Knowledge Management Framework for Autonomous Digital Twins

Miriam Zawadi Muchika[1,2,3]

[1]Orange Innovation, 22 Chemin du Vieux Chêne, 38240 Meylan, France
[2]Mines Saint-Etienne, Univ. Clermont Auvergne,CNRS, UMR 6158 LIMOS, Institut Henri Fayol, Saint- Étienne, France
[3]Mines Paris, PSL University, Center For Management Science (CGS), i3 UMR CNRS, 75006 Paris, France

Abstract

Digital twins (DTs) enable real-time monitoring and control of assets in the physical world. Thus, their application in Open Cyber-Physical Systems is becoming increasingly relevant. However, effective decision loop between digital and physical components, as well as autonomous decision-making directly embedded in DTs, remain open issues. Therefore, in this PhD project, we focus on modeling an *Autonomous Digital Twin* (ADT) that integrates its own decision-making process and acts retroactively on its physical counterpart. Thus, to build this ADT, Knowledge Graphs will be used as the representation of data from the physical world.

Keywords

Knowledge Graphs, Autonomous Digital Twins, Open Cyber-Physical System, Intelligent Agent

1. Introduction

The integration of digital technology with physical systems has resulted in the development of a novel category of systems known as Cyber-Physical System (CPS) [1]. Within a CPS, a collection of computing devices communicate with one another and interact with the physical world via sensors and actuators in a feedback loop in real time [2]. In this work, we focus on building Open CPS where components belonging to multiple stakeholders can join or leave the system unpredictably and can collaborate within the system. A stakeholder is an individual, group, or organization that has an interest or concern in the system. In the context of Smart Cities for example, the stakeholders could be the city's various transport companies (public or private). Enabling these stakeholders to collaborate requires them to exchange and share information. Semantic Web technologies [3] and standardized ontologies in particular [4] make it possible for these stakeholders to understand each other and use the same terminology when interacting. Integrating these technologies into these systems will guarantee semantic interoperability.

To support the integration of the digital and physical dimensions of an asset in the CPS, the concept of Digital twin (DT) [5] - a virtual representation of a physical asset - appears to be a promising approach. In addition, this integration has to be online because, a change of the DT's current state implies a change of state of its physical counterpart and vice versa [6]. Nowadays, DTs have been used as a tool for accessing the physical assets they represent and provide a high-fidelity representation of those assets. As mentioned in [7], DTs are the prerequisite to enable the CPS to be aware of its current state and environment, to adapt and reconfigure itself to changing conditions. In addition, DTs can detect and analyze anomalies in the system [8]. They also enable continuous monitoring and active functional improvement of physical products [9]. However, the current generation of DTs in CPS excels at reflecting the state of the physical system based on sensor data, but fails when it comes to making autonomous, proactive decisions. This limitation hinders the full potential of CPS, especially in dynamic open environments.

The goal of this PhD project is to introduce autonomous decision-making into DT to reduce human intervention by integrating a decision loop between the digital entity (DT) and its physical counterpart.

Proceedings of the Doctoral Consortium at ISWC 2024, co-located with the 23rd International Semantic Web Conference (ISWC 2024)
miriam.zawadimuchika@orange.com (M. Z. Muchika)
0009-0000-9392-0073 (M. Z. Muchika)

For this purpose, the PhD project is organized in two main parts: one dealing with the management and representation of data from the physical world, and the other with the use of this data in autonomous decision-making process of the DT.

Problem Statement. In this paper, we focus on knowledge management challenges. One important challenge is the heterogeneity of the data collected from multiple sources (e.g., stakeholders). Different entities belonging to multiple stakeholders, with different ways of representing information from the physical world, create challenges in terms of data integration. The lack of standardized representation hinders effective collaboration and autonomous decision-making in the system. Since DTs will need to collaborate with other entities, ensuring the interoperability of these heterogeneous systems would allow the DTs to interact in an unambiguous way with other systems. Another challenge is to explicitly capture and represent the changes over time in the physical environment (status, relationships between entities) within an Open CPS. These changes need to be integrated and updated in the internal state of the DT. By addressing these challenges, data representations of the physical world will enable DTs to collaborate to make autonomous and informed decisions, as well as monitor and control their physical counterparts. Therefore, the need for a knowledge management framework becomes evident in addressing these challenges.

