviations/anomalies appears in the last days/weeks for a specific station/equipment for instance Screwing, this could mean that the process is not stable anymore and could cause trouble even break down in the near or distant future. That's why we need to identify all these outliers automatically and highlight it to the maintenance team.

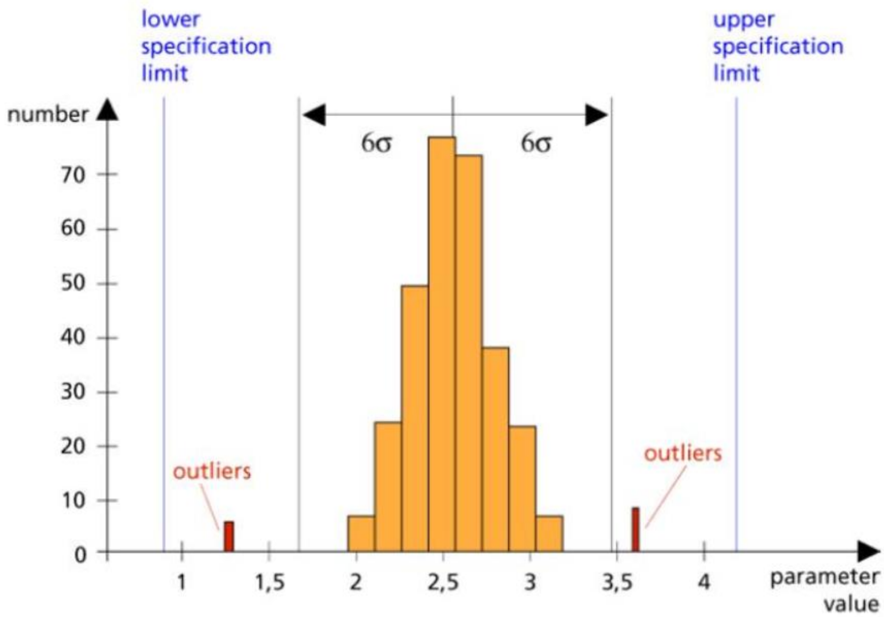


Figure 20: Outliers in data distribution

By this means, the maintenance team could check the status of the process or the "health" of the subassemblies from that station and in case if some damage (for instance, screwdriver shows to be deteriorated) is discovered, the team can plan a change of that subassembly. So, this is useful to plan also the budget for the ordering of the broken subassembly and plan the change of it at the next planned maintenance.

In this regard, we can avoid some breakdowns of the Equipment and all unplanned activities. By having this alarm and the prediction of the possible downtimes the main KPI of OEE can be improved a lot. So, the issues in the line will decrease:

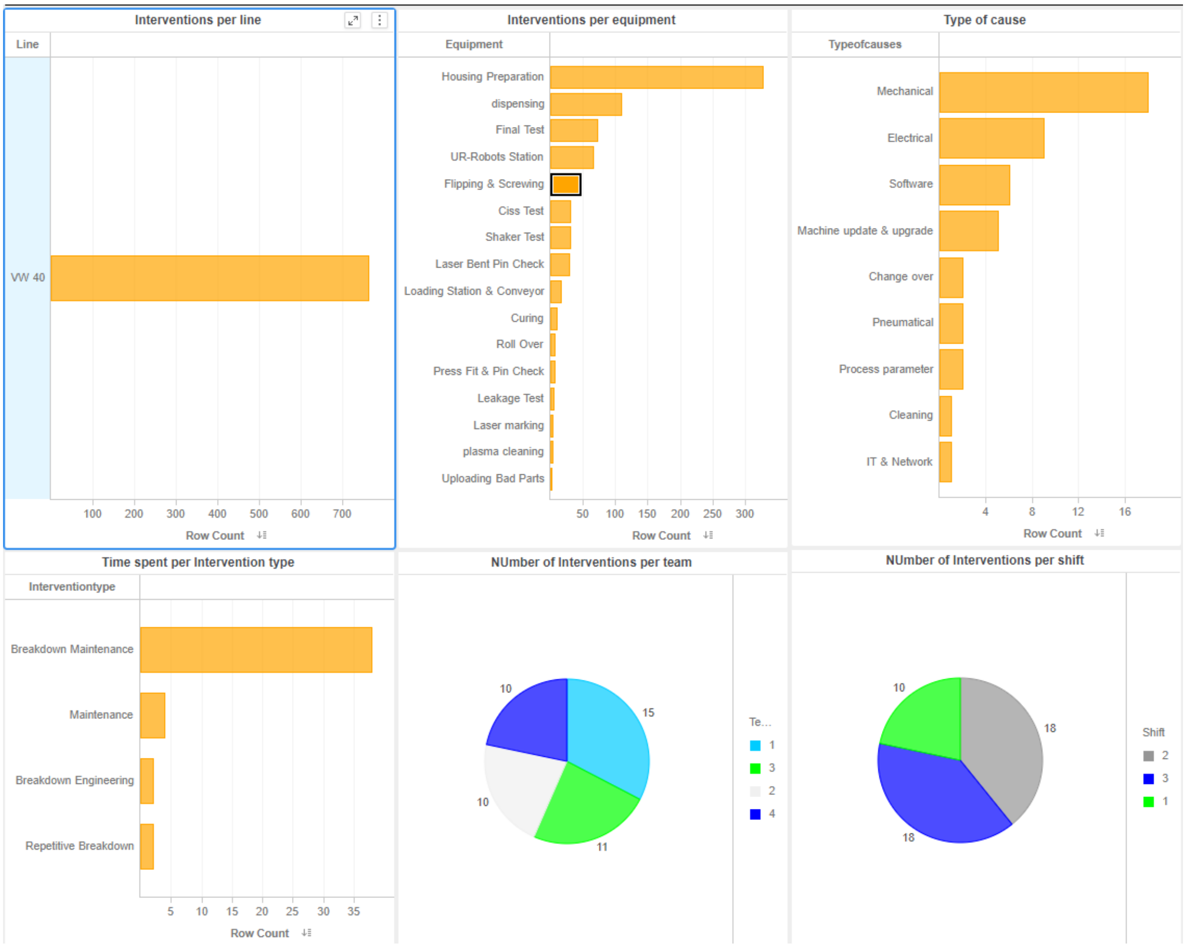


Figure 21: Sample reporting

4.5.3 Cognition cycle

In this section we will present how the information above can be used for applying the cognition framework for a concrete example. The example is about equipment in the Final Assembly Line, the Screwing stage.

4.5.3.1 Detecting variations

The screwing process is monitored using a set of parameters:

- The angle of the screw
- The height of the screw
- The torque applied during the screwing process

The values of the measurements of these parameters are grouped in 36 classes.

The measurements are used to detect anomalies in the screwing process:

- *Failures* - if the measurements are within classes C1, C2, C3, C34, C35, C36, they indicate failures. The product does not meet the required quality requirements.
- *Process instabilities* – the measurements should be distributed following a Bell curve centered on the C16 – C21 classes. From experience it is known that if the curve is flat or it is not centered as stated, an equipment breakdown is likely to happen.

In the report in the next figure there are both type of anomalies. This report represents measurements over 30 days.

