

# Towards Industry 5.0 – A Trustworthy AI Framework for Digital Manufacturing with Humans in Control

Usman Wajid<sup>1</sup>, Alexandros Nizamis<sup>2</sup>, and Victor Anaya<sup>1</sup>

<sup>1</sup>Information Catalyst for Enterprise, UK

<sup>2</sup>Centre for Research and Technology Hellas - Information Technologies Institute (CERTH/ITI), Themi, 57001, Thessaloniki, Greece

## Abstract

Despite the fact that Industry 4.0 concepts and technologies are still being developed and still being adopted, the lessons learned from the last decade have helped shape up the notion of Industry 5.0 - as the next ‘revolution’ in industrial domain. Even though Industry 5.0 shared many concepts with Industry 4.0, it is characterized by three main elements, human-centricity, sustainability and resilience. In this paper, we introduce a digital manufacturing platform architecture that extends Industry 4.0 paradigms to enable AI-based decision support with the necessary trustworthiness and human-centricity elements primed for Industry 5.0. The proposed architecture helps realize the balancing act of getting the perceived benefits from AI-centric digitalization while preserving the role of humans in key decision-making activities.

## Keywords

Industry 5.0, trustworthy AI, human-centricity.

## 1. Introduction

Nowadays, manufacturing industry is still the largest contributor to European economy. Even though the European industry contributes about 20% of European Gross Domestic Product (GDP), there are still many challenges that should be tackled. Therefore, European Commission (EC) has already prepared a series of documents that introduces Industry 5.0, in the same time that Industry 4.0 activities are still in deployment or even in development stages. In this direction, EC has already described the enablers underpinning the notion of Industry 5.0 [1], which include human machine interactions, smart materials, digital twins, data analysis, energy efficiency and Artificial Intelligence (AI). In this respect, EC has established [2] three main elements for Industry 5.0, human centricity, resilience and sustainability.

In addition, the EC’ is also active on promoting the adoption of AI across various sections and a High-Level EC Expert Group on AI (HLEG) has introduced an ethical framework and guidelines for a trustworthy AI [3]. Trustworthy AI means that a system is, lawful (respecting all applicable laws and regulations), ethical (respecting ethical principles and values) and robust (both from a technical and social-environment perspective). The HLEG guidelines indicate seven key requirements that when met, an AI system would be considered trustworthy. Moreover, a trustworthy system could be considered as human-centric that is a core element for Industry 5.0, when it implements or complies with four main principles that are Respect for Human Autonomy, Prevention of Harm, Fairness and Explicability.

Keeping in mind the close relationship between Industry 5.0 and human-centric AI, in this paper we are introducing the conceptual architecture of a digital manufacturing platform that supports

---

Proceedings of the Workshop of I-ESA’22, March 23–24, 2022, Valencia, Spain

EMAIL: usman.wajid@informationcatalyst.com (U. Wajid); alnizami@iti.gr (A. Nizamis); victor.anaya@informationcatalyst.com (V. Anaya)

ORCID: 0000-0001-7237-0658 (U. Wajid); 0000-0002-6954-4861 (A. Nizamis)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

trustworthy AI services to be provisioned in human-centric manufacturing processes based on Industry 5.0 concepts and HLEG guidelines. The architecture of the envisioned knowlEdge (Towards AI powered manufacturing services, processes, and products in an edge-to-cloud-knowledge continuum for humans in-the-loop) platform [4, 5] aims to enhance AI functionalities coming from Industry 4.0 with a human-in-the-loop approach in order to address the resilience and human-centricity elements of Industry 5.0.

To this aim, the paper presents the key elements of the introduced architecture and maps the architecture components and functionalities with HLEG requirements for a Trustworthy AI towards Industry 5.0. Following the introduction, the remainder of the paper is structured as follow: In chapter 2, the conceptual architecture is presented alongside with a short description of its layers and components. The way that the introduced architecture addresses the requirements of a Trustworthy AI towards Industry 5.0 is presented in chapter 3. The conclusions are drawn on chapter 4.

## 2. Conceptual Architecture of the knowlEdge Platform

The conceptual architecture of the knowlEdge platform is illustrated in Figure 1. It is composed of 4 hierarchal layers that realize data integration, data analysis, knowledge management and smart decision support functionalities. A vertical layer covering all horizontal aspects provide the necessary integration, security and policy support. A brief overview of the architectural layers and the components is given in Figure 1.

