DEMONSTRATION REPORT OF FAIRWORK WITH DAI-DSS

D5.3

Editor Name	Alessandro Cisi (CRF)
Submission Date	August 31, 2025
Version	1.0
State	FINAL
Confidentially Level	PU

Co-funded by the Horizon Europe
Framework Programme of the European Union

EXECUTIVE SUMMARY

This document presents the activities carried out within the FAIRWork project to validate the DAI-DSS (Democratic AI–Decision Support System) in real industrial environments. The system was designed to support complex decision-making in manufacturing contexts by integrating artificial intelligence, human expertise, and legacy systems through a modular and transparent architecture.

The demonstration materials can be accessed via the following link:

https://innovationshop.FAIRWork-project.eu/

The demonstration phase involved two industrial sites: CRF and FLEX, each with specific operational goals and challenges. At CRF, the focus was on production planning and human resource allocation, aiming to improve operational efficiency and worker well-being under dynamic production conditions. At FLEX, the activities concentrated on AI-assisted maintenance and automated validation of calibration documents, with tailored solutions for the Althofen and Timisoara sites.

The implemented solutions rely on a hybrid architecture combining cloud-based services with local components, and integrate a variety of AI algorithms, including unsupervised learning, reinforcement learning, rule-based systems, and large language models (LLMs). These tools demonstrated the ability to support transparent, adaptable, and human-cantered decision-making, contributing to improved quality and responsiveness in operational choices.

The results obtained highlight several concrete benefits across all use cases:

- Optimization of planning and resource allocation, with reduced decision-making time and increased flexibility in managing absences and operational changes.
- Reduction of manual workload and errors, thanks to the automation of repetitive processes and the simplification of user interfaces.
- Improved transparency and traceability in decision-making processes, supported by key performance indicators (KPIs) that assess the effectiveness of the proposed solutions.
- Positive validation from end users, through structured testing, usability questionnaires (SUS), and co-creation sessions involving operators, technicians, and managers.

The FAIRWork system positions itself as a strategic tool for promoting fairness, flexibility, and inclusiveness in the workplace, aligned with the FAIR principles (Findable, Accessible, Interoperable, Reusable), adapted to the industrial decision-making context. Its democratic approach to decision support, based on transparency and participation, represents a distinctive feature of the project and a potential reference for the evolution of decision support systems in the European manufacturing sector.

PROJECT CONTEXT

Workpackage	WP5: Demonstration of FAIRWork at Use Case Site
Task	T5.1: Modelling FAIRWork for Production Processes T5.2: Creating FAIRWork Knowledge Base T5.3 Setup DAI-DSS Infrastructure at Use Case Site T5.4 Install DAI-DSS at Use Case Site T5.5 Enrich DAI-DSS with AI-Algorithms T5.6 Demonstrate FAIRWork and Feedback on DAI-DSS
Dependencies	WP2, WP3, WP4, WP6

Contributors and Reviewers

Contributors	Reviewers
Alessandro Cisi, Giulio Vivo (CRF) Roland Sitar, Ciprian Alexandrescu (FLEX) Rishyank Chevuri (JOTNE) Gustavo Vieira (MORE) Robert Woitsch, Marlene Mayr, Damiano Falcioni (BOC) Sylwia Olbrych, Alexander Nasuta (RWTH) Herwig Zeiner (JR)	Rishyank Chevuri (JOTNE) Wilfrid Utz (OMiLAB) Gustavo Vieira (MORE)

Approved by: Robert Woitsch [BOC], as FAIRWork Coordinator

Version History

Version	Date	Authors	Sections Affected
1.0	August 31, 2025	Alessandro Cisi (CRF)	ALL

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Meaning
API	Application Programming Interface
CCDB	Calibration Certificate Database
CDB	Calibration Database
CoC	Centre of Competence
DAI-DSS	Democratic Artificial Intelligence - Decision Support System
DMM	Decision-Making Model
ERP	Enterprise Resource Planning
e-WI	Electronic Work Instruction
FFD	First-Fit Decreasing
HMI	Human-Machine Interface
KPI	Key Performance Indicator
LLM	Large Language Model
MAS	Multi-Agent System
MCTS	Monte Carlo Tree Search
MES	Manufacturing Execution System
MRP	Material Requirements Planning
MSA	Measurement System Analysis
PLC	Programmable Logic Controller
RAG	Retrieval-Augmented Generation
REST	Representational State Transfer
SCADA	Supervisory Control and Data Acquisition
SUS	System Usability Scale
UI	User Interface
URS	User Requirement Specification
VPN	Virtual Private Network
WI	Work Instruction

1 INTRODUCTION

1.1 Purpose of the Document

This document aims to describe and analyse the demonstration phase of the DAI-DSS system at the project's use case sites. It focuses on how the system was operated in real production environments, how AI-based decision-making was introduced, and how the reliability of such decisions was ensured. The report outlines the operational configurations, the challenges encountered, and the improvements observed in decision-making processes involving human operators, robots, and production systems.

The term FAIR is particularly well-suited to the context of WP5 for several reasons. The project aims to create a transparent, accessible, and collaborative decision-making environment where data, models, and human expertise are effectively integrated. The implementation of the DAI-DSS system, with its modular architecture, use of data lakes, access to human experts, and integration of artificial intelligence algorithms, clearly reflects the principles of Findability, Accessibility, Interoperability, and Reusability. Originally developed for data management, these principles prove equally relevant in shaping an intelligent, adaptive, and decision-quality-oriented working environment.

1.2 Document Structure

This document is structured to guide the reader through the various phases of the DAI-DSS demonstration within the FAIRWork project. It begins with an introduction that outlines the objectives of the report and situates it within the broader project context. This is followed by a detailed description of the technical and organizational activities involved in setting up the system, including the integration with existing infrastructures and the preparation of the demonstration environment.

The next section presents the demonstration scenarios developed at the CRF and FLEX sites, offering insights into the human-machine interfaces and the operational workflows adopted. The core of the document is dedicated to the analysis of the results obtained during the testing phase, incorporating both qualitative and quantitative feedback from operators, technicians, and managers. It also includes an evaluation of the reliability of AI-supported decisions and their perceived impact on work processes.

The report concludes with a summary of the main findings, a discussion of the benefits and challenges encountered, and a set of recommendations for the future adoption of the DAI-DSS system in real-world production environments. References are provided at the end of the document, along with an annex listing the tools used throughout the demonstration phase.

For ease of reading and content consistency, the case studies of CRF and FLEX are presented in parallel. This means that the various sections are often divided into two specific subsections

1.3 Change History

This document, in addition to introducing new content, also integrates the material already covered in the previous two deliverables of WP5:

- D5.1 "FAIRWork Knowledge Base at Use Case Site", in which models, data and human experts are identified and described.
- D5.2 "DAI-DSS Infrastructure and Setup Report at Use Case Site" that explains in form of show cases how to setup new infrastructure, how to integrate DAI-DSS into an existing and complex legacy environment and how to approach the human decision-makers and operators.

The main reason for this integration is that the previous documents were produced respectively for month 10 and month 20. Some scenarios have been updated and expanded during the project, making it necessary to revise and update the content developed in the earlier stages.

Some content has been reintroduced with the necessary modifications to ensure completeness and clarity. Several sections have been updated based on previous deliverables: Sections 2.1, 2.2, 5.2, 5.3, 5.6, and 5.7 draw from D5.2, while Sections 3.1 and 3.2 incorporate revisions from both D2.1 and D4.3.

2 INSTALLATION AND CONFIGURATION OF THE DAI-DSS

This section summarizes the technical and organizational preparation activities carried out to set up the test environment. Includes sensor installation, integration with legacy systems, interface configuration, and the use of cloud and local infrastructures for the system's demonstration mode. Keywords and main contents:

- Technical and organizational setup activities
- Integration with legacy systems
- Cloud and local configuration
- Infrastructure and security

2.1 Legacy Data Integration

This section revisits and expands upon the topics covered in D5.2 “DAI-DSS Infrastructure and Setup Report at Use Case Site” concerning the management of data flow from the factory to the user interface.

CRF:

For CRF the main framework is illustrated in the figure below:

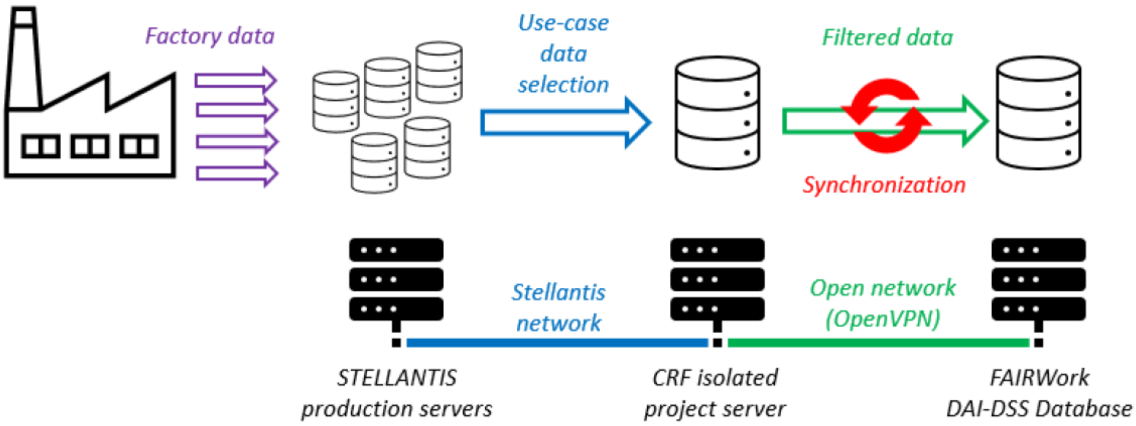

Figure 1 CRF IT Deployment View

Table below illustrates the CRF Use Cases along with the corresponding AI services designed to support decision-makers. The necessary data for these scenarios has been collected, structured, and stored in the Knowledge Base. This data is mapped to appropriate formats within the Knowledge Base, which offers both a graphical user interface (GUI) for manual updates and REST API endpoints for automated data ingestion. One of the integration tests involved updating the operator presence information to validate this functionality.

Use Case	Use Case Partner	Scenario	AI-Enrichment
Workload Balance	CRF	Assist Decisions about Fair Worker Allocation	(1) Support Understanding of Decisions through Conceptual Modelling (2) Decision Support through Decision Tree (3) Resource Allocation using Neural Networks (4) Resource Allocation using Linear Sum Assignment Solver (5) Resource Allocation MAS-based
Production Planning	CRF	Assist Decisions about Production Planning	(6) Production Planning Service with a Hybrid Approach
Delay of Material / Product	CRF	Assist Decisions for Truck Loading	(7) Truck Loading Service

Table 1 CRF Use Cases

This section outlines how data is transferred to the FAIRWork DAI-DSS Knowledge Base.

Due to strict internal IT security policies, external systems are not permitted to access data stored on Stellantis production servers. As shown in the diagram, all systems below the dotted line are restricted to internal access only, either through the company network or via a secure VPN. Access is controlled by user credentials and authorization levels, as some data is considered confidential or strategically sensitive.

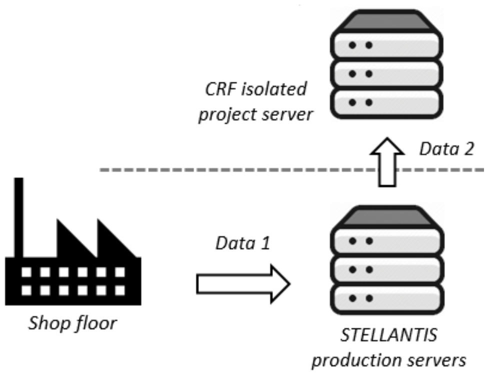

Figure 2 CRF anticipated high level IT structure

The transition of data from internal systems ("Data 1") to external systems ("Data 2") is therefore a critical step. CRF researchers are responsible for selecting and transferring relevant data from Stellantis production servers to a dedicated, isolated server used for the project. Before this data is synchronized with the FAIRWork database, it must be filtered to comply with company policies. For example:

- Employee-related data must be anonymized to prevent identification of individuals.
- Production volume data must be adjusted to avoid disclosing sensitive business information.
- Quality-related data must also be modified to align with internal confidentiality requirements.

The CRF isolated server is connected to an open network specifically designated for the project, allowing external access without major restrictions.

Once the data is uploaded to the CRF isolated server, it must be synchronized with the FAIRWork database. To achieve this, an API will be developed and hosted on the CRF side. This API will compare the contents of both databases and upload any new files to the FAIRWork system.

For the current prototypes, data has been provided in CSV or XLSX formats, and the Knowledge Base has been updated manually. At this stage, the process remains manual.

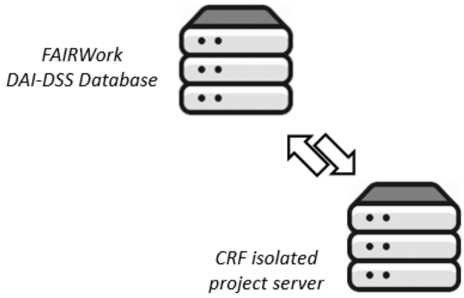

Figure 3 CRF local server to FAIRWork database information exchange

The data flow can be summarized in three main stages:

- Data is originally stored on Stellantis production servers.
- Selected data relevant to the project is transferred to the CRF isolated server via the internal network.
- After filtering, the data is synchronized with the FAIRWork DAI-DSS database over an open network.

As part of the FAIRWork initiative, we developed PresenceFAIRWork, a desktop application designed to simplify and automate the management of Excel-based attendance sheets.

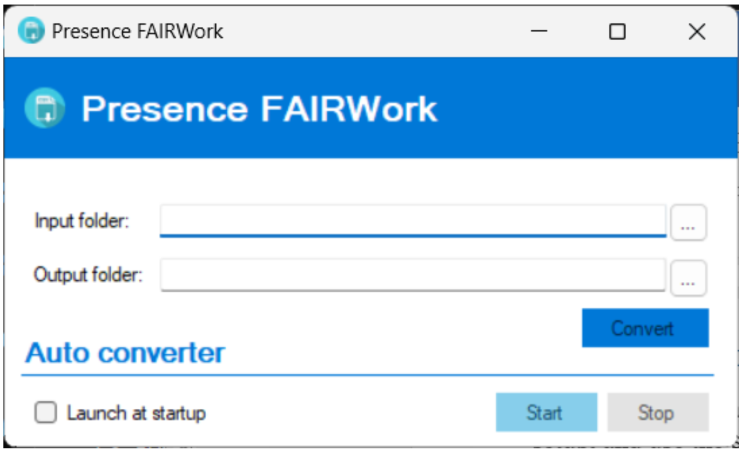

Figure 4 Tool main interface

The main goal was to reduce manual intervention in the transformation and storage of attendance data, while ensuring seamless synchronization with the EDMtruePLM system.

What is Presence FAIRWork: PresenceFAIRWork is a tool that monitors a local folder containing Excel attendance files, automatically converts them into CSV format, and uploads them to a centralized server. This process, which would typically require repetitive manual operations, is now fully automated-improving efficiency and minimizing the risk of errors.

How it Works: Once the application is launched, the user can configure two folders: one for input (where Excel files are stored) and one for output (where the converted CSV files are saved). Additionally, a configuration file (config.json) allows the user to set the necessary credentials and parameters for server synchronization.

The tool offers two main operating modes:

- **Instant Conversion:** The user can manually trigger the conversion and upload process.
- **Auto Conversion:** The application runs in the background, automatically detecting changes in the input folder and handling conversion and synchronization in real time.

There is also an option to enable automatic start-up, allowing the tool to launch silently in the system tray when the computer boots up.

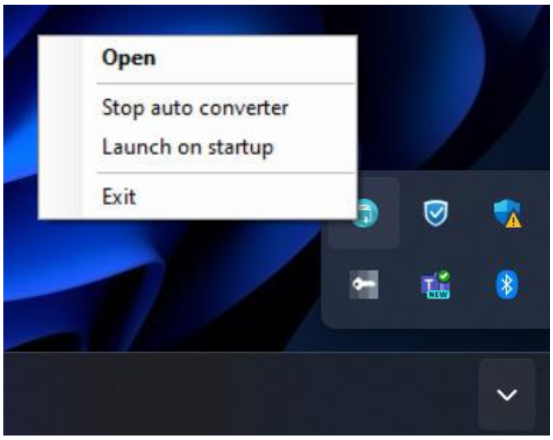

Figure 5 Menu of the application in the system tray

The application was developed in Python (version 3.12.4) and is distributed as a standalone executable (PresenceFAIRWork.exe). Configuration is user-friendly and accessible even to non-technical users, requiring only basic edits to a JSON file and folder selection through the graphical interface.

Additional components include:

- A text file (IdSap.txt) listing the SAP IDs to be monitored.
- A log directory that records all operations performed, useful for monitoring and debugging.

Thanks to PresenceFAIRWork, we have:

- Automated a repetitive and error-prone process.
- Integrated local data with a centralized system.
- Delivered a lightweight, easy-to-use, and configurable tool.

The tool is fully operational and ready for deployment in real-world environments. Its modular structure also allows for future enhancements, such as integration with other systems or the addition of advanced features.

FLEX:

As an initial approach, Flex heads towards implementation of an isolated project server for use-case data sharing. The reason for this approach is to decouple data sharing with the use-case partners and risk mitigation of interferences on the running production IT following with this. Also, Flex Global IT policies regarding data security and integrity, as well Flex's responsibility to its customers and employees do not allow any potential hazard to data integrity.

Once the systems planned and provided by the use-case partners will show a robustness in operation as well as stability in being operated by shopfloor employees, it is intended to modify the interface to allow partial access to live data. Now of report creation, no outlook can be provided if this action can be done within the project duration of FAIRWork or will be followed as extended scope at a later point in time.

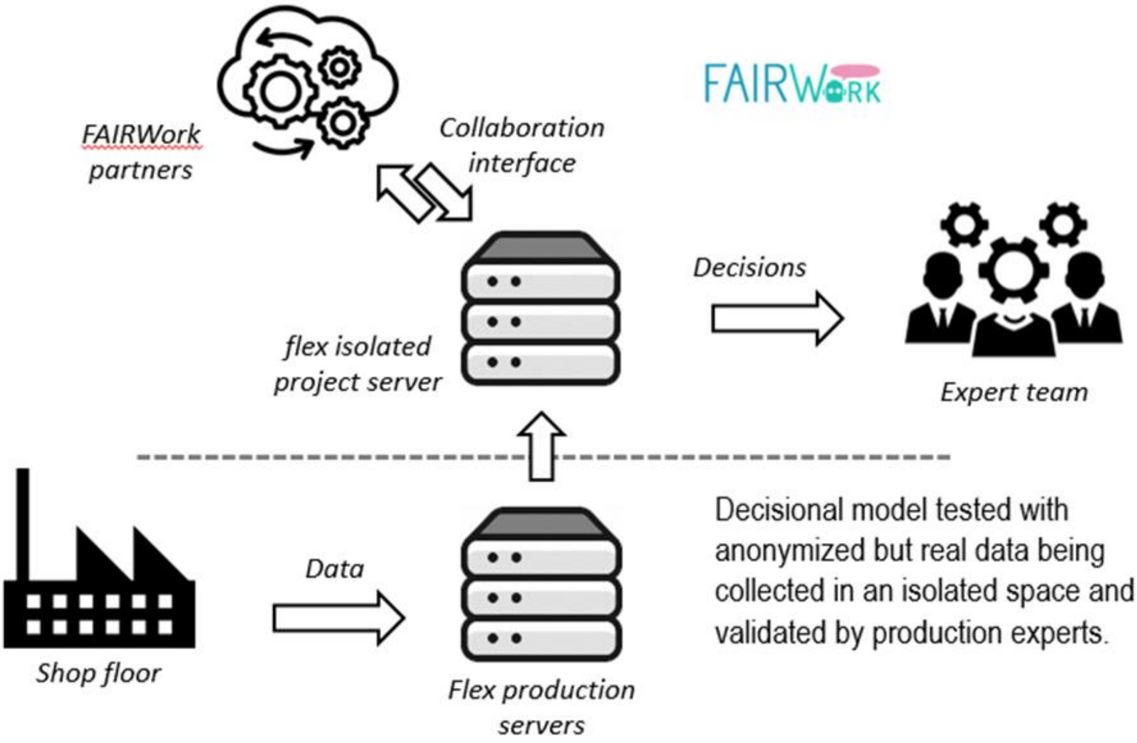

Figure 6 Flex data handling structure

All relevant data is located on the various production servers and to eliminate security risks, the required data is mirrored to a knowledge database server which is completely isolated from the company net.