Importance. CPSs are critical to the functioning of industries such as transportation, healthcare, manufacturing, and smart cities, where failure can have significant consequences [10]. The ability to effectively manage knowledge (up-to-date and homogeneous data) within Open CPSs is crucial for these systems to operate efficiently, safely, and sustainably. Consequently, entities in Open CPS should share the same data format to enable them to exchange information and collaborate to achieve the system's goal. In areas such as healthcare and transportation, knowledge management in CPS enhances safety and improves quality of life [10].

This paper is structured as follows: in Section 2, we present the research questions addressed. Related works are then reviewed in Section 3. The proposed approach is presented in Section 4. Finally, Section 5 concludes the paper and presents perspectives for future work.

2. Research Questions

The hypothesis underlying this part of the PhD project is that the use of semantic knowledge graphs to represent data in the DT enables real-time modeling of dynamic changes in the physical world and enhances the DT to manifest autonomous behaviors. The aim of the PhD project is to answer the main research question, which is: *"How to develop a knowledge management framework for digital twins that captures and represents real-time data from the physical world and enables both autonomous decision-making and unambiguous collaboration of digital entities within Open CPS?"* To answer the main research question, we will explore the following sub-questions, each aimed at addressing a distinct aspect of the development of the knowledge management framework for autonomous decision-making in Open CPS:

- (RQ.1) How to design a Digital Twin architecture that integrates a knowledge management framework for autonomous decision making?
  This question concerns the design of an architecture that will enable DTs not only to reflect the state of their physical counterparts, but also to become entities capable of making autonomous decisions. The proposed architecture will enable the integration of knowledge management and the ability to manifest intelligent behaviors. This is important, as current approaches to DTs focus on using them to reflect the state of the physical world, to monitor it or to be used for simulation purposes. The ability of DTs to manifest autonomous behaviors is therefore not addressed. The proposed architecture will be generic, so that it can be adapted to a wide range of applications; it will also integrate essential components that must work harmoniously together. Maintaining real-time synchronization between the DT and its physical counterpart is also a key objective of this architecture. This synchronization is essential for two reasons: it ensures that decisions are based on data from the physical world, and it enables the DT to act on its physical counterpart.

- (RQ.2) How to represent data from the physical world in Digital Twins?
  This question tackles the representation of data from the physical world in DTs. To explore this question, a literature review will be done to identify and evaluate different standardization techniques and ontologies used to represent physical world data in DTs. This will enable data from different sources to be structured using a common structure. The ontologies identified will then be used as upper ontologies from which new concepts will be fed to build semantic knowledge graphs that structure the data in DTs, to ensure interoperability. The knowledge graphs constructed will therefore capture not only the current state of the physical counterpart, but also its evolution (including historical states, dynamic changes in relationships between system entities), so that the DTs will be able to understand events occurring in the physical world.
- (RQ.3) How can Digital Twins access to their physical world knowledge to make autonomous decisions?
  This research question explores the mechanisms used by the DTs to query and reason over their semantic knowledge graphs that represent physical world data, to make autonomous decisions. DTs will retrieve current and historical knowledge and analyze it allowing them to understand what is happening in their physical counterparts. Thus, a decision-making process will be developed to enable DTs to evaluate different options and choose the most relevant action.
- (RQ.4) How to ensure unambiguous collaboration between autonomous entities within the Open CPS?
  This question will be addressed with the development of collaboration protocols to enable DTs to share their knowledge from the physical word and work together to achieve the system goals, or to solve problems in a decentralized way[1].