Designator		Testplan				Qty		Compressed Measurement Classes [%]																
Description	No	Group	Name	Ver	Rev	Tests	Fail	C0	C1	C2	C3	C4 C5 C6	C7 C8 C9	C10 C11 C12	C13 C14 C15	C16 C17 C18	C19 C20 C21	C22 C23 C24	C25 C26 C27	C28 C29 C30	C31 C32 C33	C34	C35	C36
Angle 1 value	Angle1	A3C0488630100	SCREWING	1	00	9389	3					100	<0.5								<0.5			<0.5
Angle 2 value	Angle2	A3C0488630100	SCREWING	1	00	9389	0					100	<0.5								<0.5			
Angle 3 value	Angle3	A3C0488630100	SCREWING	1	00	9389	4					100	<0.5											<0.5
Angle 4 value	Angle4	A3C0488630100	SCREWING	1	00	9389	5					99	1								<0.5			<0.5
Carrier no	Carrier	A3C0488630100	SCREWING	1	00	9389	0					100												
Height 1 value	Height1	A3C0488630100	SCREWING	1	00	9389	9		<0.5					11	31	11	45	2						<0.5
Height 2 value	Height2	A3C0488630100	SCREWING	1	00	9389	11		<0.5						<0.5	7	72	21						
Height 3 value	Height3	A3C0488630100	SCREWING	1	00	9389	16		<0.5			<0.5	1	12	<0.5		9	66	12					<0.5
Height 4 value	Height4	A3C0488630100	SCREWING	1	00	9389	20		<0.5					3	37	11	45	3						<0.5
Torque 1 value	Torque1	A3C0488630100	SCREWING	1	00	9389	12		<0.5			1	2	1	7	89								
Torque 2 value	Torque2	A3C0488630100	SCREWING	1	00	9389	13		<0.5							88	12	<0.5						
Torque 3 value	Torque3	A3C0488630100	SCREWING	1	00	9389	16		<0.5							90	9							
Torque 4 value	Torque4	A3C0488630100	SCREWING	1	00	9389	22		<0.5							91	9							

Figure 22: Report with anomalies

The lines that contain measurements in classes C1, C36 contain failures. Examples are lines Angle 1, Angle 3, Height 1 etc.

In the figure, the angle measurements are concentrated to the left of the interval. The measurements do not fall in the classes representing errors but they are not centered.

Also the line Torque 1 represents a process instability. The measurements are distributed to the left side of the interval.

Another representation for the measurements for Torque 1 are in the graph in the figure below:

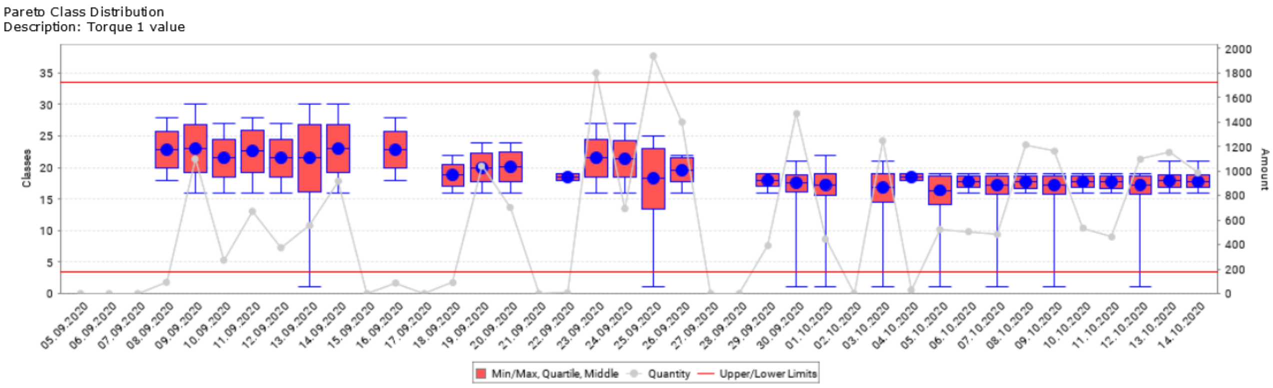


Figure 23: "Torque 1" sensor measurements

In this graph it is visible that the minimum value of the measurements goes consistently below the lower limit after 30.09.

4.5.3.2 Understanding variations

At Continental, the interventions on production lines are documented. The workers record the problems and the cause of the problems.