The knowledge database server contains data from our maintenance software ISPO-NG and the Maintenance Wiki. Furthermore, all machine documents are transferred there.

For our application, we limited ourselves to the surface mount technology area, as this is where the largest amount of data is available.

At Flex Timisoara, we initially adopted a strategy like the one used at Flex Althofen by setting up an isolated project server to facilitate data exchange for the FAIRWork Use Case. This server hosted a replica of the production environment, populated with anonymized data to safeguard sensitive Flex information and ensure customer confidentiality.

The infrastructure was designed to support the original use cases, which focused on Robot Programming, Machine Failure detection, and Workforce issue. To enable secure and efficient collaboration, the isolated server provided controlled access for stakeholders, allowing data sharing without exposing or interfering with actual Flex or customer data.

However, after these initial use cases were cancelled and replaced with new ones—specifically Document Transformation and Calibration Certificate Verification—the need for such a dedicated IT setup was no longer necessary. This shift was made possible because the DAI-DSS prototype could be developed using a smaller volume of anonymized documents, categorized into a few distinct types. This significantly reduced the need for direct access to the production environment database.

As a result, the DAI-DSS prototype was simplified and restructured into a lightweight executable file that can be easily installed and run on any local PC ensuring rapid deployment and ease of use. This streamlined approach not only minimized technical complexity but also aligned more effectively with the revised use case requirements and the existing IT infrastructure at Flex Timisoara.

2.2 Use of the OLIVE Platform in the application context

Within the FAIRWork project, the technology partners developed and deployed a set of decision support services and interactive components within the OLIVE platform, a configuration and orchestration environment based on microservices. OLIVE enables modular composition of user interfaces, configuration of workflows, and integration of AI services, facilitating the adoption of decision support solutions across heterogeneous industrial scenarios.

In the context of the selected case study, the end-user employed only the services relevant to the specific application domain, leveraging the configuration functionalities provided by the platform. The activity was structured as follows:

1. **Access to the Configuration Environment:** The user accessed the OLIVE web interface, which provides centralized management of microservices, workflows, and user interfaces.
2. **Selection of Relevant Services:** AI services and UI components developed by project partners were identified and evaluated based on their functional compatibility with the use case requirements.
3. **Composition of the User Interface:** Through the "Create a Web Application" module, a customized user interface was configured by selecting a predefined layout and integrating the required UI components. Each component was parameterized according to the specific needs (e.g., endpoints, labels, input/output formats).
4. **Execution and Validation of Services:** The configured interface was tested by uploading input data in the required formats (e.g., CSV files, Word documents) and verifying the correctness of the results returned by the services. Where applicable, results were stored in the Knowledge Base.
5. **Interaction with Workflows:** Preconfigured workflows were used to orchestrate the interaction between the user interface, AI services, and the Knowledge Base, leveraging the REST APIs exposed by the platform.
6. **Adaptation to Case-Specific Data:** Input data were adapted to match the formats required by the selected services. In some cases, configuration parameters were adjusted to reflect the operational specifics of the application context.

Below is an example of a configuration using the Olive platform to deploy the user interface of the demonstrator related to the Production Planning and Resource Allocation Service.

When the web page opens, the screen related to "Scenario Management" is displayed. The first step is to click on the icon labeled "***Create a WEB Application***":

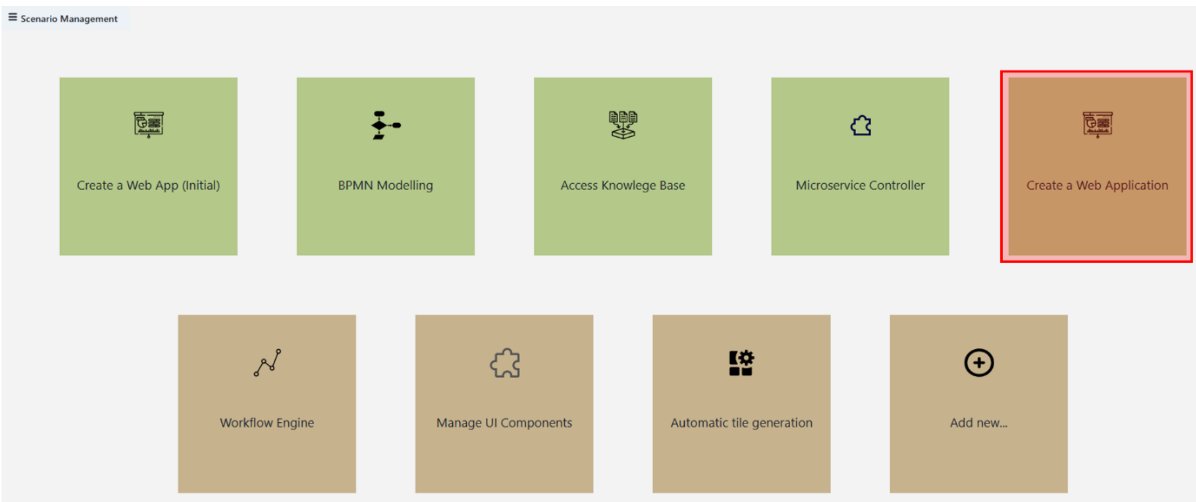

Figure 7 Scenario Management web page

At this point, you are asked what type of application you want to create. In our case, the correct icon to select is "***FAIRWork UI***":

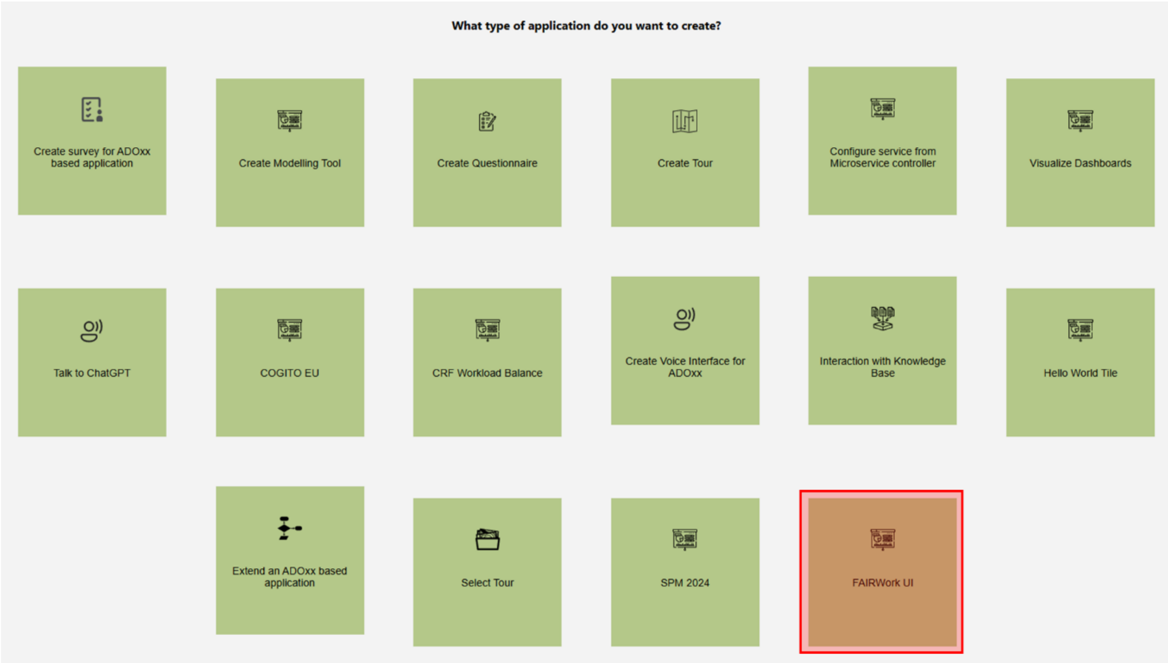

Figure 8 Application Type Creation web page

The next step is to select the correct service, from this point the application changes depending on use cases.

CRF:

In CRF case, "***Production Planning***". Once selected, it appears on the right side of the screen. To proceed, click on "***UI Builder***" written at the top of the window title.

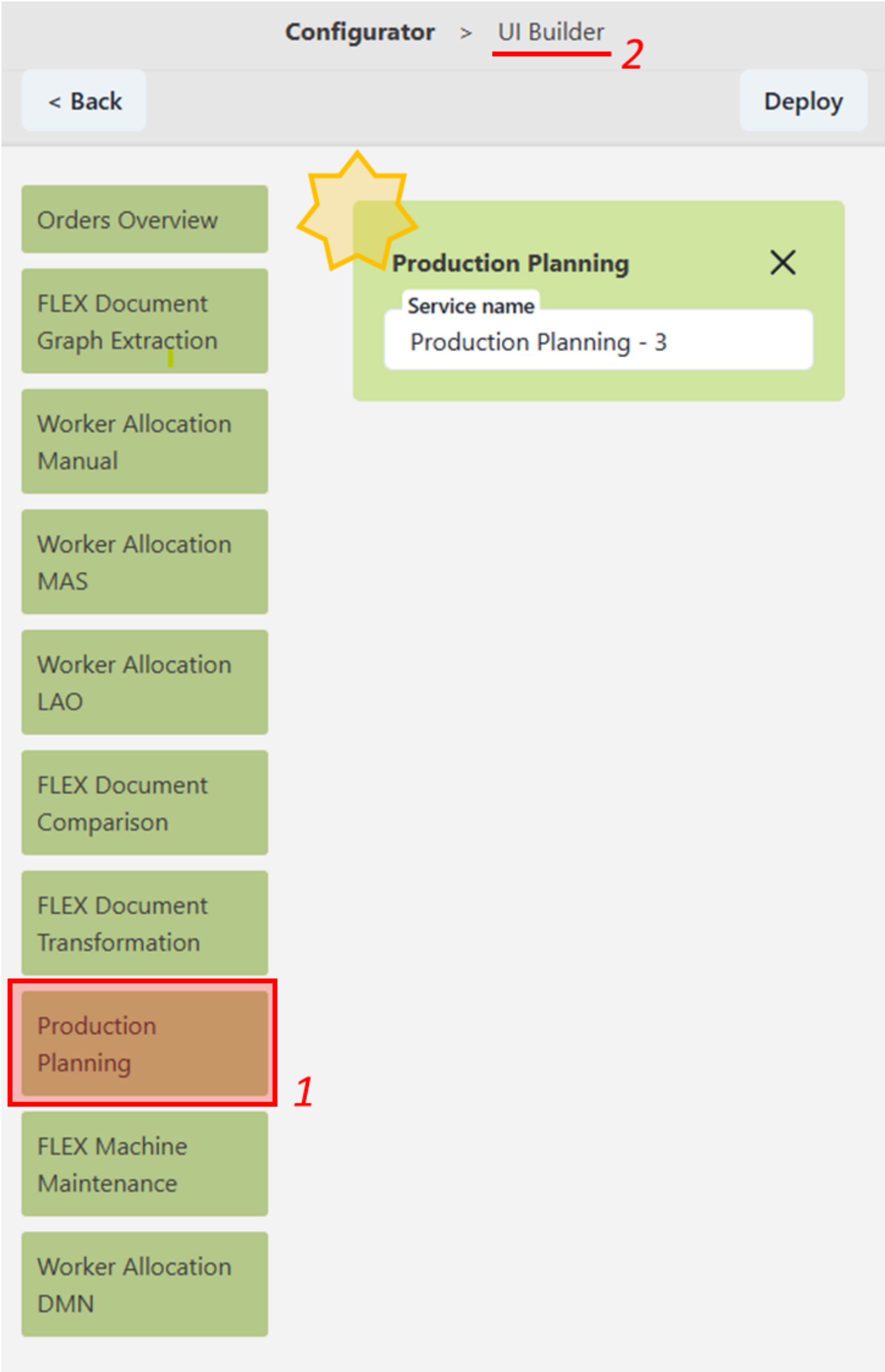

Figure 9 Service selection web page

The final step is to drag the "***Production Planning***" icon into the designated area and click on "***Deploy***" written at the top right of the window:

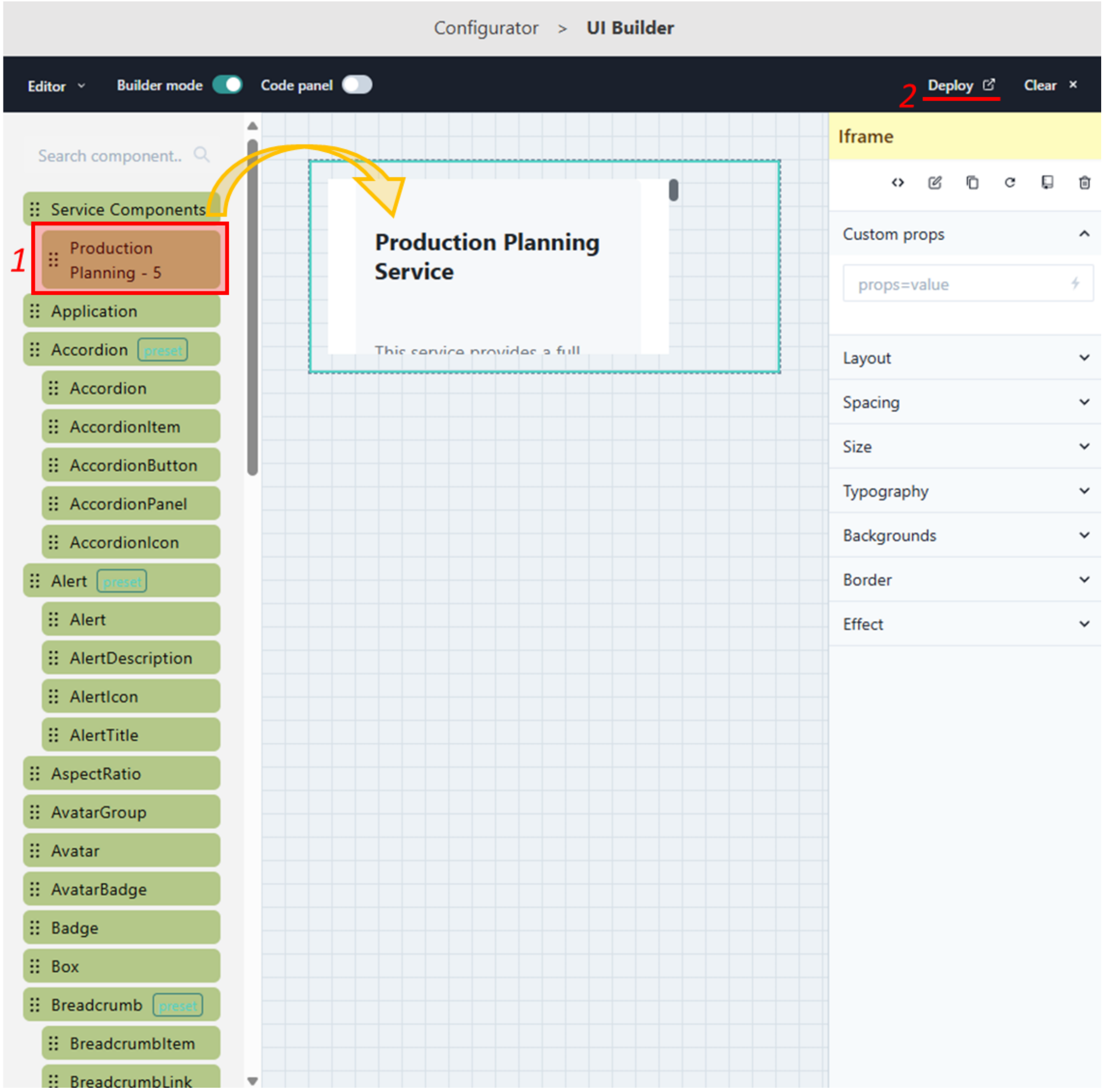

Figure 10 CRF - Application Deployment web page

At the end, a confirmation window appears. By clicking on "***Deploy***", at the top of the screen, the link to the application's web page is displayed.

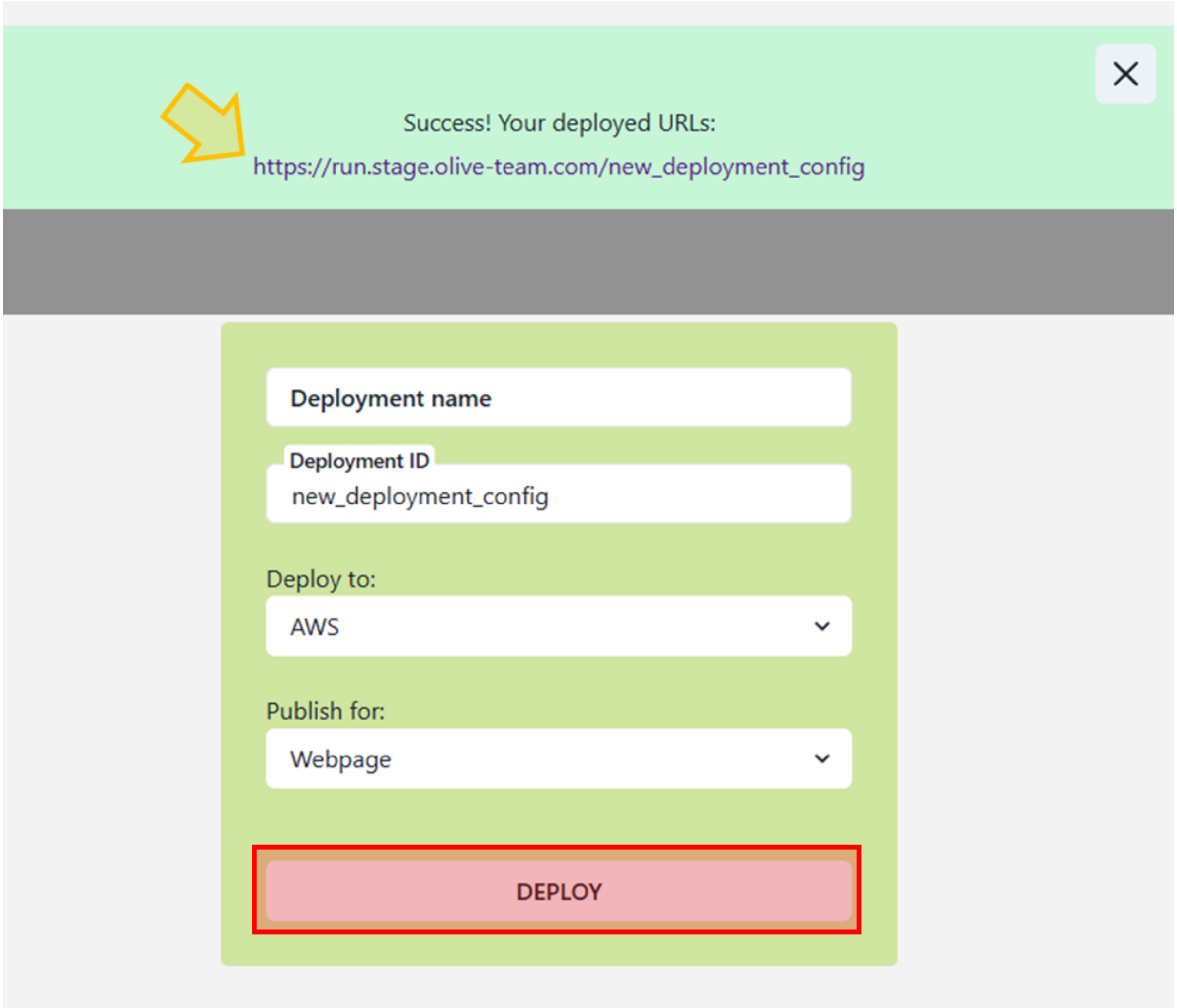

Figure 11 Deployed URLs web page

The link generated in this way is ready to be used. The user interface created will be described in detail in the dedicated chapter about Demonstration Scenarios.

FLEX:

In the case of the Flex Althofen use case, in Figure 9, the "***FLEX Machine Maintenance***" button is selected in the menu on the left. In this case, the Application Deployment web page appears as shown in Figure 12. The following steps are entirely analogous to the previous case, up to the final "***Deploy***" phase. However, in this case, the generated link refers to the FLEX Althofen case study.