3. Related Work

The integration of data collected from different sources in the DT may be challenging due to inherent heterogeneity. Therefore, some works suggest the use of Knowledge Graphs to structure and extract knowledge from data in DT. Authors in [11] states that with approaches such as Knowledge Graph, information can be linked and later retrieved by different stakeholders at any time in the DT. [11] introduces the concept of Next Generation of DT that uses semantic technologies to connect all kind of information. The vision of Cognitive Digital Twins (CDTs) is proposed by [12, 13] with the aim of enhancing traditional DTs by integrating semantic technology. In [14], authors use ontologies to model the knowledge contained in the CDTs to make the data unambiguous, shareable, and interoperable. The cognitive aspect of CDTs allows to support informed decision-making processes. However, the notion of system openness involving multi-domain knowledge management is not explored. In addition, CDTs are used as decision support tools, for example, in [14] the CDTs are used as a digital shadow: where data is automatically transmitted from the physical layer to the digital layer, without reciprocity [6]. The work in [15] introduces the Web of Digital Twins (WoDT), which is an open, distributed, and dynamic ecosystem of connected DTs functioning as an interoperable service-oriented component for applications running on top, especially smart applications, and multi-agent systems. The semantic modeling of the corresponding DT in [15], is consider as a key aspect of interoperability and openness in WoDT, which is represented by a distributed Knowledge Graph connecting independent KGs (DTs). In [15], the KG within the DT is used as a decision support. Nevertheless, integrating data from various sources and domains while maintaining semantic consistency and lack of ambiguity is a challenging task that can limit the effectiveness of distributed KG. Research on combining KG with the current generation of DTs appears to be a promising approach, enabling DTs to have a homogeneous representation of the data from the physical world, with the aim of using it for autonomous decision-making.

[1]Decentralization problem resolution will enable multiple stakeholders to participate in the decision-making process.

4. Proposed Approach

As mentioned above, DTs are currently used as decision support tools [15, 16]. To include decision making into the Open CPS, we propose an *Autonomous Digital Twin* (ADT) concept (RQ.1). In the context of Open CPS, ADT will include a decision-making process and will be able to collaborate with other autonomous entities (ADT or autonomous intelligent agent) within the system. By making autonomous decisions, these entities can address potential issues and work together towards achieving the system's goals. As an open system, autonomous entities can join and leave the system unpredictably, therefore, autonomous decision-making of this open system will be decentralized. The decentralizing decision-making in the system will enable the participation of various stakeholders in the system and increase its resilience, while minimizing the risk of single-point-of-failure and enhancing adaptability. Decentralized problem-solving also enables the system's autonomous entities to contribute to the finding of solutions that are more acceptable to all.

To model the proposed ADT, a 4-components architecture is described in Figure 1. This architecture contains the *INTERFACE* component that connects the ADT with the entities of the system. The first type of connected entities is its physical asset (PA) with the process *Synchronize* that consists of two phases i.e., *Collect* (collect data from the PA) and *Perform* (makes operation on the PA according to the change made on the ADT's current state). Furthermore, *Interface* links the ADT to systems, users, ADTs, or other entities wishing to interact with it, using the *Expose* process. *Knowledge Graph* and *Ontologies*, in this work, will be used to structure the data collected from the physical world in the *KNOWLEDGE* component of the ADT. The data to be represented will be the current state, which will contain relevant information such as the status of the physical asset, its properties, and its relationships with other entities. Historical data will also be represented in the *Knowledge Graph* and will be used to capture all the past states of the ADT and past relationships with other ADTs, and therefore those of its physical counterpart. The *KNOWLEDGE* component (RQ.2) of the ADT also contains the *Behavior Model*, which will be a numeric representation of the physical and logical behavior of the PA. The physical behavior model refers to the measurable and observable reactions of an object or system in the real world, often based on physical laws and material properties. For example, if the bike travels a certain distance in a given time, the battery level should decrease according to the distance covered and the associated energy consumption. The Logical one is based on a sequence of logical steps determined by operating rules. A logical behavior of a delivery truck might be to pick up goods before delivering to the destination. The *ACTIVITY* component allows the ADT to do operations such as *Monitor* its physical counterpart by comparing the current state of the ADT with the nominal behavior, to identify any inconsistencies; *Analyze* the data from the physical world. The *Simulate* process will enable the ADT to create a model that reproduces the behavior and characteristics of its physical counterpart, enabling different scenarios to be tested and analyzed without impacting the physical entity. The *Predict* process will involve to forecasting future states or behaviors of the physical entity by analyzing historical and real-time data. Finally, the *AGENT* component is the autonomous part of the ADT, enabling it to manifest intelligent behaviors such as decision-making and coordination with other autonomous entities to achieve the system's goals.

