Agentic AI for Digital Twin

Alexander Timms[1]*, Abigail Langbridge[2]*, Antonis Antonopoulos[3], Antonis Mygiakis[3], Eleni Voulgari[3], Fearghal O'Donncha[4]

[1]Department of Bioengineering, Imperial College London, UK
[2]Dyson School of Design Engineering, Imperial College London, UK
[3]Konnecta, Kifissia, Athens, Greece
[4]AI for Sustainable Industry, IBM Research Europe - Ireland
alexander.timms21@imperial.ac.uk, al4518@ic.ac.uk, antonis.antonopoulos@konnecta.io, antonis.mygiakis@konnecta.io, eleni.voulgari@konnecta.io, feardonn@ie.ibm.com

Abstract

The complexity of the shipping industry, dynamic operational drivers, and diverse data sources present significant scalability challenges for digital twins. Agentic Large Language Models (LLMs) augmented with external tools offer a promising solution to accelerate digital twin adoption. Using pre-trained knowledge and reasoning capabilities, these LLMs autonomously select optimal tools and data streams for user-specific queries, enabling language to serve as a universal interface between digital twins and various stakeholders, from technicians to fleet managers. This interface facilitates real-time decision making and insight generation across multiple operational workflows. In this demonstration, we present an interactive agentic digital twin designed to enhance scalability, flexibility, and efficiency in managing the extensive and intricate decision-making requirements of the shipping industry. We showcase the transformative potential of agentic LLMs in reducing complexity and improving the practical application of digital twins, ultimately enabling more efficient operations in real-world settings.

Introduction

Digital twin technologies hold transformative potential across industries such as shipping, manufacturing, and smart cities (Katsoulakos et al. 2024). However, their implementation faces challenges, particularly in architectural design, data integration, and ensuring accessibility for diverse stakeholders. Managing dynamic data, performing real-time analytics, and ensuring interoperability across cloud and edge platforms often lead to inconsistencies and maintenance challenges (O'Donncha et al. 2024). Furthermore, integrating diverse data sources—including IoT devices, operational systems, and legacy platforms in both structured and unstructured formats—while ensuring accuracy adds significant complexity. Digital twins must provide actionable insights tailored to the diverse needs of stakeholders, driving transformative impact and enabling effective change management across organizational structures.

*Work completed while at IBM Research Europe

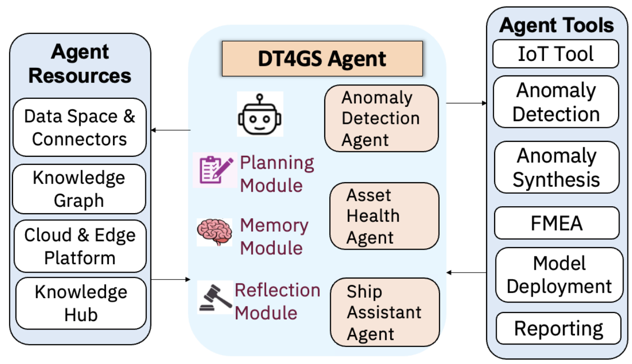

Figure 1: The agentic workflow illustrates user interaction with the agentic system. The planning module manages agent selection and orchestrates relevant tool invocations, while the reflection module evaluates each tool's effectiveness in achieving the system's objectives. The diagram highlights the agents' capabilities on the left, including access to an IoT space and a semantically aware, knowledge graph-based data dictionary. On the right, it depicts key tool calls used to interact with the environment and digital twin resources.

System Overview

Our solution leverages agent-based LLMs to enhance digital twin capabilities for shipping. This architecture incorporates a reasoning and planning agent that decomposes tasks into subtasks and coordinates various specialised agent tools to guide execution. An LLM-driven agentic approach in shipping enables diverse personas—such as fleet owners, ship owners, captains, engineers, and technicians—to seamlessly explore various aspects of the ship's systems and access tailored information, leveraging language as a generic interface.

We developed a suite of tools specifically designed for industrial applications, enhancing agentic decision-making by providing a comprehensive view of real-time ship operations along with contextual insights. These tools integrate structured information from a domain-specific knowledge graph, real-time sensor data, and technical specifications, including potential failure data. By leveraging LangGraph[1], the framework's modular and adaptable design allows it to meet diverse operational requirements and significantly improve decision-making efficiency. We demonstrate the capabilities of this framework on real-world shipping use cases, showcasing its practical applicability and scalability.

Figure 1 presents the pipeline that integrates agentic reasoning with a domain-specific knowledge graph and contextual data to streamline digital twin task selection, refinement, and execution. This framework allows the agent to leverage core digital twin functionalities, including:

- **Dataspace**: A centralised repository encompassing sensor and operational data, providing a comprehensive view of various ship operations.
- **Knowledge Graph**: A structured representation capturing the relationships between the vessel, its components, and associated operational variables. The graph interconnects critical elements such as the main engine, auxiliary engine, and cargo system with key variables like temperature, pressure, and performance metrics, forming a network that reflects the vessel's functional state.
- **Model Builder**: A system designed to handle the selection, construction, and refinement of machine learning models tailored to specific task requirements. It ensures that predictive and prescriptive analytics align with the operational goals of the digital twin.
- **Edge Orchestration Engine**: A component utilising the Linux Foundation's Open Horizon[2] to efficently manage the deployment and placement of containerised applications across a distributed fleet of edge nodes. This enables real-time processing and decision-making close to the data source.