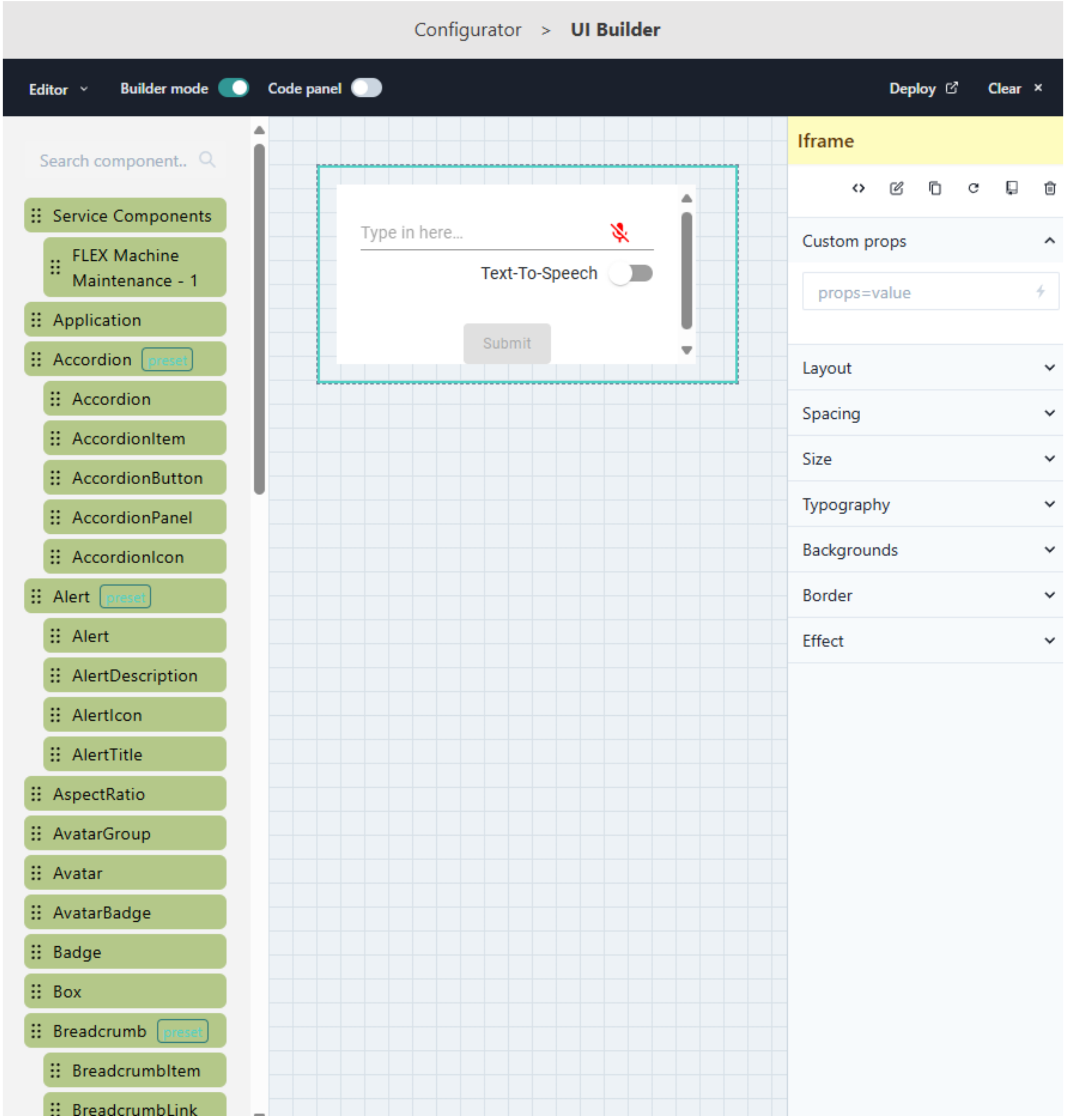

Figure 12 FLEX - Application Deployment web page

Also in this case, the user interface created will be described in detail in the dedicated chapter about Demonstration Scenarios.

At Flex Timisoara, a dedicated service was created under the Olive platform exclusively for the "Document Transformation" use case. In contrast, the DAI-DSS service developed for the "Calibration Certificate Verification" use case was implemented as a standalone executable file that can run on any local PC, and therefore did not require integration into the Olive platform.

As illustrated in Figure 9, titled "Service Selection Web Page," selecting the "FLEX Document Transformation" option from the left-hand menu navigates the user to the Application Deployment interface, shown in Figure 13.

The subsequent steps leading up to the final "Deploy" phase follow the same procedure as those outlined for the other use cases previously described in this section.

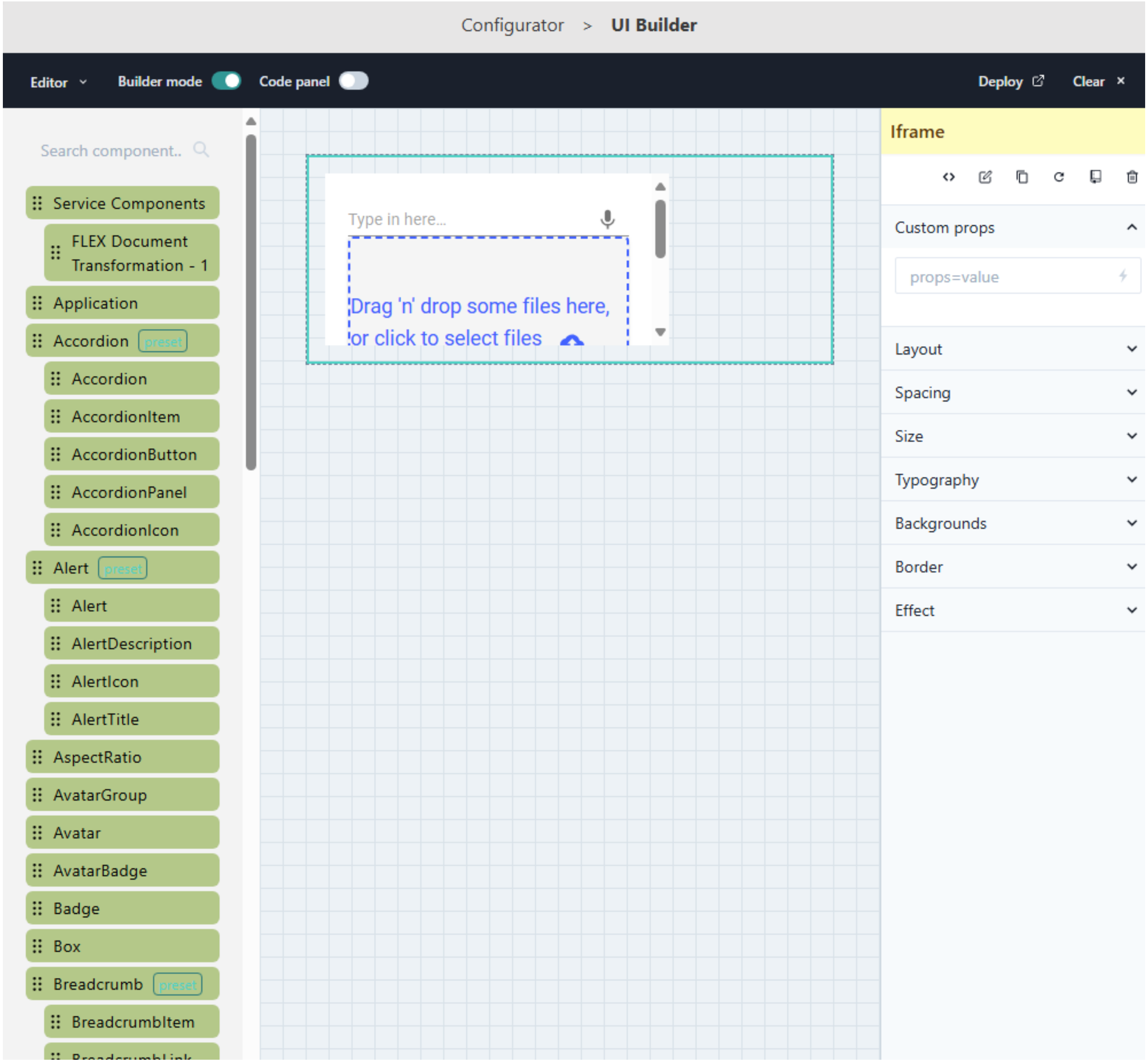

Figure 13 FLEX – Document Transformation - Application Deployment web page

The same service could be accessed directly from the following link:

3 SCENARIOS EVOLUTION

This chapter describes the initial selection of scenarios for CRF and FLEX, with the technical and operational motivations behind the choices made. It presents the final selection of scenarios, highlighting the changes made compared to the initial planning and the reasons behind them.

In the context of the FAIRWork project, the selection of decision-making scenarios for the CRF and FLEX use cases represented a crucial phase for the implementation and validation of the Democratic AI-based Decision Support System (DAI-DSS). This system, designed to support complex decisions in cyber-physical production environments, is based on the integration of intelligent agents, decision models, digital twins, and artificial intelligence algorithms, with the aim of promoting cooperative, transparent, and adaptive decision-making.

Table below, extracted from the Technical Description of the Proposal (Part B), summarizes the scenarios identified at the pilot sites. Each scenario is also associated with one or both use cases (CRF and FLEX), highlighting the initial intervention priorities.

Decision	Level	Interval / Time	Implication & Risks	CRF	FLEX	Prototype
Re-Arranging Press shop	Strategic, Executive	Once every several years	Costs, time, throughput of production process and ROI	5		Skipped (low priority)
Workforce Issues	Operation, Tactic	Several per Year	Lack of workforce		4	Covered by CRF
Machine Maintenance	Operation, Tactic	Several per Month	Reduced Lifespan, higher risk on quality issues, higher risk on machine failure		3	Covered by FLEX
Robot Programming	Operation, Tactic	Several per month	Biased Acceptance, Suitable / not suitable robot scenario		1	Replaced by Calibration Certification (FLEX)
Quality issue	Operation	Some per Month	Waste, high costs due to re-scheduling	3		Skipped (priority lowered to 4)
Production Planning	Operation	Weekly	Production plan for week, shift plan for week, Maintenance plan for week	4		Covered by CRF (priority raised to 3)
Delay of Material / Product	Operation	Several per week	Ad hoc change of work allocation, extension of logistic actions, re-scheduling of production	2		Covered by CRF
Machine Failure	Operation	Several per Week	Raised costs, reduced throughput		2	Covered by FLEX
Workload Balance	Operation	Several per day	Biased workload, delay, Inefficient maintenance	1		Covered by CRF

Table 2 Identified possible Decisions at Use Case Site

The selection of scenarios was guided by a structured methodological approach aimed at ensuring maximum industrial relevance and technical feasibility. In particular, the following general criteria were considered:

1. Decision level: scenarios were classified according to the level at which decisions are made (strategic, tactical, operational), with the goal of covering a representative range of complexity and time horizons.
2. Frequency and urgency: initial priority was given to scenarios characterized by high frequency and the need for rapid response, to maximize the system's impact on daily operations and facilitate integration into existing processes.
3. Implications and risks: the potential impacts of decisions were assessed in terms of cost, quality, safety, energy efficiency, and operator well-being, prioritizing scenarios with greater systemic relevance.
4. Data availability and quality: the selection considered the possibility of accessing reliable and meaningful data, a necessary condition for the effective training and operation of AI algorithms.
5. Interest and priorities of industrial partners: the choice was strongly influenced by direct feedback from CRF and FLEX partners, who helped identify the most critical and promising areas for system application.

During the project, the selection of scenarios underwent changes and adaptations, reflecting the evolution of industrial needs, the technological maturity of the developed solutions, and feedback collected during co-creation and experimentation phases. Some initially planned scenarios were replaced or reformulated, while others were introduced at a later stage to address new opportunities or emerging issues. The updated list of prototypes is:

Use Case	Use Case Partner	Use Case short name (internal use in consortium)	Prototype
Assist Decisions about Fair Worker Allocation	CRF	Workload Balance	Decision Support through Decision Tree; Resource Allocation using Neural Networks; Resource Allocation using Linear Sum Assignment Solver; Resource Allocation MAS-based
Assist Decisions about Production Planning	CRF	Production Planning	Production Planning Service with a Hybrid Approach
Assist Decisions for Truck Loading	CRF	Delay of Material / Product	Truck Loading Service
Improve Information Access to Support Maintenance	FLEX	Machine Maintenance	Support Machine Maintenance using RAG and LLM
Improve Reliability of "Documentation about Quality Check"	FLEX	Document Transformation	Document Transformation using LLM
Improve Information Access to Cleanroom Compliance Requirements	FLEX	Compliance with Clean Room Regulations	Support Compliance for Clean Room using RAG and LLM
Support Validation of Calibration Documents	FLEX	Calibration Certificate	Calibration Certificate Service

Table 3 Overview of FAIRWork Use Cases

3.1 CRF - Scenario Selection Evolution

In line with the general selection criteria defined in the introductory section (decision level, frequency and urgency, systemic impact, data availability, and relevance for industrial partners) three priority scenarios were identified for the CRF use case: Workload Balance, Delay of Materials, and Quality Issues. The initial selection of these scenarios is consistent with the classification proposed in the table of the Technical Description of the Proposal (Part B), which maps scenarios according to three key dimensions: decision level (strategic, tactical, operational), decision frequency, and expected impact.

Workload Balance: The workload balancing scenario was selected for its operational relevance and the frequency with which it occurs in the CRF production context. As highlighted in D2.1, balancing activities across different stations and operators is crucial to ensure efficiency, reduce downtime, and improve worker well-being. This scenario lies between the tactical and operational levels, with a daily frequency and high impact on productivity and ergonomic safety. Moreover, the availability of historical and real-time data on work cycles, execution times, and task assignments made this scenario particularly suitable for the implementation of AI-based decision models. In the table, this scenario is classified as a high-priority target for initial deployment.

Delay of Materials: Delays in material availability represent one of the main sources of inefficiency and disruption in production processes. This scenario, described in D4.3 as one of the use cases addressed by the DAI-DSS prototype, was selected for its high frequency and the systemic impact it can generate across multiple production stages. It is a tactical decision, with a weekly or event-driven frequency, and medium-to-high impact. Proactive management of delays through decision support enables activity reorganization, resource rescheduling, and minimization of idle times. The availability of logistics and material tracking data enabled effective modelling of this scenario. This scenario is also identified in the table as highly relevant for initial experimentation.

Quality Issues: Issues related to product or process quality were included among the initial scenarios for CRF as they represent a critical area both economically and reputationally. As reported in D2.1, timely detection and structured management of quality problems can greatly benefit from the support of an intelligent decision system. This scenario is situated at the tactical-strategic level, with variable frequency (from daily to monthly) and high impact, particularly in terms of scrap reduction, traceability, and compliance with customer requirements. Data from sensors, inspections, and non-conformity reports provide a solid foundation for analysing and predicting quality issues. The table includes this scenario among those with high added value for the adoption of the DAI-DSS system.

This initial selection made it possible to cover a diverse range of decisions, with different levels of complexity, frequency, and impact, laying the groundwork for effective and representative experimentation of the DAI-DSS system in the CRF context.

During the FAIRWork project, the selection of scenarios for the CRF use case was revised and updated in response to evolving industrial needs, feedback gathered during co-creation activities, and discussions held at key events such as the Consortium Meeting on May 2, 2024, in Turin. During the “Board Meeting presentation and discussion” session, CRF management expressed the need to reorient certain scenarios to better align with operational and strategic priorities.

Compared to the initial selection, some scenarios were confirmed while others were replaced or refocused, maintaining consistency with the general selection criteria (decision level, frequency, impact, data availability, and industrial relevance), as defined in the table of the Technical Description of the Proposal (Part B), which classifies scenarios according to decision level, temporal frequency, and risk/criticality.

Workload Balance (confirmed): The workload balance scenario was confirmed as a central element of the experimentation. As described in D2.1 and detailed in D4.3, it concerns the optimal distribution of operators across production lines, considering ergonomic constraints, skills, individual preferences, and resilience. This is a tactical-operational level scenario, with daily frequency and high impact on productivity, safety, and worker well-being. The availability of structured data (shifts, profiles, medical certifications, etc.) and the possibility of integrating AI algorithms and multi-agent systems make this scenario particularly suitable for DAI-DSS experimentation.

Production Planning (new, replacing Quality Issues): The Quality Issues scenario, initially selected, was removed from the experimentation for two main reasons:

- Modelling difficulty: as emerged from discussions with CRF experts, quality problems are often the result of process deviations and do not require explicit decisions, but rather standardized corrective actions.
- Low decision frequency: quality-related decisions do not occur regularly and are difficult to anticipate or simulate in a structured decision-making context.

In its place, the Production Planning scenario was introduced, at the direct request of CRF management. This scenario was deemed more relevant and synergistic with Workload Balance, as production planning is closely linked to human resource allocation: to produce a specific geometry, the appropriate personnel must be available, and in case of absences or unavailability, rescheduling is necessary. This is a tactical-level scenario, with weekly or event-driven frequency, and high impact on operational continuity and overall efficiency. As described in D4.3, the proposed solution integrates Constraint Programming and Reinforcement Learning techniques to generate adaptive and resilient production plans.

Truck Loading (refocusing of Delay of Materials): The Delay of Materials scenario, initially focused on managing incoming materials (Metal sheets, lubricants, dies and equipment, auxiliary components, etc.), was refocused into the Truck Loading scenario, which instead addresses the management of outgoing materials, i.e., components to be loaded onto trucks and shipped to customers. This refocusing reflects a greater operational criticality in the outbound logistics phase, where punctuality and shipping efficiency are essential to ensure supply chain continuity and customer satisfaction.

The priority of this scenario was explicitly indicated by the plant management and emerged during one of the periodic project progress meetings between CRF and the Press Shop, confirming its operational relevance. As highlighted in D4.3, the Truck Loading scenario was modelled as a combinatorial optimization problem (bin packing), aiming to maximize truck load efficiency and minimize the number of trips. This is an operational-level scenario, with daily frequency and medium-high impact, consistent with the categories in the table.

This final selection reflects a pragmatic and strategic evolution of the experimentation, allowing the project objectives to be better aligned with the real needs of the CRF context, while maintaining methodological consistency and coverage of the main decision-making dimensions.

3.2 FLEX - Scenario Selection Evolution

Due to organizational changes within Flex, including those in Flex Althofen and Flex Timisoara, as well as required staff adjustments and shifts in operational priorities, it has become essential to re-evaluate the initially proposed use cases to prevent any dropouts from this project.

Thus the Use cases: Robot Programming covering the decision support for the selection of different robotic application scenarios; Machine Failure interpretation of live data from various machines to predict maintenance and, Workforce Issues (actual covered by CRF) where cancelled.

AI decision support for maintenance activity was expanded in Flex Althofen to include decision support for operators in case of machine breakdown while in Flex Timisoara we focus on document verification especially for Calibration Certificate and Transformation File in case of Work Instruction.

In this content the initial use-case:

"Robot Programming": which focused on decision support for selecting various robotic application scenarios as outlined in Deliverable D5.2, was discontinued. The initiative could not be carried forward due to staffing constraints resulting from the restructuring of the Robot Department in Althofen, very low quantity of available data as well as the lack of data that is foreseen to be generated in the near future lack of scenario variations.

"Workforce issue": relates to the challenges associated with the allocation of workers particularly in situations where there are gaps or the need for rebalancing to meet production demands. This approach was not explored any further and encountered pushback within the Flex organization due to the necessity of handling highly sensitive employee information but was takeover by CRF in there "success story" with Worker allocation.

"Machine Failure": focuses on maintenance and describes a breakdown causing standstill times, which should be reduced by support of the DAI-DSS which was split in two parts:

1. Decision support of the operator before contacting a technician of the maintenance staff base on the history of interventions from engineering breakdown data base
2. Interpretation of live data from various machines to predict maintenance and, therefore, reduce or increase the time between maintenance or exchange of components.

Due to organizational and economic constraints, as well as the unstructured nature of the engineering breakdown database at the Timisoara site, it was decided that this would be jointly tackled with the "AI Support for Machine Maintenance" initiative and handled by Flex Althofen.

"Machine Maintenance" with the primary motivation to help less experienced employees of the maintenance staff resolve maintenance tasks and machine downtimes more quickly was extended to decision support of the operator before contacting a technician of the maintenance staff in case of machine breakdown, referred further as: "**Support Machine Maintenance"**

"Compliance with Clean Room Regulations" was started but due to low internal relevance not further proceeded.