To demonstrate the proposed architecture of *ADT*, we will focus on a use case related to Urban Logistics, defined by [17] as the efficient and effective transportation of goods in urban regions. According to [18], 16% to 50% of atmospheric pollutants in cities are generated by transportation, with commercial vehicles (goods transport) accounting for 20% to 30% of the total vehicle-kilometers. Therefore, Last Mile Delivery (LMD) is one of the challenges that dense cities face in terms of carbon footprint. The scope of our research focuses on this final step of the delivery process (LMD), where goods are transported from a transportation hub to their destination. The LMD process that will be used is described as follows: during the day, electric delivery trucks (e-trucks) will be used as mobile depots which are storage spaces where goods are temporarily stored before being delivered to end customers. E-trucks will meet the electric-cargo bikes (e-cargo bikes) to supply them. Finally, e-cargo bikes, will deliver goods to end-users. The meeting points between e-cargo bikes and mobile depots will be the parking spots in the city dedicated to delivery vehicles.

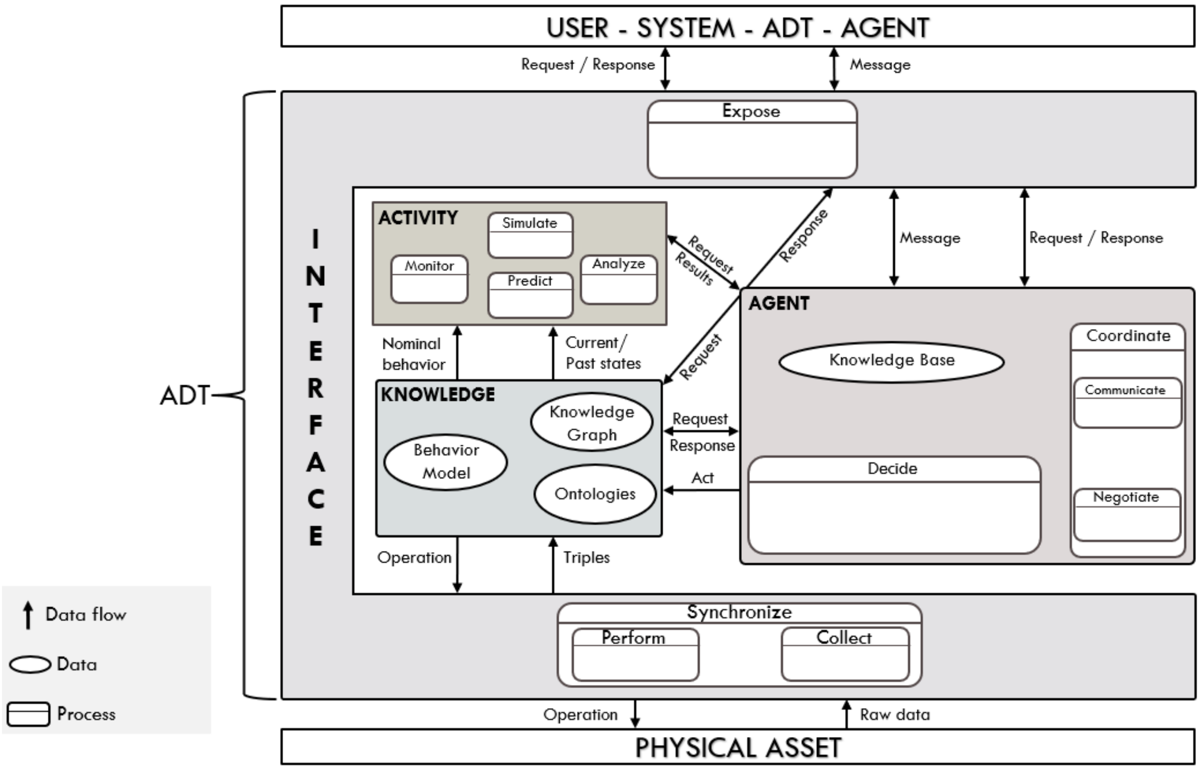

Figure 1: Autonomous Digital Twin Proposed Architecture

To validate our ADT architecture, we will assign an ADT to each entity in the logistics system; for instance, an electric delivery truck will have its ADT modeled, as will e-cargo bikes and other physical entities. One of the standard ontologies identified to model the physical assets (such as the e-trucks) is the SAREF4AUTO ontology[2] [19] (RQ.2). The ADT of an e-truck, for example, will use the current state of its physical counterpart represented in its KG (RQ.2) to monitor it. Using the results of this monitoring, it will be able to make autonomous decisions (RQ.3) that will change its current state. This change in the ADT will be reflected in the physical e-truck (in the case of last-mile delivery, we will consider that the delivery man will receive a notification on his application if the decision is a change of route, for example). Some autonomous decisions to be taken by e-truck and e-cargo bike ADTs may require their collaboration. For instance, when choosing a meeting place