The equipment from which the measurement presented above were taken had experienced some malfunctions. The interventions records are:

id	74168	74268
Team	1	1
Shift	1	3
Line	VW 40	VW 40
Equipment	Flipping & Screwing	Flipping & Screwing
Product	VW20	VW20
Problem	16.33-16.43 Screwing nok , eroare T+	22.30-23.05 Screwing nok position 1, taken from the previous shift
ActionTaken	Checked bit, stem-ok, cleaned the screwdriver guides, didn't go down enough, tracked-ok	Checked bit-ok, cleaned the screwdriver guides, cleaned the lift bolts, increased the pressure of the of the lowering mechanism of the screwdriver
TimeSpent	15	35
Intervention Type	Breakdown Maintenance	Breakdown Maintenance
DateTime	05.10.2020 18:18	12.10.2020 23:52
TypeOfCauses	Cleaning	Pneumatical
Escalation	No	No
Solved	Yes	Yes

From the descriptions of the actions taken in order to resolve the malfunctions (the line ActionTaken), one can understand that the anomalies presented in the graphs in the previous section can have multiple cause:

- The screwdriver guides collect impurities. They affect the equipment performance and finally lead to equipment malfunction.
- The lowering mechanism of the screwdriver loses pressure and this impacts the screwing quality and finally leads to equipment malfunction.

Having understood in the process the relevant parts that will be involved in taking a new course of action relevant to the maintenance, optimization can be informed in order to take into account the new changes needed in the maintenance activities. Optimization, as a consumer of the output of the cognition module but also given the necessary input (master production plan and desired predictive/scheduled maintenance activities and their time frames) tries to schedule the required maintenance activities so that production throughput is affected as low as possible. As a result, the maintenance planners can use the generated optimized schedules to decide when is the best time to perform the maintenance and how to effectively modify the existing master plan.

In order for optimization to take place there are two different types of data needed from the different FACTLOG modules, Dynamic (per week) and Stationary.

Dynamic data can include: (a) Biweekly detailed master production schedule as well as all production orders, due dates, quantities etc., (b) Machine or (station) maintenance information, that is desired time window and maintenance duration, manually or automatically detected / created. And (c) Selection of the desired objective to optimize against (if multiple) and other optimization parameters exposed to the user. This data can be defined only once if needed.

Stationary data can include: (a) Detailed depiction of the assembly line layout, (b) Detailed description of all processing stages as well as the stations included per stage. Interconnection information between stages and/or stations (e.g. travelling times from stage to stage, capacity of areas to store WIP etc). (c) Processing / Cycle times of orders / jobs on every processing stage, machine, or station. (d) The detailed flow processing from start to the end of the line for each product (e) Detailed description of the optimization objectives and other KPIs that are used to evaluate the efficiency of a production schedule.

5 Conclusion

This deliverable is the foundation for the usage of cognition, as a decision-making method/process, in resolving challenges defined in pilots, which will result in developing cognition-driven solutions.

This foundation is based on the human cognition processes, showing that complex problems can be solved by mimicking human behavior and reasoning.

One of the most important contributions of this work is related to the challenge to human cognition to monitor and detect changes in a very efficient way. Indeed, we structured the cognition process in phases which focus on particular challenges, like detecting variations, discovering the causes, understanding the impact, optimizing the reaction, which can be resolved with selected technologies: data analytics, knowledge graphs, process modelling and simulation, optimization. In that way we developed a very powerful framework, we call Cognitive Factory Framework which is the one of the central concepts of this deliverable

Main conclusion of this deliverable is that this framework can be successfully applied in all pilots, at least from the conceptual point of view. The concept should be fully implemented in WP2-WP5 and applied in WP7 in each pilot.

References

[1] Bonnaud, S., Kohler, A., Didier, C., (2019) Industry 4.0 and Cognitive Manufacturing. Architecture Patterns, Use Cases and IBM Solutions. IBM Corporation, NY, Available Online: https://www.ibm.com/downloads/cas/M8J5BA6R Accessed: 29.09.2020

[2] World Manufacturing Forum (2018) “The 2018 World Manufacturing Forum Report, Recommendations for The Future of Manufacturing” Accessed 20.5.2020 [Online]. Available: https://www.worldmanufacturingforum.org/copia-di-wmf-report

[3] P. Eirinakis et al., "Enhancing Cognition for Digital Twins," 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, United Kingdom, 2020, pp. 1-7, doi: 10.1109/ICE/ITMC49519.2020.9198492.