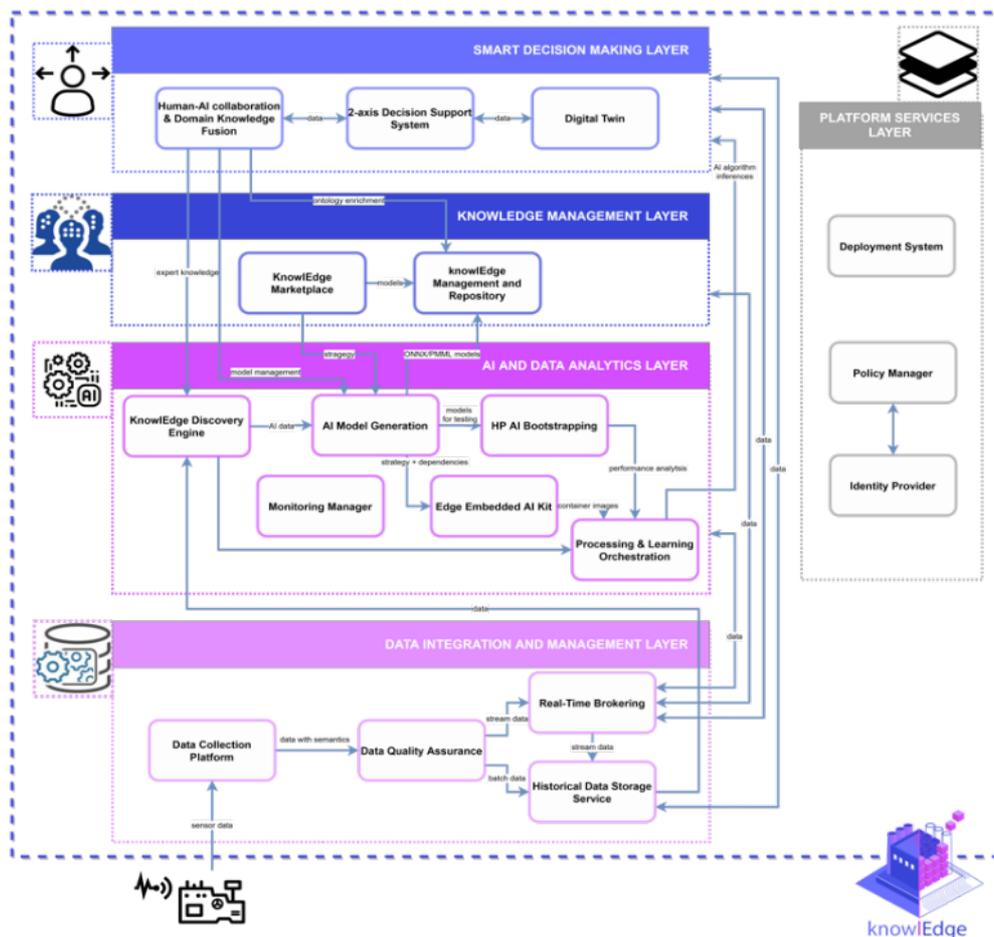


Figure 1: knowlEdge Platform Architecture.

### Data integration and management

This layer focuses on establishing connectivity with underlying (manufacturing) as-sets. The components relevant for this layer include:

- **Data Collection Platform:** This component collects the data from the shopfloor assets and make them available at the other layers. The data are transferred based on standardized data models and communication protocols
- **Data Quality Assurance:** This component is responsible to perform data quality assessment and provide some data pre-processing functionalities like cleaning etc.
- **Historical Data Storage:** A scalable time-series data storage component responsible for storing the historic data across the knowlEdge compute continuum.
- **Real-Time Brokering:** This component utilizes high-performance message broker facilitating real-time data propagation across different services in the platform using a publishing/subscribing model.

### **AI and data analytic**

This layer supports data analysis and knowledge management, components include:

- **Knowledge Discovery Engine:** This component is responsible to give feedback based on interpretable, aggregated and discoverable information.
- **AI Model Generation:** This component provides an automated continuous process of training AI-based analytic models and testing them based on synthetic datasets. The models can be executed at all stages of the compute continuum, edge, fog and cloud.
- **Edge Embedded AI Kit:** It represents an edge-ready lightweight module for inference functionalities from AI modes.
- **High Performance AI Bootstrapping:** This component supports the automated execution of AI models in the cloud with embedded performance analysis
- **Processing & Learning Orchestration:** The component enables fetching and orchestration of the AI models based on the specific needs and availability of necessary resources in the compute continuum.
- **Overall Monitoring:** This central component serves as the main access point to all the other components in this layer

### **Knowledge management layer**

This layer focuses on knowledge use and exploitation – components include:

- **Knowledge Management Repository:** This component provides technical means to ensure efficient management and access of data and knowledge generated and gathered in the platform. Furthermore, it provides the knowledge representation based on standardized models like PMML [6].
- **Knowledge Marketplace:** This component provides the means for the exploitation of (developed and tested) AI models' through search capabilities, rating and feedback mechanisms to support new business models and to enable the AI models trading through web based UIs.