for e-trucks and e-cargo bikes, their ADTs will have to exchange messages. This interaction is needed to find an acceptable parking place for them according to their respective constraints. The messages exchanged must be unambiguously and clearly interpreted by the ADTs, who may belong to different stakeholders (RQ.4). Therefore the authors in [20] recommend the use of an agent communication language SPARQL-Act[3], based on Semantic Web Technologies [3] to achieve this semantic interoperability.

5. Conclusion and Future Works

This paper outlines the advancements of the PhD project, which focuses on integrating a decision loop from virtual to physical space by incorporating an autonomous decision-making process into the DTs in Open CPS. The first stage of this PhD project consists of proposing a knowledge management framework for the *Autonomous Digital Twin* (ADT). This framework should use Semantic Web Technologies standards to model heterogeneous data from the physical world, thereby making the ADTs interoperable

[2]This ontology is one of a suite of the standardized SAREF ontologies that form a shared model of consensus intended to enable semantic interoperability between solutions from different providers and among various activity sectors in the IoT.
[3]https://gitlab.emse.fr/sparql-act/web-agents

within the system. One of the main limitations of the proposed architecture is that not all components can be fully implemented during the thesis period. Furthermore, since we are dealing with Open CPS in which components can join and leave the system unpredictably, consequently, the ADT must be able to discover when new ADTs enter the system. In addition, the system must be resilient, retaining its functionality even when an ADT leaves the system. Future work will focus on modeling the decision-making process that incorporates knowledge of the physical world. It will also involve modeling each component of the ADT architecture. Subsequently, the architecture will be implemented and validated through a realistic use case.

Acknowledgments

The author would like to thank her supervisors, Dr. Pauline Folz, Dr. Fano Ramparany, Pr. Flavien Balbo and Pr. Shenle Pan for their continuous guidance and support for this research proposal.