To stay aligned with FAIRWork's timeline and deliverables, FLEX redirected their efforts toward more practical use cases. These were chosen for their clearer scope, feasibility, and potential for measurable impact within existing constraints. This shift helped optimize resources and reinforced Flex commitment to delivering meaningful contributions to the project.

1. **Support Machine Maintenance/ Machine Failure**

   Objective:

   To empower less experienced employees to independently identify and resolve operational issues using an AI-driven support tool, thereby reducing reliance on senior staff and streamlining maintenance processes.

   Benefits:

   - **Increased Autonomy**: Employees are guided through problem-solving via targeted AI-generated questions, enabling them to resolve issues without external assistance.
   - **Reduced Maintenance Tickets**: Minor issues can be resolved directly by operators, decreasing the volume of service tickets submitted to the maintenance team.
   - **Enhanced Maintenance Planning**: The AI tool provides clear, step-by-step instructions for executing planned maintenance tasks, improving consistency and efficiency.
   - **Training Support**: The system can serve as a training aid for new employees, helping them learn procedures and troubleshoot common problems through interactive guidance.

   Challenges:

   - **Current Dependency:** Less experienced staff currently rely heavily on support from senior employees, which limits scalability and slows down operations.
   - **Adoption Curve:** Employees may require time and support to fully trust and utilize the AI tool effectively.
   - **Scope of AI Assistance:** Ensuring the AI can accurately distinguish between minor and complex issues to avoid misclassification and unnecessary delays.

2. **Document Transformation using LLM**

   Objective:

   To design and deploy a smart application powered by the DAI-DSS algorithm that automatically converts legacy paper-based Work Instruction (WI) documents into the standardized digital format required by the

Flex e-WI system. This transformation aims to streamline documentation workflows and unlock key operational advantages.

Benefits:

- **Paperless Operations:** Eliminates physical documentation, supporting sustainability and reducing clutter.
- **Improved Revision Control:** Minimizes errors caused by outdated instructions by ensuring consistent version tracking.
- **Faster Approval Cycles:** Enables electronic signatures for quicker validation and deployment of updated WIs.
- **Reduced Manual Effort:** Automates the conversion process, freeing up valuable engineering resources.
- **Lower Error Rates:** Decreases the likelihood of human mistakes during manual formatting and data entry.
- **Accelerated Rollout:** Speeds up the overall transition to digital standards across workstations and product lines.

Challenges:

- **High Structural Variability:** Existing WIs are tailored to specific workstations and product versions, resulting in diverse formats and content structures that complicate automated processing.
- **Inconsistent Formatting Practices:** Despite a shared legacy template, individual engineers applied varied styles and conventions to text and images in .doc files, making reliable parsing and transformation difficult.
- **Complex Content Mapping:** Translating nuanced procedural details into standardized digital formats requires intelligent interpretation and contextual understanding by the AI system.

3. **Calibration Certificate Service**

Objective:

Develop an intelligent system leveraging the DAI-DSS algorithm to automatically review calibration certificates. The system will detect missing critical data and failed calibration results, aiming to streamline validation, reduce manual workload, and enhance compliance oversight.

Benefits:

- **Automated Content Validation:** Ensures certificates meet required standards without manual inspection.
- **Significant Time Savings**: Reduces the need for repetitive manual checks, accelerating review cycles.
- **Minimized Human Error:** Detects omissions and inconsistencies in critical fields, improving data accuracy.
- **Improved Compliance Tracking:** Enhances traceability and audit readiness by ensuring complete and correct documentation.

Challenges:

- **Template Diversity**: Calibration certificates vary widely in format and structure across equipment types and vendors.
- **Field Position Variability**: Even within the same certificate category, key data fields may appear in different locations or layouts, complicating automated parsing.
- **Semantic Ambiguity**: Field labels and terminology may differ subtly, requiring intelligent interpretation to ensure accurate validation.

4 DEMONSTRATION SCENARIOS

This chapter provides an overview of the AI services, where a well-developed version of the prototype, with performance enhancements or partial automation, has been implemented, and real-world testing and demonstration are carried out by the use case partners.

AI-service	Maturity Level	Type of AI	Reliability
Support Understanding of Decisions through Conceptual Modelling	Basic	Rule-based	High
Decision Support through Decision Tree	Basic	Supervised Machine Learning	High
Resource Allocation using Neural Networks	Intermediate	Artificial Neural Network	Medium-High
Resource Allocation using Linear Sum Assignment Solver	Intermediate	Optimization Algorithm	High
Production Planning Service with a Hybrid Approach	Advanced	Reinforcement Learning, Constraint Programming	Medium-High
Resource Allocation MAS-based	Intermediate	Multi-Agent System	Medium-High
Truck Loading Service	Basic	Optimization Algorithm	Medium
Support Machine Maintenance using RAG and LLM	Advanced	Large Language Model, RAG	Medium-High
Document Transformation using LLM	Intermediate	Large Language Model	Medium
Support Compliance for Clean Room using RAG and LLM	Intermediate	Large Language Model, RAG	Medium-High
Calibration Certificate Service	Advanced	Rule-based, Algorithm	High
Resilience Score Service	Intermediate	Rule-based, Algorithm	High

Table 4 Summary of the AI-Services

The following pages focus on 1) the demonstration of the prototype for the use case and 2) the demonstration preparation, initial testing and feedback for each use case. First, the Human-Machine Interfaces (HMI) developed for the end users are demonstrated referring to how the user interface and prototype behaves, is used and perceived by the testers. Each interface is preceded by a brief description of the current process, highlighting the approach used in the absence of the digital tool. Second, analyses of the results obtained during the demonstration at the two sites are described including qualitative and quantitative feedback from operators, technicians, and managers, evaluations of system reliability as well as the quality of AI-supported decisions, and the perceived impact on work. This section lays the groundwork for the activities planned in WP6: “Evaluation of FAIRWork”, by defining the criteria that will be used for the validation process.

4.1 Assist Decisions about Production Planning (CRF)

4.1.1 Prototype Demonstration

This section provides a detailed description of the CRF-related components, specifically focusing on Production Planning and Resource Allocation.

Introduction

In the automotive manufacturing environment, the stamping department plays a critical role in the value chain, as it represents one of the first stages in the transformation of raw materials into components destined for final assembly. In the absence of advanced digital tools for planning and operational management, the organization of production activities and human resources relies on traditional methods, requiring a high degree of coordination, experience, and adaptability from the personnel involved.

The **production planning** process is managed manually, primarily using Excel spreadsheets, printed documentation, and direct communication between production managers, shift supervisors, and operators. Planning begins with the collection of production requirements, which are based on a combination of factors: customer orders (both confirmed and forecasted), stock levels in the warehouse, press capacity, and tool availability. Based on this input, the production manager defines which part geometries need to be stamped, prioritizing them according to delivery urgency, technical complexity, and the need to optimize tool changeover times. Similar parts are grouped together to minimize the frequency and duration of die changes, which are time- and resource-intensive operations.

The quantities to be produced are calculated by considering the hourly capacity of each press, minimum batch sizes for economic efficiency, and the availability of personnel for each shift. The production plan is typically created on a weekly basis, with daily updates to reflect operational changes. Without an MRP (Material Requirements Planning) system, traceability and visibility of priorities depend heavily on interdepartmental communication and manual documentation. In parallel with production planning, human resources are assigned to the end-of-line stations (**Workforce Allocation**). Each active press requires a team of one to six operators, depending on the size and complexity of the part. These operators are responsible for unloading the stamped components, performing a basic visual quality check, and placing the parts into the appropriate containers. They do not carry out any additional tasks such as packaging, labelling, or dimensional inspection.

The assignment of operators is managed by the shift supervisor or team leader at the beginning of each shift, based on staff availability, the complexity of the scheduled jobs, and the experience of the workers. In some cases, operators may be rotated between stations to maintain flexibility and reduce fatigue caused by repetitive tasks.

One of the most critical aspects of the current process is the management of unpredictable variables, such as last-minute absences, machine breakdowns, tool non-conformities, or urgent changes in customer orders. In these situations, replanning is carried out in real time through quick decisions made by supervisory personnel, often without the support of digital systems. This results in a strong reliance on the experience and responsiveness of the team, as well as a certain rigidity in resource management. For example, in the event of an absence, the shift supervisor may decide to temporarily reduce the number of active presses, reassign available operators across multiple stations, or request support from personnel in other departments. These decisions are typically communicated verbally or via whiteboards in the production area and manually updated in production logs.

In summary, the current operating model is based on manual planning and a workforce management approach that depends heavily on the experience and flexibility of the personnel. While this method allows for a certain degree of adaptability to daily operational conditions, it also presents significant limitations in terms of traceability, efficiency, and responsiveness to unforeseen events. These challenges highlight a clear opportunity for the introduction of digital tools to support and optimize decision-making processes.

Human Machine Interface description:

Building upon the model generated through OLIVE (described at the end of section 2.2), this chapter describes the user interface (HMI) developed to support the Production Planning and Resource Allocation service within the DAI-DSS system, with a specific focus on the CRF use case. The interface serves as the main point of interaction between the end-user and the underlying AI services, offering a clear and interactive visualization of the results produced by the decision engine. This HMI has been designed to facilitate the analysis and management of the production plan over a two-week horizon, integrating data related to orders, production lines, and worker allocation. The goal is to provide an intuitive and flexible tool that enables CRF production managers to quickly evaluate the system-generated proposals, intervene when necessary, and store approved solutions in the Knowledge Base.

The following sections illustrates the main functionalities of the interface, the interaction flow with the system, and the operational steps that allow users to upload data, trigger the AI service, and visualize the results in both graphical and tabular formats:

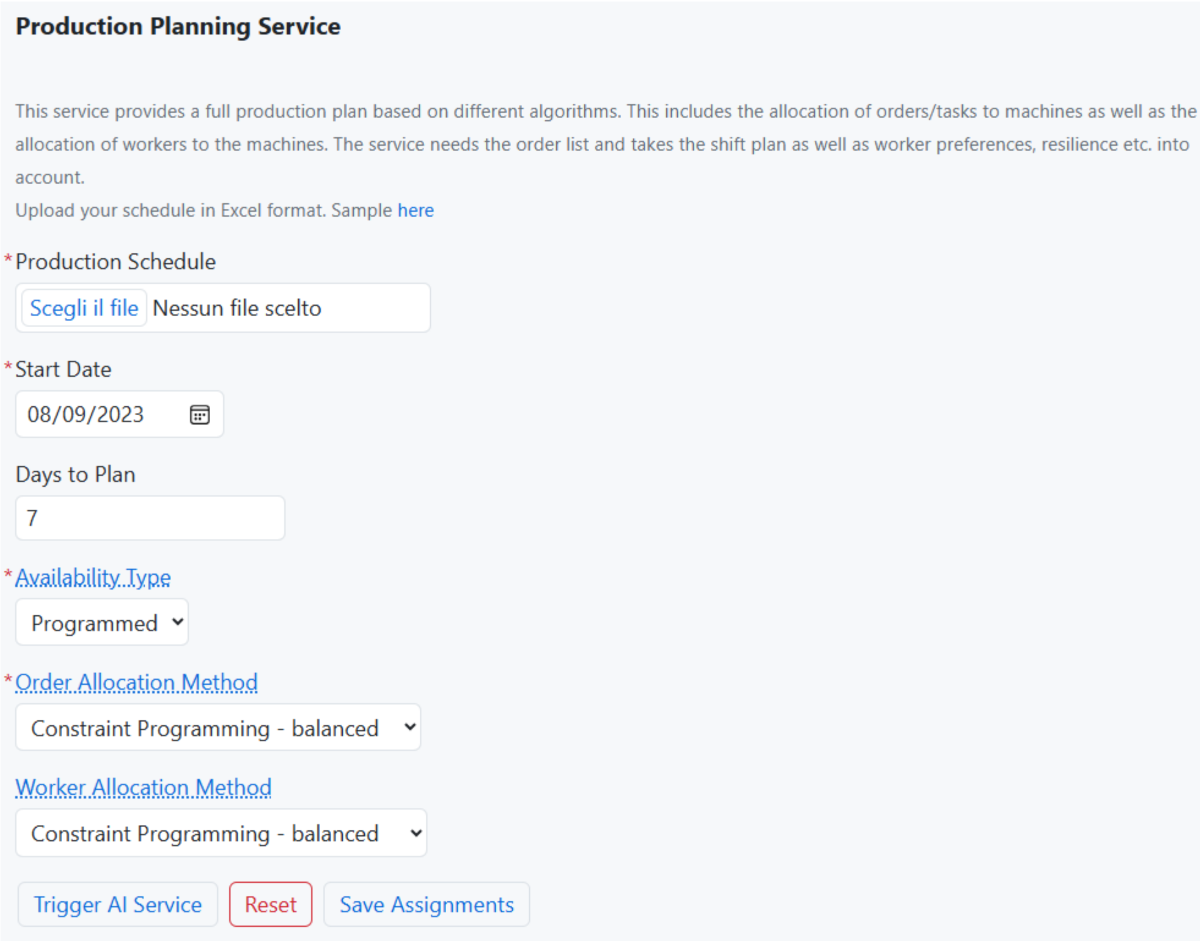

Figure 14 HMI - CRF starting window

The following section provides a detailed description of each individual component of the interface.:

Production Schedule: The first step consists in uploading the Excel file containing the orders. This operation is performed by clicking on “Choose File.” An example of such a file is shown in the Figure 15. Each row corresponds to a geometry that must be produced and delivered. The file includes the following fields:

- ORDER: the Stellantis code of the order issued by a specific customer (an order may include multiple types of components, and therefore multiple geometries).
- DUE DATE: the deadline by which the order must be fulfilled.
- PRIORITY: a TRUE/FALSE field indicating whether priority should be given to that geometry during production scheduling.
- DRAWING (GEOMETRY): the Stellantis code of the geometry (i.e., the drawing number of the component).
- DESCRIPTION: a brief description of the component.
- QTY (ORDERED): the quantity, in units, of that geometry to be produced and delivered.
- MOLDS: the number of molds required to produce the component.

Regarding this file, it should be noted that it is a dynamic document, updated daily based on the remaining quantity of components that still need to be produced. This means that, for example, if an order of 1,000 components of a given geometry is due by the end of the week, and 400 components are produced today, the updated file for the following day will indicate only 600 components remaining to be produced by the end of the week. Therefore, the quantities shown in the file represent the residual amounts.

	A	B	C	D	E	F	G
1	ORDER	DUE DATE	PRIORI	DRAWING (GEOMETR	DESCRIPTION	QTY (ORDERED	MOLD
2	SEV - 36	8-set	FALSE	533908540	Scat. Mont Parab Sx	6290	6
3	SEV - 36	8-set	FALSE	1342266080	Est.Sup.Ant.Dx	3000	6
4	SEV - 36	8-set	FALSE	1343314080	Est.Lat.Sup.Post.(Lungo)	3000	5
5	SEV - 36	8-set	FALSE	531359140	Mont.Post.Cab.Dx	2800	5
6	SEV - 36	8-set	FALSE	1343327080	Est.Sup.Ant.Sx	3000	6
7	SEV - 36	8-set	FALSE	534259180	Oss.Mont.Centr.sx	6500	6
8	SEV - 36	8-set	FALSE	534259080	Oss.Mont.Centr.Dx	6000	6
9	SEV - 36	8-set	FALSE	1342236080	Est.Lat.Inf.	6300	4
10	CAS - 36	8-set	FALSE	505597580	Est.Portell.Parte inf	1800	5
11	SEV - 37	15-set	FALSE	531359140	Mont.Post.Cab.Dx	2800	5
12	SEV - 37	15-set	FALSE	1343314080	Est.Lat.Sup.Post.(Lungo)	3000	5
13	SEV - 37	15-set	FALSE	1340538080	Est.lat.inf.int.Corto	3150	4
14	SEV - 37	15-set	FALSE	533908540	Scat. Mont Parab Sx	6290	6
15	SEV - 37	15-set	FALSE	534259180	Oss.Mont.Centr.sx	6500	6
16	SEV - 37	15-set	FALSE	534259080	Oss.Mont.Centr.Dx	6000	6
17	MIR - 37	15-set	FALSE	521377700	Baule	800	5

Figure 15 Orders file

Start date: it refers to the date from which production scheduling is intended to begin. Typically, this corresponds to the current date; however, a future date may also be specified.

Days to Plan: it indicates the number of days for which production is to be scheduled, starting from the specified start date. The system considers only the orders that fall within the defined time window.

Availability type: in this field, it is necessary to choose whether to use the theoretical attendance plan or the actual attendance data for allocating personnel at the end of the production line.

- **Programmed**: Planned presence of operators follows a regular weekly rotation of shifts: typically, third-second- first; at Mirafiori, first and second shifts alternate, while the third shift remains fixed.
- **Real:** Actual presence often deviates due to absences (holidays, illness, redundancy fund).

Actual attendance refers to the data obtained at the beginning of each shift from the attendance files generated by the badge-scanning system at the line entrance. As a result, there will be a theoretical production schedule, for example covering the entire week, and a more reactive schedule based on the actual presence of personnel, updated shift by shift, as it becomes available.

Order Allocation Method: this section outlines the various strategies available for generating a production schedule based on a given set of orders. All options are based on Constraint Programming (CP) techniques and differ according to the specific optimization objectives they pursue.

Solvers define the algorithmic approach used to generate the production plan:

- Constraint Programming: Generates a production schedule for a given set of orders using a constraint programming strategy. As with other CRF services, high-priority orders are scheduled first.
- Optimize: Optimizes the order allocation with a constraint programming strategy. This approach will only optimize the order allocation without considering the worker allocation to optimize this part of the production plan without the need to redo all the allocations.

Scheduling goals define the criteria used to assign workers to production lines:

- Balanced: The optimization weights tardiness and makespan equally (1:1), meaning that avoiding one hour of tardiness is treated as equally valuable as reducing makespan by one hour. This balanced approach results in good machine utilization while still avoiding many late deliveries. It may occasionally delay early starting orders in favour of completing others on time to meet due dates more effectively.
- Tardiness: The scheduling goal is to minimize tardiness defined as the time by which an order misses its deadline. Makespan is not considered in this optimization. As a result, the system maximizes on-time deliveries, even if it reduces machine utilization by prioritizing deadline adherence over throughput.
- Makespan: Orders are assigned to production lines in a way that minimizes the makespan, that is the total time required to complete all orders. This service does not consider delivery deadlines. Its sole focus is on minimizing makespan, resulting in high machine utilization, even if it means missing some deadlines in favour of throughput efficiency.

Worker Allocation Method: This section describes the strategies available for assigning personnel to production lines, based on different criteria and optimization goals. The available options are structured into two levels: solvers (the computational engines used) and allocation methods (the decision-making logic applied).

Solvers define the algorithmic approach used to solve the allocation problem:

- Constraint Programming: Uses a Constraint Programming solver (CP) to allocate workers to the production lines.
- Reinforcement Learning: Uses Reinforcement Learning (RL) to allocate workers to the production lines. The RL-based allocation approach requires moderate computational resources to generate solution instances, but it is still more efficient than some other simulation-based methods.
- Monte Carlo Tree Search: Uses Monte Carlo Tree Search (MCTS) to allocate workers to the production lines. MCTS approaches are computationally heavy and result in the longest solution times among all allocation methods.

Allocation methods define the criteria used to assign workers to production lines:

- Balanced: Balances three objectives (experience, preference, and resilience) each weighted equally (1:1:1). This balanced approach ensures that skilled workers are placed in roles where they are effective, personal preferences are respected.
- Experience: Prioritizes assigning workers to tasks and lines where their experience is most relevant. This improves productivity and quality by ensuring skilled personnel are matched with the most suitable tasks.

- Preference: Prioritize the worker maximum satisfaction by considering preferred tasks or production lines.
- Resilience: Prioritize the worker resilience.

Trigger AI Service: by pressing this button, the generation of the Production Planning is initiated. The results of this operation will be presented later in this document.