The agents are equipped with a suite of external tools that enhance its ability to interact with digital twin resources and expand its observational capabilities. These tools include

- **Anomaly Detection Tool**: An online unsupervised anomaly detection approach which leverages an efficient formulation of the optimal transport (OT) problem in one dimension to detect anomalies in noisy, seasonal time series data (Langbridge et al. 2024). The tool leverage the counterfactual explanations for detected anomalies generated using the OT mapping, to improve the auditability of the approach when deployed in practical real-world settings (You, Cao, and Nilsson 2024).
- **Work Order Generation Tool**: Evaluates identified anomalies and leverages Failure Mode and Effects Analysis (FMEA) to suggest potential root causes and mitigation activities.
- **Time series forecasting tool**: Leverages a pre-trained time series foundation model (Ekambaram et al. 2024), to forecast the health of a component based on historical data and contributory features. The model can be deployed in zero-shot mode or further refined on historical data.
- **Edge Orchestration Tool**: A tool which automates the creation of JSON policy documents managing the edge deployment of applications on a specific ship based on details such as ship name, geography, company name, or specific requirements.

This demo shows the capabilities of an agentic LLM framework in transforming digital twin applications for the shipping industry across two use cases. Each scenario showcases the system's reasoning ability, self-reflection, and efficient tool execution, using a combination of data aggregation, machine learning, and edge deployment for actionable insights.

One use case focuses on an anomaly detection agent (Timms, Langbridge, and O'Donncha 2024) designed to identify anomalies in a ship's main engine and translate these findings into actionable maintenance work orders. When a user initiates a query, the agent interprets the intent and extracts the necessary asset and operational context to determine which tools to activate. The framework invokes the anomaly dection tool to detect deviations in the identified component, utilising sensor data retrieved through a query to the knowledge graph. Following this, a time series visualisation tool is called to generate relevant plots for further analysis.

After each tool invocation, a reflection agent evaluates the outputs, determining whether additional actions are required. When maintenance tasks are warranted, the agent triggers the work order generation tool, which generates work orders tied to specific standardized failure modes and recommends appropriate mitigation measures. This integrated workflow transforms raw data into actionable insights, enabling technicians to address issues with precision, minimize downtime, and optimize vessel performance.

The second use case highlights the framework's capability to build and deploy an asset health model for the auxiliary engine directly to the edge, enabling real-time monitoring and maintenance close to the data source. This is particularly critical in shipping, where bandwidth is constrained, and operations are mission-critical. The agent autonomously identifies key health indicators using a knowledge graph and historical data, then constructs a zero-shot predictive model using a pretrained time series foundation model to assess wear and tear. Once optimised, the model is containerized with Docker for seamless operation across edge devices. Open Horizon orchestrates the deployment, strategically placing the model on onboard edge devices based on latency and bandwidth constraints. This approach minimises data transfer delays, ensures continuous monitoring, and facilitates predictive maintenance directly at the point of need.

These use cases showcase how agentic LLMs drive operational efficiency and proactive maintenance in shipping fleets by combining advanced reasoning, adaptive tool usage, and iterative task refinement. Through seamless integration of data, analytics, and orchestration, the framework empowers stakeholders to make timely, data-driven decisions, enhancing fleet performance while minimizing costs associated with equipment failures.

[1]https://langchain-ai.github.io/langgraph/
[2]https://lfedge.org/projects/open-horizon/

Acknowledgements

Research described in this report has received funding from the European Union's Horizon Europe research and innovation program under grant agreement no. 101056799.

References

Ekambaram, V.; Jati, A.; Nguyen, N. H.; Dayama, P.; Reddy, C.; Gifford, W. M.; and Kalagnanam, J. 2024. TTMs: Fast Multi-level Tiny Time Mixers for Improved Zero-shot and Few-shot Forecasting of Multivariate Time Series. *arXiv preprint arXiv:2401.03955*.

Katsoulakos, T.; Tsiochantari, G.; O'Donncha, F.; Kaklamanis, E.; Maccari, A.; and Mucharski, M. 2024. Shipping Digital Twin Landscape. In *State-of-the-Art Digital Twin Applications for Shipping Sector Decarbonization*, 1–25. IGI Global.

Langbridge, A.; O'Donncha, F.; Rayfield, J.; and Eck, B. 2024. Optimal Transport for Efficient, Unsupervised Anomaly Detection on Industrial Data. In *2023 IEEE International Conference on Big Data (BigData)*. IEEE.

O'Donncha, F.; Sheehan, J. D.; Touma, M.; Zemouri, S.; and High, R. H. 2024. Towards Intelligent Ship-Edge Computing Enabling Automated Configuration of Ship Models and Adaptive Self-Learning. In *State-of-the-Art Digital Twin Applications for Shipping Sector Decarbonization*, 73–93. IGI Global.

Timms, A.; Langbridge, A.; and O'Donncha, F. 2024. Agentic Anomaly Detection for Shipping. In *NeurIPS 2024 Workshop on Open-World Agents*.

You, L.; Cao, L.; and Nilsson, M. 2024. DISCOUNT: Distributional Counterfactual Explanation With Optimal Transport. *arXiv preprint arXiv:2401.13112*.