References

[1] R. Baheti, H. Gill, Cyber-physical systems, In:Samad T and Annaswamy AM (eds) The impact of control tech- nology, vol.12 (2011) New York: IEEE, 161–166.
[2] R. Alur, Principles of Cyber-Physical Systems, MIT Press, Cambridge, MA, USA, 2015.
[3] B. Matthews, Semantic web technologies, E-learning and Digital Media 6 (2005) 8.
[4] D. Fensel, D. Fensel, Ontologies, Springer Berlin Heidelberg, 2001.
[5] G. Lukowski, A. Rauch, T. Rosendahl, The virtual representation of the world is emerging, Management for Professionals (2018). doi:10.1007/978-3-319-77724-5_14.
[6] A. Fuller, Z. Fan, C. Day, C. Barlow, Digital twin: Enabling technologies, challenges and open research, IEEE access 8 (2020) 108952–108971.
[7] K. Josifovska, E. Yigitbas, G. Engels, Reference framework for digital twins within cyber-physical systems, in: 2019 IEEE/ACM 5th International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS), 2019, pp. 25–31. doi:10.1109/SEsCPS.2019.00012.
[8] P. Pileggi, J. Verriet, J. Broekhuijsen, C. V. Leeuwen, W. Wijbrandi, M. Konsman, A digital twin for cyber-physical energy systems, 2019 7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES) (2019) 1–6. doi:10.1109/MSCPES.2019.8738792.
[9] C. Stary, M. Elstermann, A. Fleischmann, W. Schmidt, Behavior-centered digital-twin design for dynamic cyber-physical system development, Complex Syst. Informatics Model. Q. 30 (2022) 31–52. doi:10.7250/csimq.2022-30.02.
[10] V. Gunes, S. Peter, T. Givargis, F. Vahid, A survey on concepts, applications, and challenges in cyber-physical systems, KSII Transactions on Internet and Information Systems (TIIS) 8 (2014) 4242–4268.
[11] S. Boschert, C. Heinrich, R. Rosen, Next generation digital twin, in: Proc. tmce, volume 2018, Las Palmas de Gran Canaria, Spain, 2018, pp. 7–11.
[12] J. Lu, X. Zheng, A. Gharaei, K. Kalaboukas, D. Kiritsis, Cognitive twins for supporting decision-makings of internet of things systems, in: Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing: AMP 2020, Springer, 2020, pp. 105–115.
[13] X. Zheng, J. Lu, D. Kiritsis, The emergence of cognitive digital twin: vision, challenges and opportunities, International Journal of Production Research 60 (2022) 7610–7632.
[14] Y. Liu, S. Pan, P. Folz, F. Ramparany, S. Bolle, E. Ballot, T. Coupaye, Cognitive digital twins for freight parking management in last mile delivery under smart cities paradigm, Computers in Industry 153 (2023) 104022.
[15] A. Ricci, A. Croatti, S. Mariani, S. Montagna, M. Picone, Web of digital twins, ACM Transactions on Internet Technology 22 (2022) 1–30.
[16] A. Bujari, A. Calvio, L. Foschini, A. Sabbioni, A. Corradi, A digital twin decision support system for the urban facility management process, Sensors 21 (2021) 8460.

[17] M. Savelsbergh, T. Van Woensel, 50th anniversary invited article—city logistics: Challenges and opportunities, Transportation science 50 (2016) 579–590.
[18] L. Dablanc, Goods transport in large european cities: Difficult to organize, difficult to modernize, Transportation Research Part A: Policy and Practice 41 (2007) 280–285.
[19] SAREF, Saref extension for automotive, 2020. Https://saref.etsi.org/saref4auto/v1.1.1/, (accessed: 27.04.2024).
[20] M. Z. Muchika, O. Kem, S. B. Abbes, L. Temal, R. Hantach, Achieving interoperability in energy systems through multi-agent systems and semantic web, in: 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE, 2023, pp. 441–448.