Reset: by pressing this button, all previously entered preferences are reset, and the web page returns to its initial state

Save Assignments: by pressing this button, an Excel file is saved containing the planned production along with the previously selected preferences. The file includes the following information:

- Task: the order identification and the geometry
- Start / Finish: start and end times of the task
- Resource: the assigned production line
- Geometry: the code of the geometry to be produced
- is_setup_timebox: indicates whether setup time is required
- produced_amount: quantities produced
- produced_until_now: produced so far
- order_total_amount: quantities total required
- Required_workers: number of assigned workers
- Workers: names of assigned workers
- Warning: any operational alerts or notes

The continuation of the user interface and content description includes, by way of example, a 7-day scheduling scenario, with the Order Allocation Method set to 'Constraint Programming - Balanced' and the Worker Allocation Method also set to 'Constraint Programming - Balanced'. Upon pressing the 'Trigger AI Service' button, the window shown in the figure appears, indicating that the tool is running:

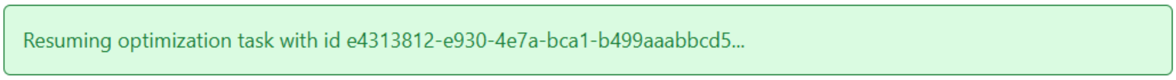

Figure 16 HMI - CRF running window

After a few seconds (the duration may vary depending on the selected period and methods, the proposed setting results in an execution time of approximately 20 seconds), the result appears in the lower section of the page:

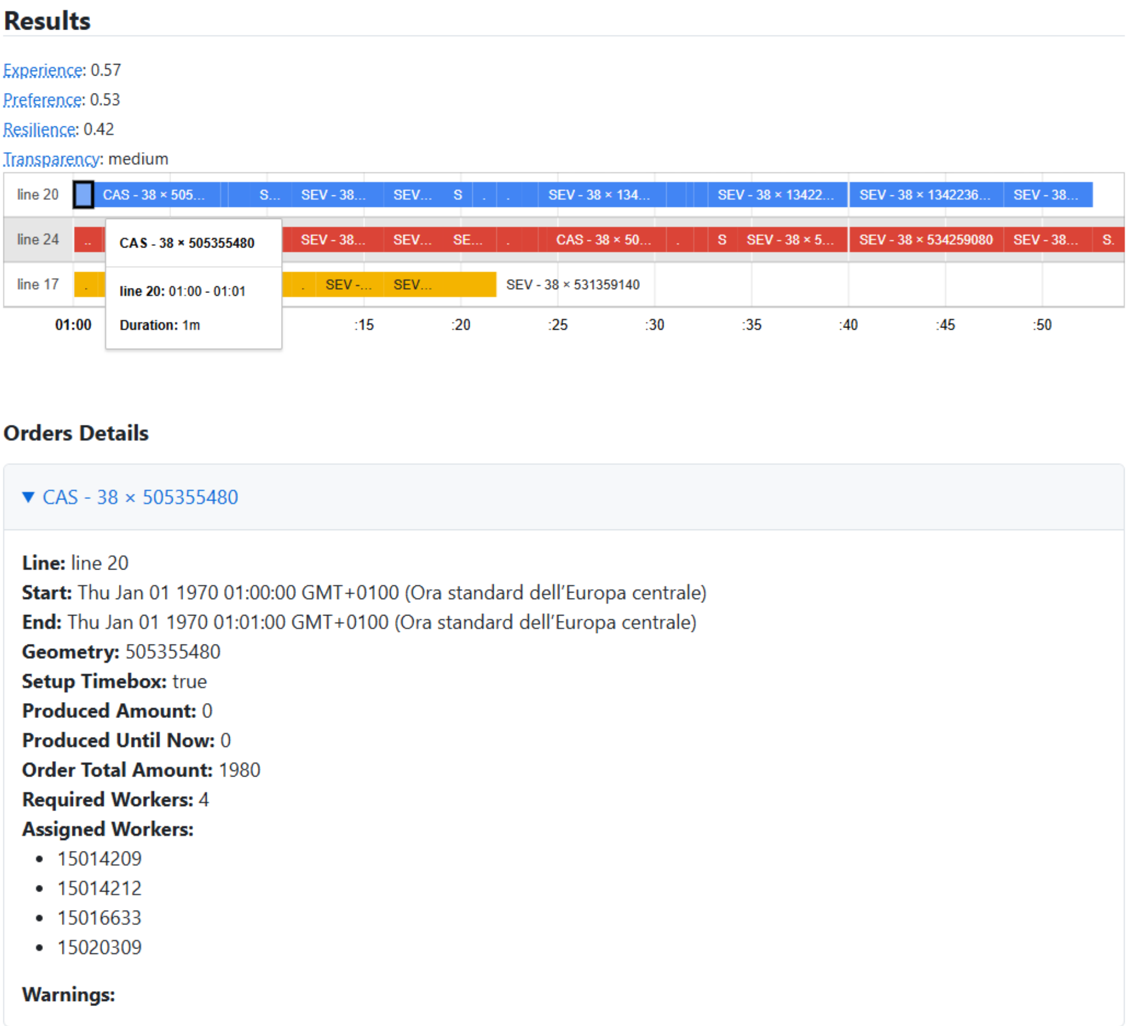

Figure 17 HMI - CRF result window

At the top of the screen, the indicators resulting from the operator allocation are displayed:

Experience: the KPI measures the overall experience levels present in the solution allocation. It is calculated as a weighted sum of all experience values, where the weights correspond to the duration of each task. The experience values are scaled such that:

1) The Experience KPI equals 1.0 when all workers have an experience value of 1 for all their tasks.
2) The Experience KPI equals 0.0 when all workers have an experience value of 0 for all their tasks.
3) If all workers have an experience value of 0.5 for all their tasks, the Experience KPI would be 0.5.

Preference: the KPI measures the overall resilience levels present in the solution allocation. It is determined as a weighted sum of all resilience values, where the weights are based on the duration of each task. The resilience values are scaled such that:

1) The Resilience KPI equals 1.0 when all workers have a resilience value of 1 for all their tasks.
2) The Resilience KPI equals 0.0 when all workers have a resilience value of 0 for all their tasks.
3) If all workers have a resilience value of 0.5 for all their tasks, the Resilience KPI would be 0.5.

Resilience: the KPI measures the overall resilience levels present in the solution allocation. It is determined as a weighted sum of all resilience values, where the weights are based on the duration of each task. The resilience values are scaled such that:

1) The Resilience KPI equals 1.0 when all workers have a resilience value of 1 for all their tasks.
2) The Resilience KPI equals 0.0 when all workers have a resilience value of 0 for all their tasks.
3) If all workers have a resilience value of 0.5 for all their tasks, the Resilience KPI would be 0.5.

Transparency: the KPI is a value assigned to a service and serves as a measure of its transparency. Each service is analysed in terms of both global and local transparency. Global transparency refers to the system. In the case of services, the specific use case is crucial for global transparency. Local transparency describes how the service operates and how its results are derived.

A service based on analytical expressions and first-principles methods is considered more transparent than one relying on statistical models. Additionally, external factors such as information about the authors and reviewers of the services also play a role. All these aspects are combined into the overall transparency rating of the service.

Below the indicators, the actual scheduling across the various lines over time is displayed. In the example provided, production is scheduled across three lines (17, 20, and 24). The bars represent the utilization of each line over time and are segmented into blocks, with each block corresponding to a specific task.

When hovering over the blocks, a popup window appears displaying key information (Task, Line, Start/Finish, Duration), while clicking on a block opens a detailed view window. There are as many detail windows as there are blocks. The information contained in each detail window corresponds to that previously listed in the "Save Assignments" section, as each block generates a row in the Excel file.

Among the warnings, the message 'Too few workers assigned' may appear. In such cases, the task field is preceded by a yellow triangle containing an exclamation mark. This indicates that it is not possible to assign the exact number of operators required for unloading the line. It is then the responsibility of the line manager to address this issue, either by requesting temporary support from other departments or by rescheduling the task.

4.1.2 Demonstration Set-up, Testing and Feedback

The initial section of this chapter presents the preliminary testing activities carried out on the demonstrator by CRF, in collaboration with the Mirafiori press-shop personnel. These tests were conducted during the development phase of the demonstrator, covering the period from the initial prototype to the final version, prior to validation.

The objectives of these activities were as follows:

- to introduce the demonstrator and the associated AI-related topics.
- to collect key insights for steering the development process (e.g., required modifications, system integrations).
- to gather early feedback from plant personnel, including doubts, concerns, and potential resistance.

4.1.2.1 Preliminary test and feedback

During the phase of test, the tool AI was introduced to a selected group of managers and line operators. The objective was to observe the interaction with the system, gather spontaneous and structured impressions, and identify any barriers to adoption. A summary of the feedback collected is presented below:

1. Understanding and Trust in the System

- *"It is not clear how the system arrives at certain decisions."* Many users expressed difficulty in understanding the logic behind the AI suggestions, requesting more transparency (e.g., explanations or motivations for choices).
- *"I trust my experience more than an algorithm."* Some managers showed skepticism, preferring to rely on their intuition developed over the years.

2. Interface and Usability

- *"The interface is simple, but some functions are not intuitive."* Usability was generally appreciated, but requests for improvement emerged to make operational steps clearer.
- *"I would like to manually modify the suggestions."* More flexibility was requested to intervene on AI proposals, maintaining decision-making control.

3. Perception of Added Value

- *"It is useful to have an overview, especially in complex shifts."* Some users recognized the value of the system in managing complex situations, such as assigning workers based on preferences and skills.
- *"If it helps me save time, I will use it gladly."* Adoption seems linked to the perception of a concrete benefit in terms of operational efficiency.

4. Cultural and Organizational Resistance

- *"Not everyone is ready to be told what to do by a machine."* Cultural resistance emerged, related to the perception of AI as a "controller" rather than as support.
- *"Training is needed to understand how to use it really well."* Users emphasized the need for training sessions to fully understand the potential of the system.

About the conclusions, the tests revealed a mix of curiosity, skepticism, and openness. To encourage the adoption of the system, it will be essential to:

- improve the transparency of AI decisions (e.g., contextual explanations),
- offer targeted training,
- ensure flexible and non-binding interaction,
- actively involve users in the development process.

A summary of the suggestion for improvement collected is presented below:

1. Greater Transparency and Explainability

- *"I would like to understand why the system proposes certain assignments."* Request to introduce functionalities for explaining decisions (e.g., "why was this worker assigned to this line?").

2. Customization of Parameters

- *"Every day is different; it would be necessary to adjust priorities."* Suggestion to allow users to modify optimization criteria, for example, giving more weight to experience or worker preference based on the context.

3. Integration with Existing Systems

- *"It would be useful if the tool were connected to the management system we already use."* Request for integration with existing MES systems to avoid data duplication and simplify the operational flow.

4. Greater Operational Flexibility

- *"Sometimes I have to change assignments at the last minute."* Need to manually modify AI proposals easily and quickly, even in unforeseen situations.

5. Continuous Training and Support

- *"A tutorial or step-by-step guide would be useful."* Proposal to introduce interactive training material, such as videos, simulations, or contextual help.

6. More Intuitive Interface

- *"Some buttons are not clear; it took me a while to understand what they do."* Suggestion to simplify the user interface, with clearer labels, intuitive icons, and guided paths.

About the conclusion, these suggestions highlight the value of a user-centred approach in tool development. Integrating these suggestions can increase system acceptance, improve user experience, and facilitate large-scale adoption.

4.1.2.2 HMI preparation to validation

This document is the continuation of the draft prepared and inserted in D5.2 (DAI-DSS Infrastructure and Setup Report at Use Case Site) from page 22 to page 24.

As anticipated in the previous document, the usability test is divided in 3 phases, discussed below:

- Test preparation
- Actual test
- Results analysis

To test the HMI, the number of test users has been chosen to be 10. In accordance with literature a usability test with 9 users allows to capture 90% of the critical elements of interaction with a product.

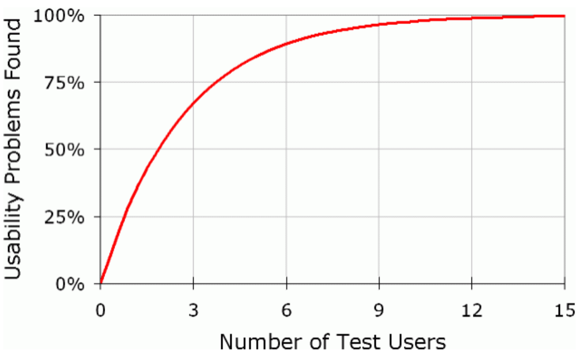

Figure 18 Usability problems found on number of test users

The users will be volunteer colleagues of the CRF, as the skills of the working group are very similar to those present in the line staff with decision-making tasks:

1. *User1* (Flexible & adaptive systems specialist)
2. *User2* (Flexible & adaptive systems specialist)
3. *User3* (Flexible & adaptive systems specialist)
4. *User4* (Production technologies specialist)
5. *User5* (Production technologies specialist)

6. *User6* (Process integration specialist)
7. *User7* (Project manager)
8. *User8* (Factory sustainability manager)
9. *User9* (Manufacturing methods specialist)
10. *User10* Manufacturing methods specialist)

Regarding the evaluation team, normally it's made by cognitive ergonomists, psychologists, and software engineers. In our case it will be made by:

1. *Evaluator1* (Project responsible)
2. *Evaluator1* (Eco-factory senior specialist)
3. *Evaluator1* (external, Ergonomics senior specialist)

The next step is to define a series of task that the user must do (protocol definition), this is done by a team composed by the end-user (CRF) and the partners of WP4 and WP5. Then the team writes a questionnaire about the interface use. A first draft is reported below:

FAIRWork 101069499
Flexibilization of complex Ecosystems using Democratic AI based Decision Support and Recommendation Systems at Work

HMI evaluation questionnaire

Production Planning and Resource Allocation Service

Quantitative test: to be filled by the expert observer

Instruction: for each of the following tasks, observe if the volunteer completes it with **S**uccess, **P**artial success, **F**ailure:

Task 1. *Open the application*	**S**	**P**	**F**
Task 2. *TBD*	**S**	**P**	**F**
Task 3. *TBD*	**S**	**P**	**F**
Task 4. *TBD*	**S**	**P**	**F**
Task 5. *TBD*	**S**	**P**	**F**
Task 6. *TBD*	**S**	**P**	**F**
Task 7. *TBD*	**S**	**P**	**F**
Task 8. *TBD*	**S**	**P**	**F**

Qualitative test (SUS): to be submitted to the volunteer

Instruction: for each of the following statements, make one box that best describes your reaction to the software:

	Strongly disagree				Strongly agree
1. *I think that I would like to use this software frequently*	1	2	3	4	5
2. *I found this software unnecessarily complex*	1	2	3	4	5
3. *I thought this software was easy to use*	1	2	3	4	5
4. *I think that I would need assistance to be able to use this software*	1	2	3	4	5
5. *I found the various function in this software were well integrated*	1	2	3	4	5
6. *I thought there was to much inconsistency in this software*	1	2	3	4	5
7. *I would imagine that most people would learn to use this software very quickly*	1	2	3	4	5
8. *I found this software very cumbersome/awkward to use*	1	2	3	4	5
9. *I felt very confident using this software*	1	2	3	4	5
10. *I needed to learn a lot of things before I could get going with this software*	1	2	3	4	5

Figure 19 HMI evaluation questionnaire draft

The test will be done by every user, separating, applying this protocol:

1. General explanation of the project, of the demonstrator, of the actions to be carried out in the test, of the purpose of the test.
2. Description and teaching of the tasks to be performed, with a step-by-step practical demonstration (volunteers were allowed to take notes).
3. First execution of the tasks by the volunteer, with the assistance of expert staff.
4. Second execution of the tasks by the volunteer, individually and without assistance (in this phase the expert staff has filled in the first part of the questionnaire).
5. Completion of the second part of the questionnaire by the volunteer.

The questionnaires will be anonymous, and no photographic material will be produced during the execution of the tests.

The first section of the test regards the "quantitative test", it is used to objectively define how many of the volunteers were able to carry out the assigned tasks correctly. The result is expressed as a percentage (0% - 100%), for each volunteer. The scores of each task are added together and divided by the total number of tasks. Each evaluation has the following score:

- S=1
- P=0.5
- F=0

The target value for this KPI normally is **80%**, which corresponds to a very good rating of the software.

The second section of the questionnaire is related to the "qualitative test". The System Usability Scale (SUS) is a quantitative method that collects numerical data on users' subjective experiences. Although the responses reflect personal perceptions, the standardized structure and scoring procedure classify it as a validated quantitative tool for usability evaluation. It consists of a 10-item questionnaire with five response options for volunteer. It is about the feelings of the volunteer using the interface. Also, in this case, the result is expressed as a percentage (0% - 100%), for each volunteer. The scores of each task are added, then the result is multiplied by 2.5. Each item score has a range from 0 to 4 and it is calculated as follows:

- For odd items (1, 3, 5, 7, 8): score = valuation – 1
- For even items (2, 4, 6, 8, 10): score = 5 - valuation

A value equal to **85%** is typically corresponds to an "excellent" rating in the SUS scale and a value greater than **73%**, corresponding to a "good" rating, in the SUS scale, is generally used to indicate a satisfactory usability.

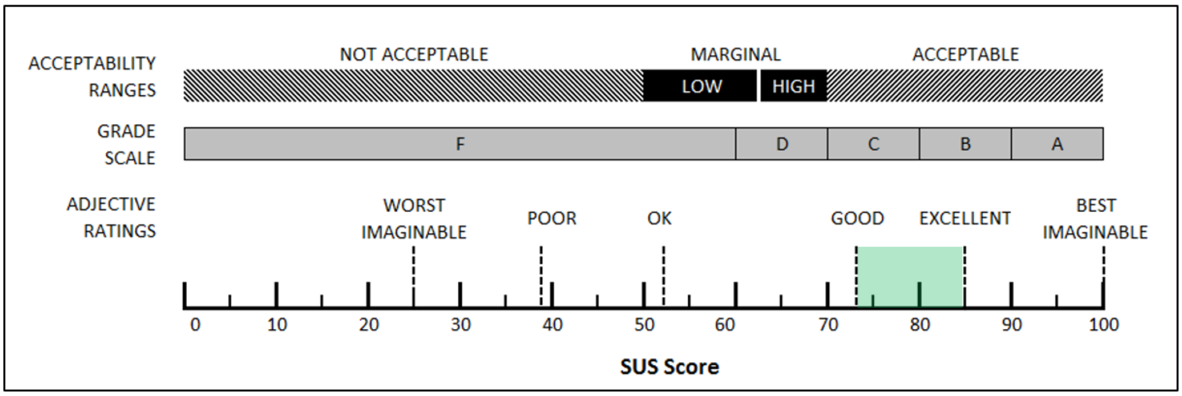

Figure 20 SUS test - evaluation scale

4.1.2.3 Decision-making model preparation to validation

This section describes the evaluation procedure of the decision model. In this case, the aim is to evaluate the prototype's ability to propose correct decisions.

This activity requires a high level of experience in the production field, so the plant personnel will be involved. The person in charge to this activity is:

1. *Responsible1* (Circularity CoC senior specialist)

This is due to his professional background and direct knowledge of the Press shop; he can act as an intermediary towards the Plant management. For information purposes, the plant organization chart is shown below for the Production Planning & Control:

Figure 21 Press shop organization chart

The main job functions, who will be involved in the analysis under the manager's designation, are:

- Handling Operations
- Lines and Materials Planning
- Manpower Analysis

As in the previous case, the activities will be divided into 3 phases, described in the following paragraphs:

- Test preparation
- Actual test
- Results analysis

The first step to be done is informative and formative. We will have two separate sessions with the plant personnel:

Informative session: in this phase the FAIRWork project will be presented, describing its actions and objectives. It will be explained what is meant by "democratic decision-making model" and how it was thought to apply it to the real industrial case. The demonstrators will be illustrated, which inputs they start from and what type of output they are able to generate. Attention will be paid to the aspects related to the fairness of the proposed solution and what are the obtainable advantages.

Formative session: in this phase, the staff will be trained to use the demonstrator. The user interface will be presented in all its aspects, describing each option. Particular attention will be paid to describing the necessary inputs, i.e., the necessary information, their formats, how they are used by the system.

Regarding the last phase, it is emphasized that it will be a highly practical session. The staff will be encouraged to test the prototype, to vary the parameters and inputs to better understand the effect on the results. It is believed that one day of preparation will be sufficient to complete the training.