6 Appendices

6.1 Appendix 1.1 – BRC Digital Signal analysis

			Data Dependancies						
Sensor Data	Description	Possible Failure detection	Required additional data inputs	Drift Error (anomaly)	Possible Causes	Possible outcomes	Step Change Error (anomaly)	Possible Causes	Possible outcomes
This column is divided into to sets of data, analogue (green) and digital (yellow) signals, and describes the actual data being stored from each of the sensors being monitored. So the sensor maybe the Hydraulic cooler on but the data we are measuring is teh time it is active for.	This is a description of the sensor signal being monitored	These are the failures that can be detected on the machine from the different sensor data being collected	These are the additonal inputs that are required to correctly decide whether a change in a sensor signal is an anomaly or just a change in operating mode. For instance if the product dimensions change then the opeating parameters will also change so we will need to consider the product along with the variations in sensor signals	Gradual change in measured sensor signal outside normal operating conditions 	This column is a description of the machine components that could cause the measured sensor data to drift out of normal limits	This column is used to estimate the probable effects of the drift errors/anomalies.	Large change in measured sensor data, normally caused by a major change in the machine operational performance 	This column is a description of the machine components that could cause the measured sensor data to suddenly jump outside of normal limits. Although the causes are often the same as for the drift error the outcome is normally more critical	This column is used to estimate the probable effects of the step change errors/anomalies
Digital Signals									
Safety Relay closed	machine safety circuit is OK and machine can have automatic selected	Machine has been emergency stopped or guard doors are open, possible maintenance actions in progress	NA	NA	NA	NA	NA	NA	NA
Automatic Operation	Machine in automatic mode	Machine on stop	NA	NA	NA	NA	NA	NA	NA
Hydraulic Cooler ON time	Cooling for Hydraulic system is running	Hydraulic system going over temperature if cooler is on for extended periods	Machine in automatic and safety circuit OK.	The time the cooler is on when the machine is running is varying outside normal operational parameters so system is less efficient	1) Hydraulic pump wear, 2) Hydraulic valves not operating correctly. 3) Cooler is inefficient, 4) faulty temperature detector	Hydraulic Oil temperature will go outside safe limits and stop machine over an extended time	Hydraulic oil temperature is increasing rapidly and cooler switches on more rapidly	1) Hydraulic pump wear, 2) Hydraulic valves not operating correctly. 3) Cooler is inefficient, 4) faulty temperature detector	Hydraulic Oil temperature will go outside safe limits and stop machine within a very short space of time. (1 - 2 machine cycles)
Shears Cut timestamp	Shears cut command signal	Failure of shears mechanism	Shears open and closed signals	gradual drift in shears operation time (open and close)	1) Shears jammed, 2) Shear solenoid operation, 3) shear position detection, 4 Hydraulic pump	eventual machine stoppage due to shear operation failing	Step change in shears operation time (open and close)	1) Shears jammed, 2) Shear solenoid operation, 3) shear position detection, 4 Hydraulic pump	machine stops at next shear operation
Shears Open timestamp	Shears are in open position	Failure of shears mechanism	shears cut signal	as for shears cut	as for shears cut	as for shears cut	as for shears cut	as for shears cut	as for shears cut
Shears Closed timestamp	Shears are in closed position	Failure of shears mechanism	shears cut signal	as for shears cut	as for shears cut	as for shears cut	as for shears cut	as for shears cut	as for shears cut
Chiller Overload timestamp	chiller motor in overload condition	Machine hydraulic cooler is in overload state	None	Time between successive overload trips could indicate a issue with the overload or the chiller operation	1) failure of chiller unit 2) failure of overload trip 3) Blocked heat exchanger	Machine stops with overload fault that will require to be reset by maintenance	sudden reduction in time between successive overload trips could indicate a issue with the overload or the chiller operation	1) failure of chiller unit 2) failure of overload trip 3) Blocked heat exchanger	Machine stops with overload fault that will require to be reset by maintenance
Hanger Motor Overload timestamp	Hanger motor in overload condition	Machine hydraulic cooler is in overload state	None	Time between successive overload trips could indicate a issue with the overload or the Hanger motor operation	1) failure of Hanger Motor unit 2) failure of overload trip, 3) excessive load	Machine stops with overload fault that will require to be reset by maintenance	sudden reduction in time between successive overload trips could indicate a issue with the overload or the Hanger motor operation	1) failure of Hanger Motor unit 2) failure of overload trip, 3) excessive load	Machine stops with overload fault that will require to be reset by maintenance
SMU Motor Overload	SMU motor in overload condition	Machine hydraulic cooler is in overload state	None	Time between successive overload trips could indicate a issue with the overload or the SMU Motor operation	1) failure of SMU Motor unit 2) failure of overload trip, 3) excessive load	Machine stops with overload fault that will require to be reset by maintenance	sudden reduction in time between successive overload trips could indicate a issue with the overload or the SMU Motor operation	1) failure of SMU Motor unit 2) failure of overload trip, 3) excessive load	Machine stops with overload fault that will require to be reset by maintenance
Hydraulic Oil level switch change of state time	Hydraulic oil level low level switch	Hydraulic oil level is low, indicating a probable leak	None	Time between level switch operation and reset indicates response time. Number of occurrences in a period indicate level of leak	Oil seepage from pipes or valves	Machine will stop for hydraulic oil low level	sudden reduction between level switch operation and reset indicates response time. Number of occurrences in a period indicate level of leak	1) catastrophic oil leak 2) Oil level switch failure.	Machine will stop for hydraulic oil low level
Cycle Start time	The time when the shears retract and the next bar starts to feed	Machine stopped	machine in automatic	NA	NA	NA	NA	NA	NA