### **Smart decision making**

This layer bring together data processing and AI analytic functionalities in a set of human-centric decision support mechanisms.

- **Digital Twin:** This component simulates the behaviour of the manufacturing processes based on AI algorithms, with output going in 2D and 3D dashboards
- **Human-AI collaboration & Domain Knowledge Fusion:** This component offers an interface for domain experts to interact with AI models and underlying configuration so to enter ground truth, configuration adaptation and model selection
- **2-axis Decision Support System:** This component represents a dynamic and interactive dashboard composed of several topics, such as machine use, alerts to identify areas where a fault or malfunction occurs and management info (maintenance time, operational costs, machine performance, etc.) as measured by key indicators. The visualized KPIs can be created and calculated based on the need of each shop-floor and multiple levels (product and process-level) to evaluate their success at reaching targets.

### Platform services layer

This layer packages the above layers to realize a multi-side platform. Components are:

- Policy Manager: This component guarantees the availability of knowledge and the integrity and confidentiality of the information shared in the platform. The policies defined (e.g. for access and information use) need to be enforced by all components.
- Identity Provider: This component provides a single sign-on mechanism for the users who are interested in interacting with the other component. It also provides role management functionalities.
- Deployment Manager: This component enables users to deploy software on remote edge devices, providing utilities to support bulk deployments, remote command execution, cloud deployment and monitoring of the deployments.

### 3. Addressing the Requirements of a Trustworthy AI System

This section analyses the compliance of the proposed knowlEdge platform with regards to the HLEG key requirements (see Figure 2.) for trustworthy AI systems, which is one of the main pillars for Industry 5.0.



**Figure 2:** Requirements for a Trustworthy AI. Figure from [3]

The following table explains how the components of the knowlEdge platform address these seven key requirements of a trustworthy AI system.

Requirement	Platform's Response
1. Human agency and oversight	The 2axis Decision Support System brings human in-and-on the loop through the implementation of eXplainable AI (XAI) concepts that ensure that information about AI models e.g. AI system's results, suggestions and notification are presented to the human users as decision support. A SotA user interface can enable users to interact with the system, improving the understanding of the data, as well as the evaluation process of the AI decisions and proposals. <b>Human-AI Collaboration &amp; Domain Knowledge Fusion</b> enables users to inject human knowledge to AI system through easy-to-use and intuitive UI. Consequently, human users will have the chance to intervene in every step of the decision cycle of the system. The human involvement can cover not only the ability to inject knowledge into the AI models, but