The image below refers to the HMI created for the Production Planning and Resource Allocation Service, that will be one of the topics of the training:

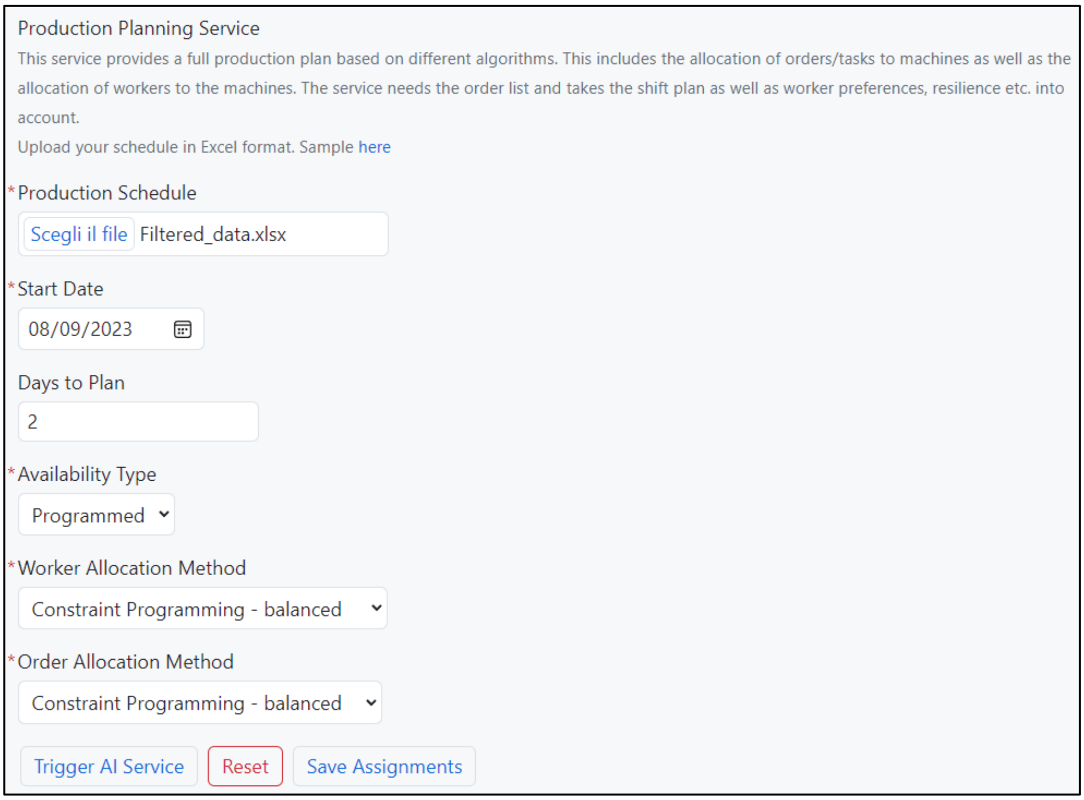

Figure 22 HMI for Production Planning Service

At this stage, the questionnaires that will be used to collect feedback during the test, will also be prepared. The figure below shows an example for Production Planning and Resource Allocation Service that will be explained in the next paragraph:

FAIRWork 101069499
Flexibilization of complex Ecosystems using Democratic AI based Decision Support and Recommendation Systems at Work

FAIRWORK

DMM evaluation questionnaire

Production Planning and Resource Allocation Service

Instruction: Choose a week for the test starting from Monday and set the "Days to plan" equal to 5 days. Start with the scheduled availability of the personnel. Perform subsequent tests on the reference week by changing some parameters (actual availability of the staff starting from a given day/shift, new priority of an order, change of order quantities, ...), please report the method used for allocation. Fill in from time to time, giving a rating from **1** to **10** to the result, insert comments if desired. The final section is edited by CRF personal for overall response.

Test 1	**Week:** 18-24 September		
1: Programmed personnel availability *Method:* Worker: 1.2 + Order: 1 *Comment:* I would have preferred to start with the most expensive order		*Planning* 8	*Allocation* 9
2: Real personnel availability at day 2 *Method:* *Comment:*		*Planning*	*Allocation*
3: *Method:* *Comment:*		*Planning*	*Allocation*
4: *Method:* *Comment:*		*Planning*	*Allocation*
5: *Method:* *Comment:*		*Planning*	*Allocation*
6: *Method:* *Comment:*		*Planning*	*Allocation*
Global response:		*Planning*	*Allocation*

Figure 23 DMM evaluation questionnaire draft

Once the test preparation phase is completed, the actual test phase will be conducted. As previously mentioned, the Production Planning & Control Manager will identify the right people, a specialist for each function (Planning, Allocation).

Candidates will not carry out the assessment alone but will be supported by CRF personnel experienced in using prototypes. This is because we want to avoid that any difficulty in using the HMI affects the results (in this phase we want to test the results produced by the prototypes, not the HMI).

As reported in the "Production Planning and Resource Allocation Service" evaluation questionnaire, for each test session, the candidate will:

- Choose a week for the test, starting from Monday
- Set the “Days to plan” equal to 5 days.
- Start with the scheduled availability of the personnel.
- Report the method used for allocation.
- Fill in from time to time, giving a rating from 1 to 10 to the result
- Insert comments if desired.
- Perform subsequent tests on the reference week by changing some parameters (actual availability of the staff starting from a given day/shift, new priority of an order, change of order quantities, ...)

At the end of the test session, the final section is edited by CRF personal for overall response, calculating the average of the scores and recording a final comment.

This procedure will be repeated at least 10 times, considering different weeks. So, at the end of the tests, we will have as many questionnaires as there are test sessions. It is believed that one day will be sufficient to complete the tests.

For each Service all the sheets will be collected and the average of all the “global response” will be calculated. So, at the end we will have a final score about the prototype's ability to propose correct decisions about:

- **PLANNING [1...10]**: ability to optimize production by considering multiple factors
- **ALLOCATION [1…10]**: ability to allocate people appropriately on the lines

Up to now nothing has been said about the target of these indicators, a result equal to or higher than **7** will be considered good.

In addition to the evaluation described so far, the KPIs defined in D4.1.1_DAI-DSS Architecture will also be taken into consideration. For all the scenarios analysed we have 2 KPIs easy and immediate to evaluate:

DECISION TIME [sec]: Time will be measured from the instant in which all the data necessary to allocate workers on the lines will be available. The input data will therefore be the presence of people in the plant for the work shift, the database of the operators' features, the desired production (components and quantities). This KPI applies not only for the initial production planning, at the beginning of the shift, but also in the case of replanning due to unforeseen events during production. The target value is: **UNDER 1 MINUTE**

ALTERNATIVE PLANS [n]: the model must be able to provide multiple alternative solutions, clearly assessable by the line manager. This does not mean that a solution is always possible, but if the personnel present in the plant is not suitable to meet the planned production, this must be clear. In the best-case scenario, where scheduled production is possible, the model must provide **5 ALTERNATIVES**.

In addition, we have 2 medium-long term KPIs. For these 2 KPIs the evaluation will not be immediate, but will go beyond the end of the Project, evaluating the improvements obtained in the end-of-period final data:

UNPLANNED ABSENTEEISM [%]: the way to measure worker well-being is linked to the level of unplanned absenteeism. The formula to calculate it is reported into the Deliverable 2.1 and is defined as the total person-hours (paid, unpaid) where the hourly worker was not at work as expected. The acceptable target in this case is a **NO WORSENING**, in the best case a **2.5% REDUCTION**. This KPI is related to the “Workload Balance” scenario.

ENERGY CONSUMPTION [kwh/Ton]: the formula used to measure the energy consumption is reported is into the Deliverable 2.1. It is a measure of the total energy consumed by a facility to produce components. The values that will be taken into consideration will be the ones of the press lines in relation to the quantities produced. The desirable target in this case is a **5% REDUCTION**. This KPI is connected to “Production Planning” scenarios.

4.2 Assist Decisions for Truck Loading (CRF)

4.2.1 Prototype Demonstration

This service has a 'basic' level of maturity, therefore no full demonstration and validation has been performed.

Introduction:

In the automotive manufacturing environment, the truck loading phase represents a critical step in the outbound logistics process, ensuring that finished components are shipped to customers in a timely and efficient manner. Despite its strategic importance, this activity is currently managed manually, relying heavily on the experience and availability of a few key personnel. The truck loading process is coordinated by a dedicated operator who is responsible for organizing the loading of containers based on shipment schedules, part dimensions, packaging constraints, and customer delivery requirements. This task involves complex decision-making, including the selection and arrangement of parts to maximize space utilization, minimize the number of trips, and ensure compliance with loading priorities.

The absence of digital tools means that all planning is carried out using printed documentation, spreadsheets, and direct communication with warehouse staff and logistics coordinators. The operator must manually assess the available space in each truck, match it with the outgoing parts, and determine the optimal loading sequence. This process is time-consuming and highly dependent on the operator's knowledge of part geometries, packaging configurations, and shipping constraints.A major operational risk arises when the designated operator is unavailable due to absence or reassignment. In such cases, the lack of standardized procedures or digital support tools can lead to delays, suboptimal loading, or even missed shipments. The risk of not loading all containers correctly or on time becomes significant, potentially disrupting the supply chain and affecting customer satisfaction.

Over the past few years, plant management has explored various software solutions to support or automate the truck loading process. However, none of the tools evaluated were found to be sufficiently flexible or tailored to the specific needs of the press shop's logistics operations. As a result, the process remains manual and vulnerable to inefficiencies.

In summary, the current truck loading model is based on manual planning and expert knowledge, which allows for a certain degree of adaptability but also introduces critical limitations in terms of scalability, repeatability, and resilience. These challenges underscore the need for a digital decision support system capable of optimizing truck loading operations and reducing dependency on individual expertise.

HMI description:

This section describes the HMI to support decision in stacking transport containers inside trucks. The user interface allows for the upload of an excel file with input data regarding the shipment order, containers specifications (size, stackability and weight) and trucks specifications (size and weight limit). After uploading it with the click of a button, the container arrangement recommendation is processed in the cloud.

Figure 24 HMI - CRF truck loading file upload view

This approach follows a heuristic strategy of First-Fit Decreasing (FFD) in which the largest containers are stored first, taking the maximum of available space in the given truck. It recommends the allocation based on the size, weight and stackability (maximum number of supporting stacking containers on top of each other) of the containers.

As output, it gives a visualization in 3D of the number of necessary trucks to ship the order, how each container should be positioned in the trucks and the weight saturation of each truck. This prototype aims to offer an improved recommendation of packing containers shipment in delivering trucks compared to the current strategy based on implicit human knowledge.

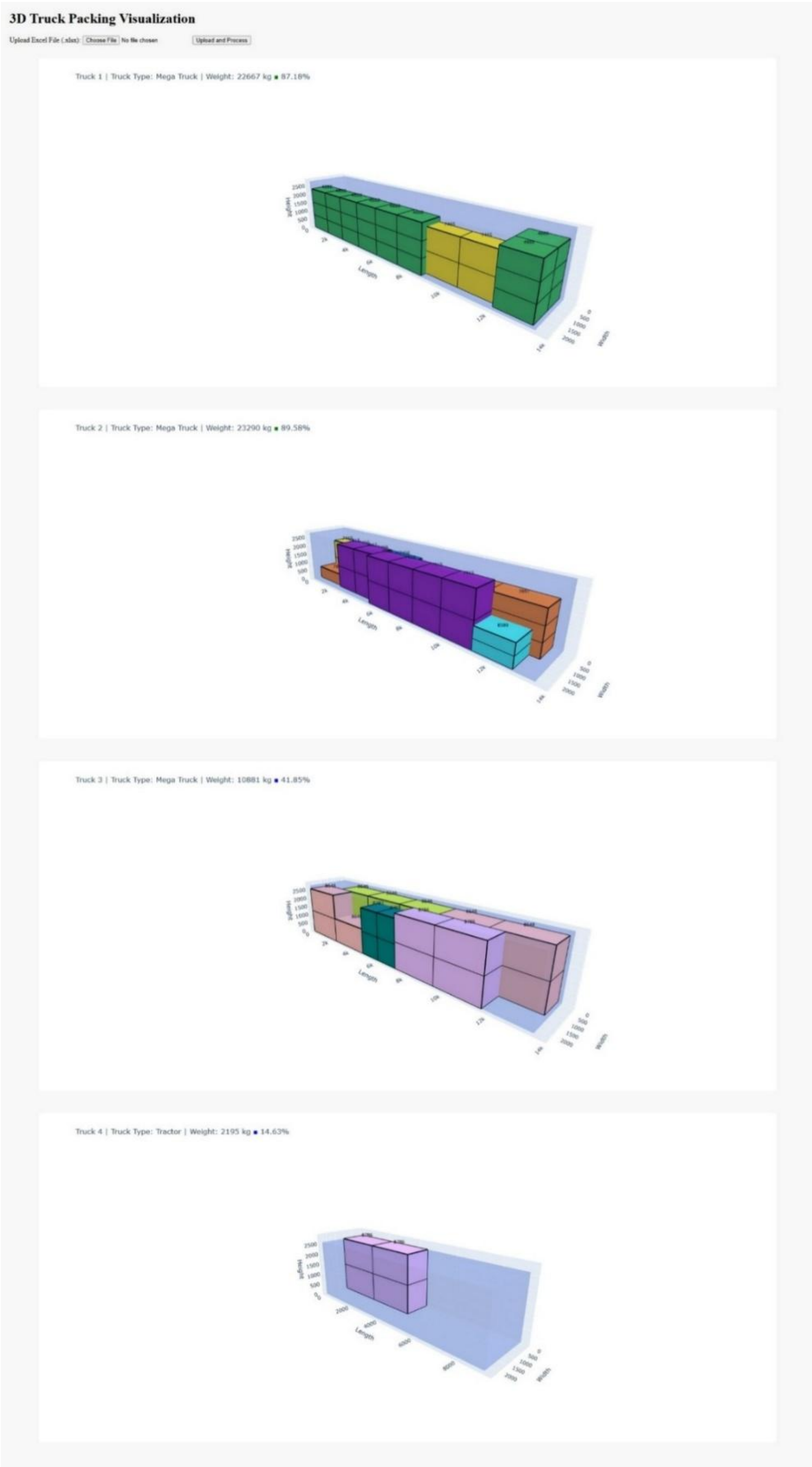

Figure 25 HMI - CRF visualization of container placement in delivering trucks

The input file used for the demonstration, named truck_loading.xlsx, is structured into three sheets:

- Orders: lists the item codes and the quantities to be shipped.
- Containers: specifies, for each code, the physical dimensions, weight, and stackability.
- Trucks: defines the dimensional characteristics and maximum load capacity of the available vehicles.

This structure enables the system to generate a loading proposal that is consistent with real-world logistical constraints.

Despite its simplicity, this solution represents a first step toward the digitalization of a logistics process traditionally managed manually. The automatic data processing and the graphical and tabular visualization of results facilitate the understanding of the proposed solutions and support their validation by operators.

4.2.2 Demonstration Set-up, Testing and Feedback

The Truck Loading Service was not included among the validated demonstrators mainly due to its "basic" maturity level throughout the project. However, the prototype was partially tested and analysed by CRF and Mirafiori plant personnel, who acknowledged its potential. The final version of the service was completed only towards the end of the project, which limited the possibility of structured validation. The decision to not proceed with full testing was driven by the need to focus resources on more mature services, such as Production Planning and Workload Balance, which reached an advanced level and were thoroughly validated with press shop personnel.

4.3 Improve Information Access to Support Maintenance (FLEX Althofen)

4.3.1 Prototype Demonstration

This section provides a detailed description of the FLEX-related components, specifically focusing on the Support Machine Maintenance Service implemented at the Althofen site.

Introduction:

Since Flex operates as an electronic manufacturing service company, we have many different machines and processes that are becoming more and more complex and difficult. Due to the highly automated machines, the problems of increasing maintenance effort and the reliability and susceptibility of the machines are also increasing. Due to this fact, the effort required by the maintenance team to keep production running is increasing.

Therefore, it would be desirable and necessary to give the machine operator more competence to be able to handle minor machine problems themselves and not be tempted to call in the maintenance team.

A further challenge is the lack of experience of newer colleagues who need help with even the smallest problems and thus reduce the availability of the maintenance team.

The idea is to improve this situation with an AI tool and to provide employees with a knowledge database with intelligent communication.

A prerequisite is to create a support tool that requires no or minimal training effort.

HMI description:

Building upon the model generated through OLIVE (described at the end of section 2.2), this chapter describes the user interface (HMI) developed to support the Machine Maintenance Service within the DAI-DSS system, with a specific focus on the FLEX use case in the Althofen site.

As described at the end of section 2.2 the frontend is:

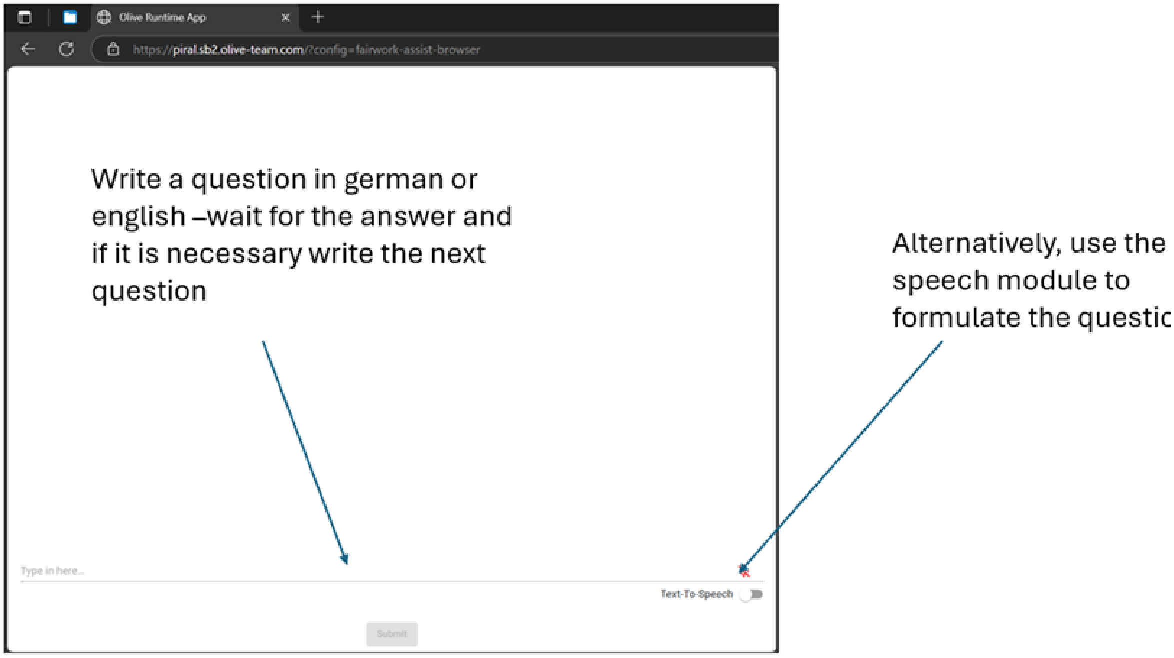

Figure 26 HMI - FLEX Althofen

After opening the communication page, the operator can start communicating by writing or speaking. The requirement of a simple to use HMI was solved excellently by the partners - to be used like Chat GPT.