6.2 Appendix 1.2 – BRC Analogue Signal analysis

			Data Dependancies						
Sensor Data	Description	Possible Failure detection	Required additional data inputs	Drift Error (anomaly)	Possible Causes	Possible outcomes	Step Change Error (anomaly)	Possible Causes	Possible outcomes
Analog Signals									
Cycle Time	Time for each bar to be produced	Cycle time varying for similar products. Changes in cycle time for successive operation for same product	Product dimensions from MES for historical comparison	Gradual deterioation of machine operation compared to historical data. Successive operation for same component	1) wear in feed rollers 2) hydraulic motor failing 3) feed speed proportional valve failing or needs calibration, 4) hydraulic pump failing 5) machine setup variation	Machine will stop eventually with possibility of predicting stoppage	Step change in cycle time for successive components, or compared to historical data for same product	1) wear in feed rollers 2) hydraulic motor failing 3) feed speed proportional valve failing or needs calibration, 4) hydraulic pump failing 5) machine setup variation	Machine stops due to failure
Mains frequency	Real number for the frequency of electrical input	Not dependent on machine operation	NA	NA	NA	NA	NA	NA	NA
Mains voltages	The voltage of the power supply going into the machine	Not dependent on machine operation, but used to generate power readings	NA	NA	NA	NA	NA	NA	NA
Mains current	The current of the power supply going into the machine	Not dependent on machine operation, but used to generate power readings	NA	NA	NA	NA	NA	NA	NA
Mains phase current	The individual phase current for the machine supply	Mismatch in phase current demand, possible failure in the hydraulic pump motor	None	hydraulic pump motor phases are becoming unbalanced increasing mechanical stress on motor	motor windings are becoming worn and shorting between windings	motor will eventually fail due to bearing failure	hydraulic pump motor phases are become suddenly unbalanced	motor winding has become short or open circuit	Motor and machine stops immediately
Mains Power	Instataneous operating power for the machine.	energy used for successive cycles can be used to monitor machine performance for the same product type. Historical data will allow performance/ product to be evaluated	Product dimensions and shape code to enable a table of power vs product to be generated	Successive cycles: machine is using more energy to produce same product. Historical data: machine is using more energy to produce same product	1) wear in feed rollers 2) hydraulic motor failing 3) feed speed proportional valve failing or needs calibration, 4) hydraulic pump failing 5) machine setup variation	Machine will stop eventually with possibility of predicting stoppage	Instatneous change in power reading, indicating the hydraulic motor has failed.	motor winding has become short or open circuit, motor circuit breaker trips	Motor and machine stops immediately
Hydraulic pressure	The hydraulic pressure measurment of the system	pressure changes from normal operation indicate machine operational changes	Product dimensional data, Hydraulic oil level switch	Changes in the average hydraulic pressure will indicate the changes in machine efficiency	1) pump wear, 2) pressure transducer unstable, 3) low oil level	a gradual loss in pressure will eventually lead to a total loss of pressure and the machine to stop	Sudden change in hydraulic pressure will indicate a pump or motor failure or a failutre of the pressure transducer	1) pump wear, 2) pressure transducer unstable, 3) low oil level	sudden step change in hydraulic pressure will indicate an imminant failure and machine stoppage