	also the ability to select models from Marketplace, compare them and select the best one.
2. Technical Robustness and safety	The <b>Identity Manager</b> component enforces privacy-by-design to minimize damage from potential security threats. Together the Identity Manager and Policy Manager components can provide a security framework for data security, role and access management. The <b>Digital Twin</b> framework will enable the virtual representation of manufacturing process; the users will be able to explore the outcomes of various decisions before applying them on real world scenarios. In this respect, DT can provide a mechanism for testing, verification and validations of system's reliability and robustness. Robustness and safety can also be ensured through <b>Human-AI Collaboration &amp; Domain Knowledge Fusion</b> , which allows the injection of human knowledge for security and safety considerations as well optimization of relevant parameters.
3. Privacy and data governance	The <b>Policy Manager and Identity Provider</b> components can provide security and access protocols to protect sensitive and private data from unauthorized access e.g. Marketplace, DSS etc. can only be accessed by authorized users. The <b>Data Quality Assurance</b> component can also help with data integrity and governance through relevant functions. Moreover, the <b>Edge Embedded AI Kit</b> will help with management of data at the source and ensure only valuable data goes upward.
4. Transparency	The <b>Knowledge Management Repository</b> can ensure traceability of data and AI techniques by storing relevant information and metadata about data, models and algorithms. Furthermore, the use of blockchain technology in the <b>Knowledge Marketplace</b> can log immutably the various versions of AI models and the users' updates. AI model bootstrapping over <b>Digital Twin</b> can ensure the best possible degree of transparency related to selected models and actions in hypothetical scenarios. Special focus can be given in the design of <b>Decision Support System's XAI</b> part for providing further information to end user about models used and rules followed for the delivery of any suggestions, recommendations and notifications through DSS interfaces.
5. Diversity, non-discrimination and fairness	The <b>Policy Manager and Identity Provider</b> can ensure that knowlEdge services are accessible in a fair way and not denied to people based on their age, gender, abilities or characteristics. The platform also promotes the stakeholders' active participation and use of the system by providing human-centric components.
6. Societal and environmental well-being	<b>DSS and AI Model Generation</b> components, can consider various aspects such as optimization of resource allocation and energy consumption as key priorities for models' evaluation and analysis. AI Bootstrapping and Processing & Learning Orchestration can provide best practices on deployment decisions, parallelization, code optimization to maximize efficiency and environmental friendliness across different computational frameworks
7. Accountability	The responsibility and accountability for AI modules, their auditability etc. can be done in components like <b>Data Quality Assurance and AI Bootstrapping</b> . Simulation scenarios within the <b>Digital Twin</b> component can also support assessments. Rating, feedback and comments about available AI models in the knowlEdge Repository can also be available through <b>Marketplace</b> .

## 4. Conclusions

This paper introduces the conceptual architecture of a digital manufacturing platform that balances AI-based decision making with human-centric explainability and trustworthiness – as desired in Industry 5.0. The components of the architecture are introduced, together with the justification on how they can ensure compliance with the recommendations of the EC Expert Group on Trustworthy AI. The architecture establishes a link between the digitalization in Industry 4.0 with the human-centricity, resilience and sustainability concepts being promoted under Industry 5.0.

As next steps are considered the development of the described concepts and technologies and their application and validation to different pilot cases related to human-AI collaborative smart planning and production processes' optimization based on edge computing and human-knowledge injection.

To sum up, knowlEdge platform was initially designed for the needs of Industry 4.0. However, based on the platform's responses to requirements for a Trustworthy AI it is perceived that Industry 5.0 it is not that far from technological perspective as most of the technologies are already there. Therefore, what is needed in order to move from Industry 4.0 to 5.0 is to use the available technologies targeting concepts like human-centricity, sustainability and resilience.

## 5. Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957331 - KNOWLEDGE.

## 6. References

- [1] J. Müller, Enabling Technologies for Industry 5.0 - Results of a workshop with Europe's technology leaders, 2020. URL: <https://op.europa.eu/en/publication-detail/-/publication/8e5de100-2a1c-11eb-9d7e-01aa75ed71a1/language-en>.
- [2] M. Breque, L. De Nul, A. Petridis, Industry 5.0 - Towards a sustainable, human-centric and resilient European industry. 2021. URL: [https://ec.europa.eu/info/news/industry-50-towards-more-sustainable-resilient-and-human-centric-industry-2021-jan-07\\_en#:~:text=Search-,Industry%205.0%3A%20Towards%20more%20sustainable%2C%20resilient%20and%20human%2Dcentric,drivers%20of%20this%20dual%20transition](https://ec.europa.eu/info/news/industry-50-towards-more-sustainable-resilient-and-human-centric-industry-2021-jan-07_en#:~:text=Search-,Industry%205.0%3A%20Towards%20more%20sustainable%2C%20resilient%20and%20human%2Dcentric,drivers%20of%20this%20dual%20transition).
- [3] European Commission, High-Level Expert Group on AI (HLEG): Ethics Guidelines for Trustworthy Artificial Intelligence, 2021. URL: <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>.
- [4] S. Alvarez-Napagao, B. Ashmore, M. Barroso, C. Barrue, ... L. Ziliotti, knowlEdge Project – Concept, Methodology and Innovations for Artificial Intelligence in Industry 4.0, in: Proceedings of the 19th International Conference on Industrial Informatics, IEEE, New York, 2021, pp. 1-7, doi: 10.1109/INDIN45523.2021.9557410.
- [5] knowlEdge Project, Welcome to knowlEdge Project Homepage!, 2022. URL: <https://www.knowledge-project.eu/>.
- [6] DMG, PMML 4.4.1 – General Structure, 2022. URL: <https://dmg.org/pmml/v4-4-1/GeneralStructure.html>.