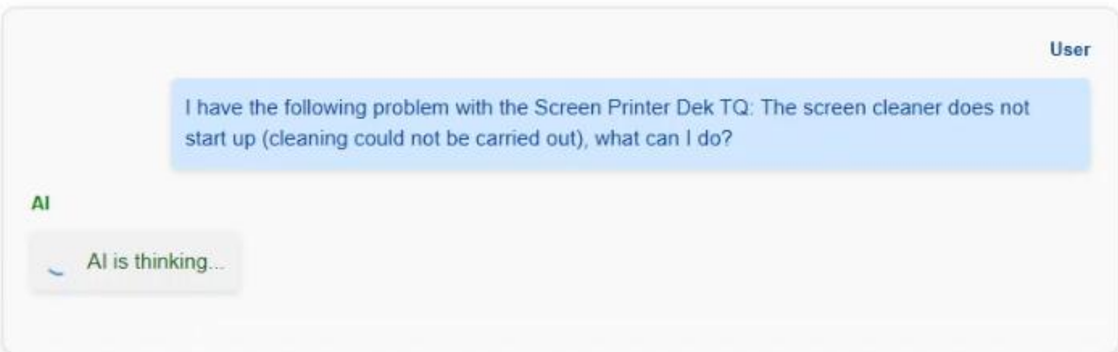

Figure 27 Chatbot usage demonstration (part 1)

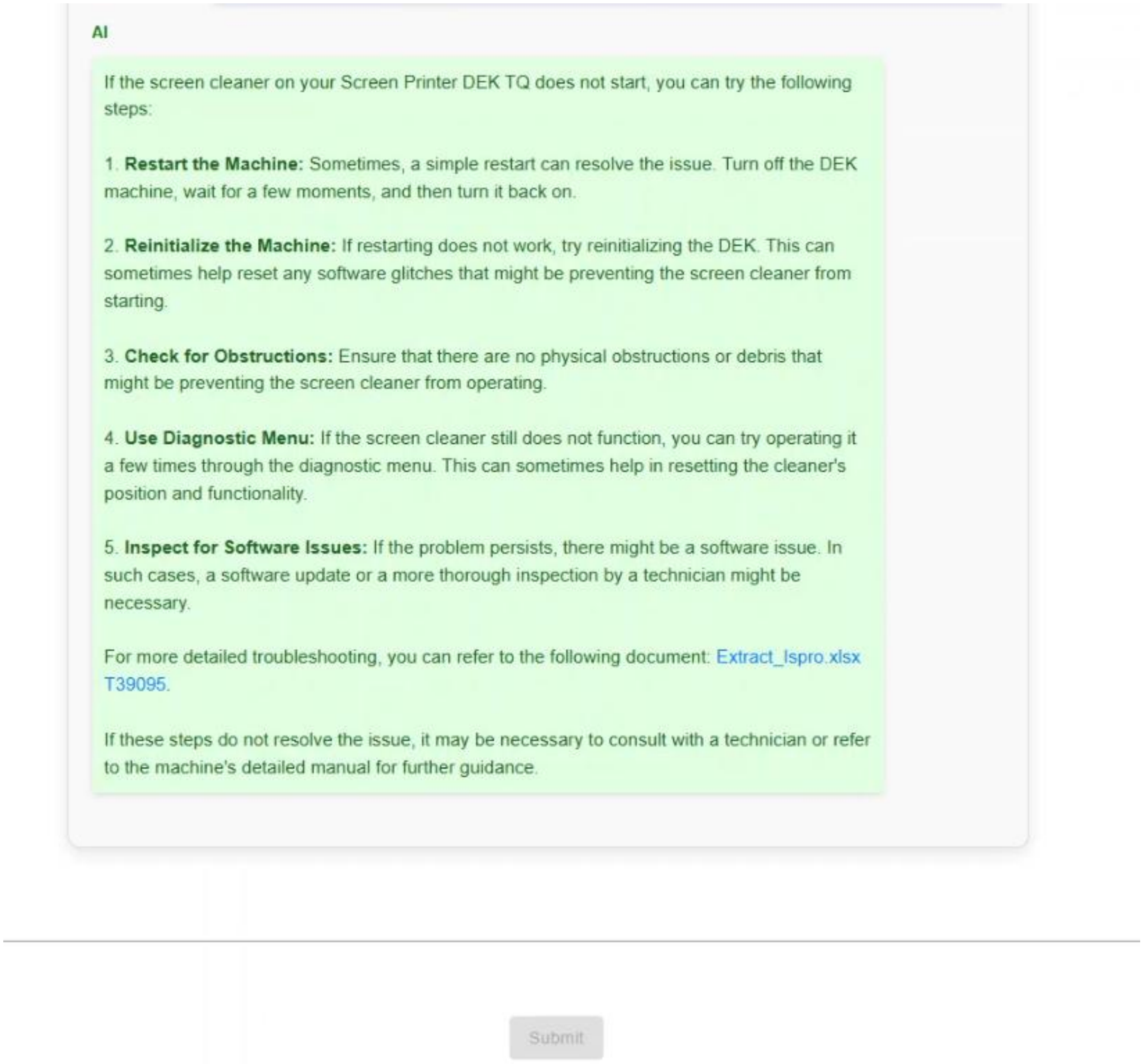

Figure 28 Chatbot usage demonstration (part 2)

4.3.2 Demonstration Set-up, Testing and Feedback

As part of the FAIRWork project, FLEX contributed two distinct case studies developed at its production sites in Althofen and Timisoara. Each site addressed specific operational challenges by integrating AI-based decision support tools into their workflows.

4.3.2.1 Preliminary test and feedback

The Althofen site focused on enhancing machine maintenance processes through AI, aiming to reduce downtime and empower less experienced operators to handle minor issues independently. During the initial testing phase, several key insights emerged:

1. AI systems require large volumes of clean, well-labelled data to perform effectively. Rare or complex faults are often underrepresented, making model training more difficult and potentially requiring additional resources.
2. Integrating AI with industrial systems such as PLCs, SCADA, MES, or ERP (e.g., Flex Flow) presents technical challenges and demands careful planning and robust infrastructure.

3. The Human-Machine Interface (HMI) was found to be simple and user-friendly. However, the effectiveness of the system depends heavily on the specificity and clarity of the questions posed by the user.
4. To improve usability, a language module was integrated upon user request, enabling natural language interaction, and making the tool more accessible, like a conversational assistant.

4.3.2.2 HMI preparation to validation

At the Althofen site, the AI tool was designed with a strong emphasis on simplicity and ease of use. Unlike the CRF case, which involved a complex interface for production planning and resource allocation requiring structured usability testing, the solution implemented at Althofen relies on a chat-like interface that mimics the interaction style of popular conversational agents.

This intuitive design allows operators to communicate with the system either by typing or speaking, making the tool immediately accessible even to users with limited technical experience. The interface was developed to require minimal training effort, with the goal of supporting maintenance tasks and troubleshooting in a fast and user-friendly manner.

Given the elementary nature of the interface and the positive feedback received during initial trials, a formal usability evaluation was considered unnecessary. The tool was successfully adopted by shopfloor personnel without the need for structured onboarding or detailed validation protocols. Its effectiveness lies in its ability to guide users through targeted questions and provide relevant support without overwhelming them with complex functionalities.

This approach reflects the specific operational context of Althofen, where the priority was to deliver a lightweight, accessible solution that could be deployed quickly and used effectively by non-expert staff. The simplicity of the HMI proved to be a key factor in its acceptance and usability.

4.3.2.3 Decision-making model preparation to validation

At the Althofen site, a dedicated validation framework has been defined to assess the performance of AI systems supporting maintenance operations. This framework focuses on evaluating the AI's ability to assist in machine fault diagnosis, repair recommendations, maintenance planning, and spare parts forecasting.

The framework is applicable to real-world use cases such as:

- Predictive maintenance
- Root cause analysis
- Repair and maintenance planning
- Spare parts demand forecasting

Key stakeholders involved in the validation process include:

- Maintenance engineers and technicians
- Production managers
- IT specialists
- Asset management teams

Unlike other FAIRWork use cases, such as CRF, which involve complex optimization models and structured usability testing, the Althofen case is centred on a lightweight, conversational AI tool designed for immediate use on the shop floor.

The goal of the validation is not to assess the user interface, but rather to evaluate the quality and reliability of the AI-generated decisions. Specifically, the framework aims to measure how useful, relevant, and applicable the AI recommendations are in real maintenance scenarios.

The evaluation will be based on a structured set of technical metrics (e.g., accuracy, response time, coverage of fault types) and operational metrics (e.g., downtime reduction, operator autonomy, impact on productivity), which will be detailed in the following sections.

This methodological approach ensures a comprehensive and realistic assessment of the AI system's impact on day-to-day decision-making, considering both its technical performance and its effectiveness as perceived by end users.

Technical Metrics

To evaluate the AI system's performance in supporting fault diagnosis and maintenance decision-making at the Althofen site, a set of technical metrics has been defined. These metrics focus on the system's ability to deliver accurate, timely, and reliable outputs under realistic operating conditions.

Criterion	Metric Description	Target
Diagnosis Accuracy	% of correctly identified root causes	From 60 to 100% descript in test strategy
Recommendation Quality	Expert agreement rate on suggested actions	> 85%
Response Time	Time to generate diagnosis	< 10s
System Uptime	Availability of AI service	> 99%

Table 5 Technical Metrics

These metrics will be used during the validation phase to assess the system's technical robustness and its ability to support real-time decision-making in maintenance operations.

The evaluation of diagnosis accuracy follows a structured three-step strategy, applied to the second prototype of the AI tool. Each step involves five participants interacting with the system using both written and spoken input (in English and German). Only written responses are scored; spoken input is recorded for qualitative analysis.

Test Strategy for Diagnosis Accuracy:

Step 1 - Standardized Questions

- Each participant submits 10 identical, predefined questions.
- Target: 100% correct answers (written input only)

Step 2 - Semi-Structured Questions

- Each participant formulates 2 free-form questions on five predefined topics.
- Target: 80% correct answers (written input only)

Step 3 - Open Questions

- Each participant submits 10 freely chosen questions based on real maintenance issues.
- Target: 60% correct answers (written input only)

This progressive approach allows the system to be tested across increasing levels of complexity and realism, ensuring that it can handle both structured and unstructured user input effectively.

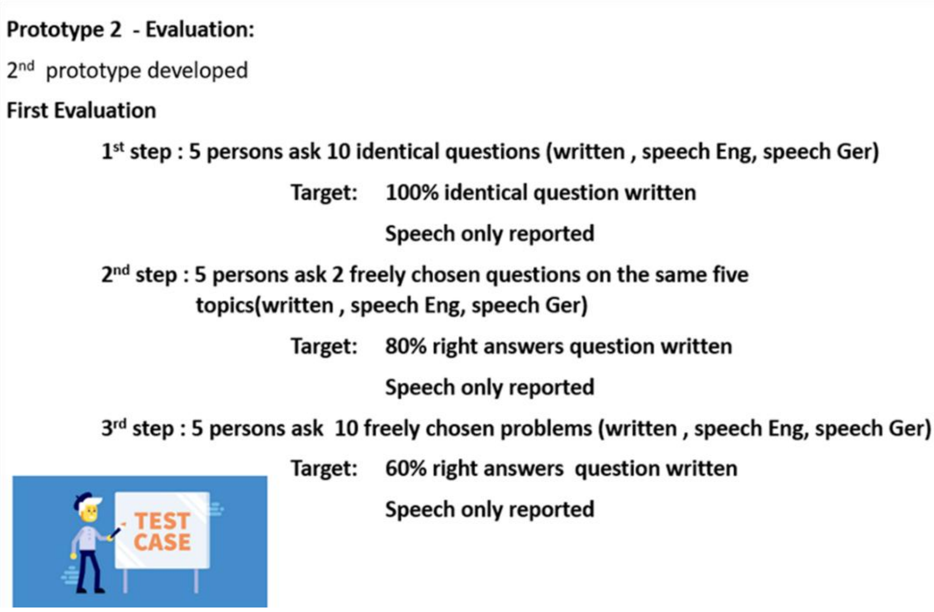

Figure 29 Test Strategy for Diagnosis Accuracy

Operational Metrics

In addition to technical performance, the validation framework for the Althofen use case includes a set of operational metrics aimed at measuring the real-world impact of the AI system on maintenance activities. These metrics focus on improvements in efficiency, responsiveness, and workload reduction for maintenance teams.

Criterion	Metric Description	Target
Maintenance Efficiency	Reduction in unplanned downtime	> 20%
Repair Time Reduction	Avg. time saved per intervention	> 15%
Service tickets	Reduction of Tickets	> 20%

Table 6 Operational Metrics

These criteria represent long-term KPIs, which cannot be fully evaluated within the timeframe of the FAIRWork project. Their measurement requires extended observation periods and integration into ongoing operational monitoring. All results related to these metrics, along with the outcomes of the technical validation, will be reported in detail in Deliverable 6.1.

4.4 Improve Reliability of "Documentation about Quality Check" (FLEX Timisoara)

4.4.1 Prototype Demonstration

This section outlines the FLEX - Document Transformation Service powered by large language models (LLMs). As noted earlier, the service currently demonstrates an intermediate level of maturity and a moderate degree of reliability (refer to Table 4 Summary of the AI-Services). Due to several challenges encountered during the demonstrator evaluation phase, development of the prototype was discontinued, and the solution has not progressed to the validation stage.

Introduction:

Work Instructions (WIs) are essential operational documents that clearly and precisely describe, step by step, how to perform a task within a manufacturing process. Using text and images, they guide operators at specific workstations, ensuring consistency, quality, safety, and effective training across the production flow.

At Flex Timisoara, most current WIs are still maintained in paper format, a legacy approach that poses challenges in today's fast-paced, digitally integrated manufacturing environment. As part of our factory digitalization strategy, we are transitioning to an electronic Work Instruction system (e-WI). This initiative involves converting the existing paper-based WIs into a standardized digital template compatible with the e-WI platform.

Switching from paper-based to electronic Work Instructions offers several key advantages:

- **Eliminates paper usage:** Reduces printing costs, physical storage needs, and environmental impact.
- **Improves revision control:** Ensures operators always access the latest version, minimizing errors from outdated instructions.
- **Speeds up approval processes:** Enables electronic signatures, streamlining document validation and deployment.
- **Enhances traceability:** Tracks changes and user interactions more effectively for audits and compliance.
- **Supports remote access:** Allows authorized users to view and update instructions from anywhere.
- **Enables system integration:** Connects with MES, ERP, and quality management platforms for a more responsive and connected manufacturing environment.

To accelerate the transition from paper to digital, a DAI-DSS LLM-based solution to automate the conversion process was proposed for evaluation. This will dramatically reduce the time, effort, and human error involved in manual conversion, allowing teams to focus on validation and optimization rather than transcription.

The proposed DAI-DSS solution has two main components:

1. **Extraction** of information and pictures from paper-based template file.
2. **Conversion** to new e-WI template

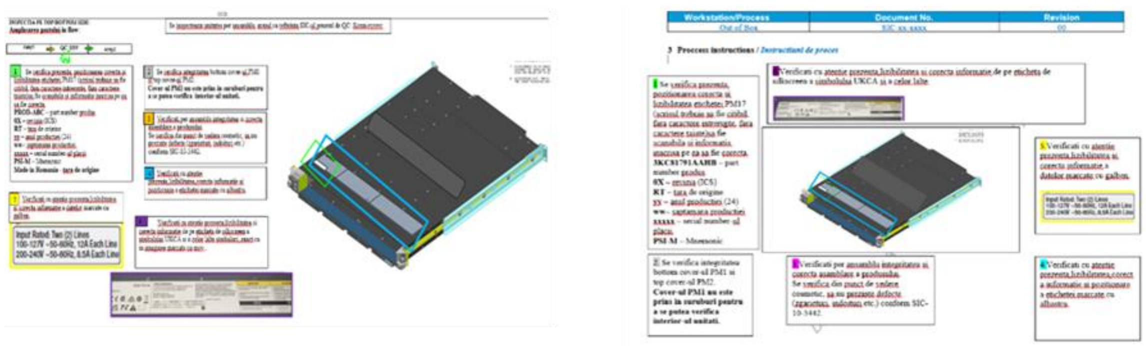

Figure 30 Work instruction example paper-based (left) vs e-WI (right)

HMI description:

Building upon the model generated through OLIVE (described at the end of section 2.2), this chapter describes the user interface (HMI) developed to support the Document Transformation service within the DAI-DSS system, with a specific focus on the Flex Timisoara use case. The interface serves as the main point of interaction between the end-user and the underlying AI services, offering an interactive visualization of the results produced by the decision engine.

The following sections will illustrate the main functionalities of the interface, the interaction flow with the system, and the operational steps that allow users to upload data, trigger the AI service, and visualize the results.

Opening AI Extraction component from Olive platform>

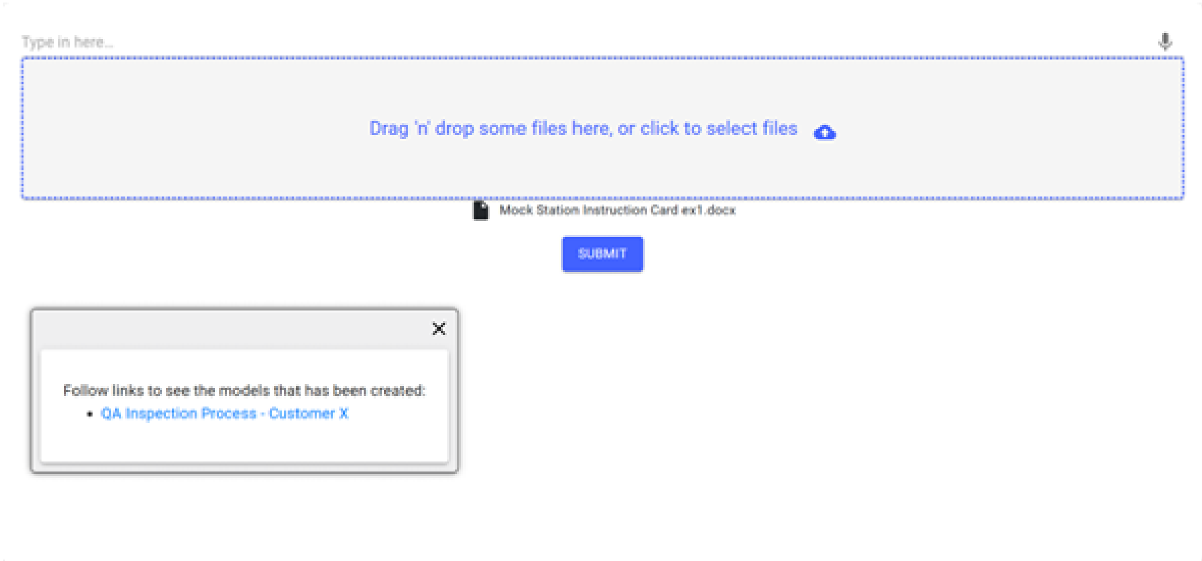

Figure 31 Upload WI doc in document transformation webpage

Then open a WI doc:

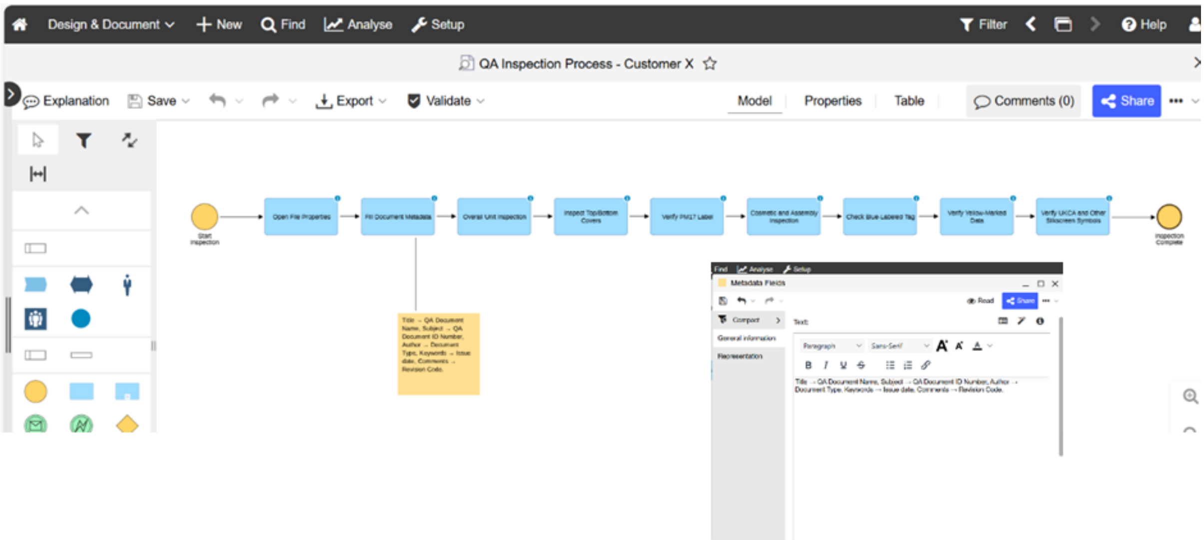

Figure 32 AI Extraction output

4.4.2 Demonstration Set-up, Testing and Feedback

The Timisoara site concentrated on improving the validation of calibration certificates and document transformation related to quality control. The goal was to enhance the reliability and traceability of technical documentation using AI-powered parsing and rule-based validation. This approach aims to reduce manual errors, streamline verification processes, and support compliance with industry standards. As mentioned in previous sections, the proposed DAI-DSS solution for “Document Transformation” is developed in 2 steps having 2 main components:

1. Extraction of information and pictures from paper-based template file.
2. Conversion to new e-WI template.

The first part focusing on the extraction of information and pictures from word templates was developed and evaluated. Before implementing second part and transferring information into the new template, first part must return the intended outputs, otherwise it is not feasible to proceed with second part.