Step time	The time taken between each of the steps in a cycle for example feed,cut, feed,bend, feed, bend, feed and cut	Variations in step time will indicate the changes in individual component operation	Product dimensional data	Step time drift between successive cycles will indicate possible failure of individual components dependant on the step operation. Drift compared to previous product runs will accentuate the comparison	1) Feed roller wear, 2) Bend unit operation 3) Shear operation, 4) Hydraulic pump	Gradual deterioation in operaiting efficiency and loss of production	A step change in the step operation time will indicate a major failure change in the respective component efficiency	1) Feed roller wear, 2) Bend unit operation 3) Shear operation, 4) Hydraulic pump	a step change in the step time will indicate an imminant failure of the respective component.
Step power	The amount of power input for each step taken	Variations in step power/energy will indicate the changes in individual component operation	Product dimensional data	Step power drift between successive cycles will indicate possible failure of individual components dependant on the step operation. Drift compared to istorical product runs will accentuate the comparison	1) Feed roller wear, 2) Bend unit operation 3) Shear operation, 4) Hydraulic pump efficiency	Gradual deterioation in operaiting efficiency and loss of production	A step change in the step operation power will indicate a major failure change in the respective component efficiency	1) Feed roller wear, 2) Bend unit operation 3) Shear operation, 4) Hydraulic pump	a step change in the step power will indicate an imminant failure of the respective component.
Step Feed length	The amount of bar fed in for each step, value is taken from the proportional valve speed command signal for the feed drive integrated to give length	Variations will indicate if the feed rate is stable and possible variations in the proportional valve control	Product dimensional data, feed roller encoder signal	a gradual change in step feed length will indicate that the stability of the drive system is varying	1) Roller wear, 2) proportional valve calibration	variations in product quality and scrap levels	A step change in the feed length signal will indicate a failure of the the speed control loop.	1) encoder feedback, 2) hydraulic drive motor	feed drive will become erratic and product will be out of spec.
Step angle bent	The angle measurement bent in a step, value taken from the bending unit speed command value for the proportional valve integrated to give total angle	Variations in angle will indicate the health of the bending unit operation	Product dimensional data	a gradual change in step angle bend will indicate that the stability of the drive system is varying	1) bending unit failure, 2) feedback signal degredation	variations in product quality and scrap levels	A step change in the bend angle signal will indicate a failure of the the speed control loop.	1) bending unit failure, 2) feedback signal degredation	variations in product quality and scrap levels
Roller wear	The measurment of the wear on machine infeed and straightening rollers	Not possible at present							