Due to the high level of unstructured input information during evaluation of the first part AI Extraction service performed by BOC and analysed together with FLEX the following challenges were identified:

1. Input documents must have a .docx format to be processable.
2. Color-codes cannot be interpreted and mapped. In WI example - cannot be linked the label from yellow square to step 5 and label from magenta square to step 6.
3. Relations of textboxes within textboxes cannot be identified. In WI example - for step1 – text box with number 1 cannot be associated/linked with textbox with the textual description, they are treated as two different objects with no links between them.
4. Coloured boxes placed over the main picture in WI, highlighting specific areas are considered as 2 different objects which are not included because are not identified as part of the picture.
5. Objects added on top of the pictures when edit WI word document, labels from main picture cannot be linked to the pictures they are considered as separated objects and are not included when AI algorithm is looking for the pictures.
6. The mechanism behind the applied AI interprets the text and defines the number of steps in the process non-deterministically meaning that output can vary for each request.

To enable better results of the AI extraction service, the input documents must be pre-processed to ensure a higher level of structure so that a better machine-readable format can be provided to the AI service. Some recommendations for improvement are:

- change word format from .doc to the .docx,
- transform multiple boxes placed over the pictures to single pictures,
- avoid color-coding,
- place less emphasis on pictures and more on textual descriptions.

The impact of the changes on the resulting output should be evaluated through iterative testing.

Given the challenges outlined above, it was necessary to re-evaluate both the feasibility of applying AI to this use case and the appropriateness of using large language models (LLMs) for document transformation.

To implement the recommended preprocessing of original document files so that AI can use it, demands considerable effort increasing manual workload which is equivalent to not using AI. Given the uncertainty of achieving meaningful or high-quality AI outcomes, the effort invested in preprocessing—and the additional development required for the AI system—may not justify the expected benefits, as the overall workload could remain comparable to traditional methods. As a result, the development of the Document Transformation service was halted after its initial phase, and the use case was ultimately stopped and no further efforts in development of HMI and decision-making model validation was performed.

4.5 Support Validation of Calibration Documents (FLEX Timisoara)

4.5.1 Prototype Demonstration

This section provides a detailed description of the FLEX – Calibration Certificate Service based on Rule Based Algorithm which reach an advance level of maturity and a high degree of reliability (refer to Table 4: Summary of AI Services) which was implemented at the Timisoara site.

Introduction:

The current manual verification of calibration certificates is a time-consuming, repetitive, and error-prone process that demands meticulous attention to critical data points. Employees must manually check instrument details (model, type, manufacturer, serial number), calibration dates (calibration and expiry), calibration results (measured values, deviations, uncertainties), measuring equipment information (for traceability), and formal aspects (signatures, page numbering, completeness). This exhaustive manual review, exacerbated by the sheer volume of certificates, significantly increases the risk of human error. Overlooked discrepancies can have severe repercussions, potentially leading to the use of faulty measuring equipment and compromising product or service quality.

To address these challenges and optimize the process, we propose developing an AI-powered application to automate the verification of calibration certificates. This application will leverage AI to analyse certificates, identify discrepancies, and automatically generate a comprehensive report highlighting any missing or incorrect information. This automated solution is expected to drastically reduce review time and eliminate the likelihood of human error, leading to a much higher level of accuracy and reliability in the certificate review process. Ultimately, this will bolster internal quality assurance and allow staff previously dedicated to manual verification to be reassigned to more value-adding activities.

It was defined 14 URS (User Requirement Specification) to be checked by AI Service during Calibration Certificate verification:

URS	Name	Verification
URS01	Issue Date	not earlier then Calibration Date
URS02	Device Name	same as in CCDB
URS03	Device Manufacturer	same as in CCDB
URS04	Device Model	same as in CCDB
URS05	Device Serial Number	same as in CCDB
URS06	Device ID	same as in CCDB
URS07	Calibration Date	same as in CCDB
URS08	Calibration Due Date	same as in CCDB
URS09	Standards Due Date	not later than Calibration Date
URS10	No. of Pages	presence of all pages

URS	Name	Verification
URS11	Overall Results	all summary results are Pass
URS12	Measurement Results	all individual measurements results are Pass
URS13	Executor Signature	presence of signature
URS14	Approval Signature	presence of signature

Table 7 URS definition

The application identifies and interprets data from the Calibration Certificate PDF file, localizing each field associated with the respective URS. It then compares these fields against our Calibration Database (CDB), which is represented as an Excel file reflecting our local SQL Database, to verify the presence and accuracy of information as specified for each URS in the table above.

An example of mapping of URS to the Calibration Certificate, below an example of URS mapping:

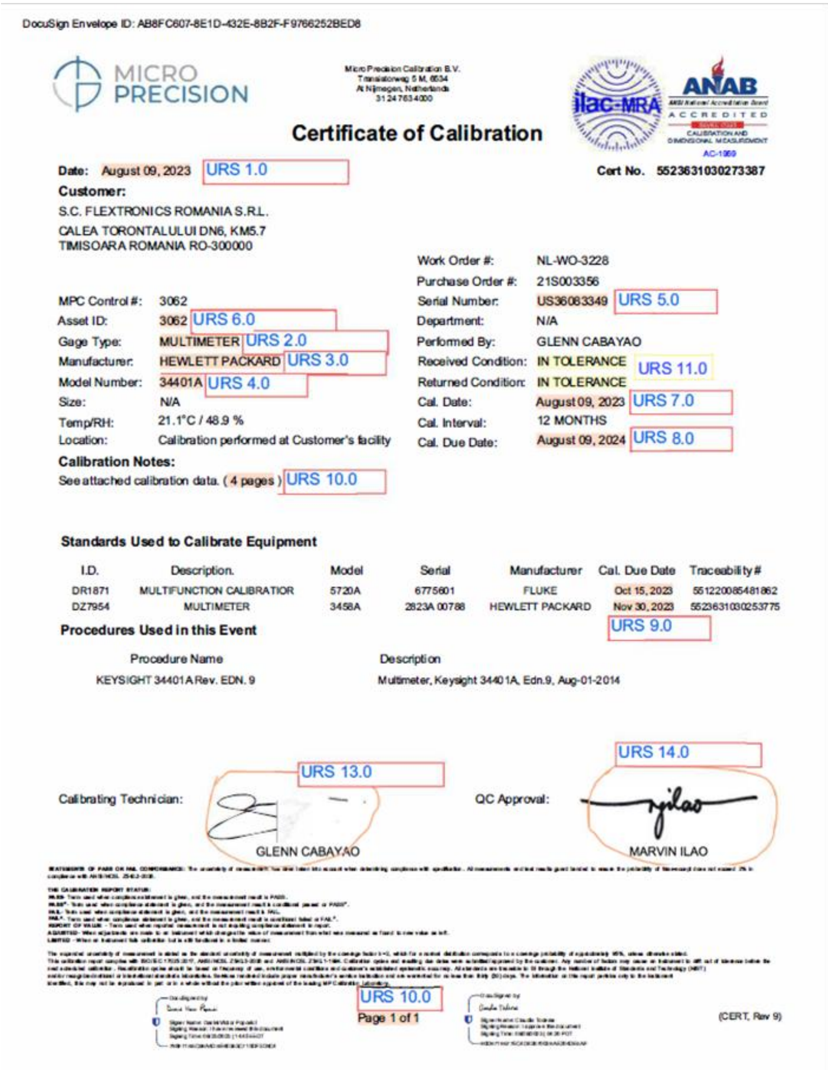

Figure 33 Example of URS mapping

HMI description:

Building upon the model generated through OLIVE (described at the end of section 2.3), this chapter describes the user interface (HMI) developed to support the Calibration Certification Service within the DAI-DSS system, with a specific focus on the FLEX use case in the Timisoara site.

In the case of the Calibration Certificate Service, the application is a standalone application.

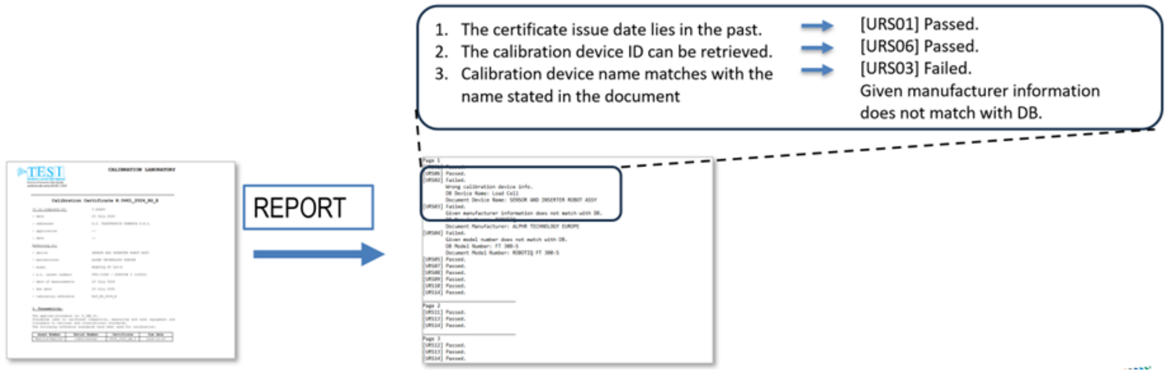

Figure 34 Calibration Certificate Service HMI Sketch

The user interface is designed to support the export user. In the application, a series of documents can be transferred to a specific file. The files are then analysed, and a report is generated. A report is then generated for each individual calibration document, which can be viewed in the browser. This shows which requirements have been successfully met in the test.

4.5.2 Demonstration Set-up, Testing and Feedback

The Calibration Certificate Verification Service, currently operating as a standalone executable, was successfully deployed on two devices for initial testing:

- A desktop PC located in the Flex Calibration Laboratory
- A laptop designated for evaluation and field testing

4.5.2.1 Preliminary test and feedback

Initial Testing - Small Batch Evaluation: a preliminary batch of up to 50 Calibration Certificates from the first targeted category was processed using the prototype. The goal was to:

- Observe the service's response and behaviour
- Verify the accuracy of both individual certificate reports and the aggregated summary report

Identified Areas for Improvement: based on the initial results, several enhancement opportunities were identified to improve the tool's performance and usability:

- Improved URS Detection: enhance recognition of User Requirement Specifications (URS) and implement logic to ignore or correctly process irrelevant characters (e.g., case sensitivity, extra spaces, special formatting)
- Use Case Refinement: reassess and redefine certain use cases to better align with real-world scenarios and certificate variations
- Enhanced Output Reporting: improve log structure for faster failure identification and include summary statistics to support data-driven analysis

- Calibration Database Cleanup: standardize and adjust entries in the Calibration Certificate Database to ensure consistency and compatibility with the tool

Extended Testing - Large Batch Validation: after implementing the preliminary corrections, a larger batch of up to 1,000 certificates was processed. The results were highly encouraging:

- The tool demonstrated very good reliability and validity
- Output consistency and detection accuracy were significantly improved
- The concept proved scalable and ready for extension to additional types of Calibration Certificates

This successful evaluation confirms the potential of the service to support broader digitalization efforts and reduce manual workload in calibration certificate validation.

4.5.2.2 HMI preparation to validation

The Calibration Certificate Service currently operates as a standalone executable file, without an integrated Human-Machine Interface (HMI). In its present form, switching between different types of calibration certificates requires manual editing of a traditional configuration file, an approach that is functional but not user-friendly.

To enhance usability and efficiency, future improvements are being considered, including the development of a graphical HMI. This interface would offer:

- A visible progress bar, especially useful when processing large volumes of certificates
- An intuitive method for switching between certificate categories, eliminating the need for manual configuration file edits

These enhancements aim to streamline the user experience and support broader adoption of the service across various calibration certificate types.

4.5.2.3 Decision-making model preparation to validation

Following successful preliminary testing, the Calibration Certificate Verification Service has been identified as the most promising candidate for full validation within the FAIRWork project in Flex Timisoara. This AI-powered tool is designed to automatically detect missing critical information or failed results in calibration certificates, helping reduce manual effort and human error.

To ensure the robustness and reliability of the service across different certificate formats, a structured multi-step validation process will be carried out:

Step 1: Initial Behaviour Testing

A small sample of randomly selected calibration certificates will be processed to observe the prototype's behaviour. This phase will focus on evaluating the stability and repeatability of the tool's output under controlled conditions.

Step 2: Large-Scale Evaluation

A full batch of calibration certificates from a single category (year 2024) will be tested to assess performance at scale. This step will help identify potential issues such as false calls, inconsistencies, or edge cases that may arise when processing large volumes of data.

All findings will be documented and tracked using the CoC Action Tracker.

The evaluation will be based on certificates from the 2024 dataset, using an Excel export of the web-based SQL Calibration Database (CDB).

Step 3: Measurement System Analysis (MSA)

To rigorously assess the tool's stability and repeatability, an Attribute Study will be conducted using the MSA methodology. The study will include:

- Sample size: 20 calibration certificates: 10 verified certificates with no known issues and 10 certificates with intentionally induced errors across different fields, covering all User Requirement Specifications (URSs)
- Each certificate will be processed 23 times to ensure statistical confidence, targeting 90% confidence and 99% reliability levels.

Step 4: Output Optimization

To achieve a false call rate below 5%, fine-tuning will be required in two key areas:

- Enhancing AI detection capabilities: Improving the algorithm's accuracy in identifying errors and missing data
- Standardizing the Calibration Certificate Database: Cleaning and harmonizing data formats to reduce variability and improve processing consistency

This validation roadmap is essential to ensure that the Calibration Certificate Verification Service can be reliably deployed across different certificate categories, supporting Flex Timisoara's broader goals of digitalization and operational excellence.

All results related to these metrics, along with the outcomes of the technical validation, will be reported in detail in Deliverable 6.1.

5 SUMMARY AND CONCLUSIONS

This chapter summarizes the main results of the demonstration, highlights the observed benefits and emerging issues. Provides recommendations for the future adoption of DAI-DSS in real production contexts and suggests possible future developments of the system.

The demonstration phase of the FAIRWork project has confirmed the feasibility and added value of deploying the DAI-DSS system in real industrial environments. Across all use cases, the system has shown its potential to support human-centered decision-making by integrating AI services, legacy systems, and human expertise through a modular and transparent architecture.

The hybrid deployment model, combining cloud-based orchestration with local data handling—proved effective in addressing both technical and organizational constraints. The involvement of end-users throughout the co-creation and validation phases was essential to ensure usability, trust, and alignment with operational needs.

Despite differences in context and maturity across the use cases, all partners contributed to validating the system's ability to:

- Enhance decision transparency and traceability.
- Improve responsiveness to dynamic production conditions.
- Support less experienced personnel in complex tasks.
- Reduce manual effort in repetitive or error-prone processes.

5.1 CRF

The CRF use case focused on three main scenarios: Production Planning, Workload Balance, and Truck Loading. These were implemented through two AI-powered services:

- **Production Planning Service with a Hybrid Approach**, integrating Constraint Programming and Reinforcement Learning.
- **Truck Loading Service**, based on heuristic optimization.

The demonstrators were tested through structured usability and decision-making validation protocols involving real plant personnel. The validation methodology, covering usability (SUS), task completion, and decision quality, has been fully defined and is described in detail in Deliverable D5.3.

The actual results of the validation activities will be reported in Deliverable D6.1, which will include quantitative and qualitative assessments of system performance, user acceptance, and decision support effectiveness.

Benefits:

- Improved planning accuracy and flexibility.
- Enhanced transparency in decision-making.
- Potential reduction in unplanned absenteeism and energy consumption (to be verified in D6.1).

Critical Issues:

- Need for integration with existing MES systems.
- Importance of explainability to foster trust in AI-generated suggestions.
- Cultural resistance to automated decision support in some roles.

5.2 FLEX Althofen

As part of a strategic innovation initiative, the Althofen site focused on the development and implementation of an AI-supported solution for machine maintenance. The primary objectives were to significantly reduce unplanned downtime and to provide targeted assistance to less experienced operators through intelligent support systems.

The developed solution combines a structured knowledge base with a conversational user interface. This interface leverages advanced technologies such as Large Language Models (LLMs) in conjunction with Retrieval-Augmented Generation (RAG) to deliver context-aware, accurate, and traceable information in real time. This enables efficient fault diagnosis and guided support during maintenance and repair activities.

The demonstrators were tested directly with plant personnel, hands-on and using clearly defined protocols for usability and decision-making. The full approach (including SUS scores, task completion checks, and decision quality assessments) is described in detail in Deliverable D5.3.

The outcomes of WP5 confirm the technical feasibility of the solution and its potential to support maintenance operations effectively. The activities carried out in this work package laid the foundation for the subsequent validation phase in WP6.

The real-world evaluation and performance assessment of the system will be addressed in WP6. Deliverable D6.1 will present the quantitative results, user feedback, and a comprehensive overview of how the system performs under operational conditions, how it is perceived by users, and how effective the AI support proves to be.

5.3 FLEX Timisoara

The Timisoara site addressed document validation, with a focus on calibration certificates and document transformation. The goal was to support digitalization efforts and improve operational efficiency, the reliability and traceability of quality-related documentation using AI-based document parsing and rule-based validation.

- **Document Transformation using LLM**, based on large language models (LLMs), aimed to convert legacy paper-based Work Instructions into a standardized digital format compatible with the Flex e-WI system.
- **Calibration Certificate Service**, powered by a Rule Base algorithm for DAI-DSS, was developed to automatically detect missing, invalid data or fail results in calibration certificates.

The demonstrators were evaluated through internal testing and prototype validation activities. While the Calibration Certificate service showed promising results and go further, the Document Transformation service faced significant challenges and was paused prior to full validation. Details of the evaluation methodology and service maturity are outlined in Deliverable D5.3, while final validation results for Calibration Certificate service, including performance metrics, user feedback, and decision support effectiveness, will be presented in Deliverable D6.1.

Benefits:

- Increased reliability and speed in certificate verification
- Reduction of manual effort and human error
- Support for digital transformation goals

Critical Issues:

- High variability and inconsistency in WI document formats (Document Transformation)
- Limited feasibility of automated transformation without extensive preprocessing (Document Transformation)
- Need for user-friendly interfaces and integration with existing SQL Calibration Data Base (Calibration Certificate)

5.4 Recommendations and Outlook

Building on the insights gained from the demonstration activities, this section outlines a set of recommendations for future adoption of the DAI-DSS system, as well as possible directions for its further development. These reflections are based on the feedback collected from end-users, the technical integration experience, and the lessons learned across the different use cases.

The experience gained during the demonstration phase has highlighted several key factors that should be considered to ensure the successful adoption of the DAI-DSS system in real industrial environments.

First and foremost, the early involvement of end-users, from operators to decision-makers, has proven essential. Their feedback not only helped shape the design of the interfaces and services but also increased acceptance and trust in the system. Future deployments should continue to adopt a co-creation approach, ensuring that the tools developed are aligned with real operational needs.

Another important aspect is the explainability of AI decisions. While the system has demonstrated strong technical capabilities, some users expressed the need to better understand the rationale behind the suggestions provided. Enhancing transparency through contextual explanations or visual cues will be crucial to foster confidence and encourage broader use.

From a technical standpoint, integration with existing MES and ERP systems remains a priority. Seamless data exchange and interoperability will reduce duplication of effort and ensure that the DAI-DSS becomes a natural extension of current workflows, rather than an isolated tool.

Training and onboarding also emerged as critical success factors. Even the most intuitive interfaces benefit from targeted training sessions, tutorials, or embedded guidance, especially when introducing AI-based decision support in traditionally manual environments.

Looking ahead, several development directions can be envisioned. These include:

- Extending the system to cover additional decision-making scenarios and departments.
- Enabling real-time data integration, allowing the system to react dynamically to changes on the shop floor.
- Improving personalization and adaptability, so that the system can tailor its suggestions based on user roles, preferences, or historical behaviour.
- Exploring cross-site learning, where insights and models developed in one facility can be adapted and reused in others, fostering knowledge sharing across the organization.

In conclusion, the FAIRWork DAI-DSS has demonstrated its potential to become a cornerstone of intelligent, inclusive, and resilient decision-making in manufacturing. With continued refinement and alignment to user needs, it can evolve into a robust support system capable of adapting to the complexity and variability of modern production environments.