PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

PhD Student: Alessandra Somma

Cycle: XXXVII

Training and Research Activities Report

Academic year: 2022-23 - PhD Year: Second

Tutor: Prof. Alessandra De Benedictis

Date: October 23rd, 2023

Information:

- **PhD student: Alessandra Somma** **PhD Cycle: XXXVII**
- **DR number: DR995852**
- **Date of birth: 28/12/1996**
- **Master Science degree: Computer Engineering at University of Naples "Federico II"**
- **Scholarship type: UNINA**
- **Tutor: Prof. Alessandra De Benedictis**
- **Co-tutor: none**

Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
"Publish Open Access Articles with IEEE under the CARE CRUI Agreement"	Seminar	1/2	0.1	09.11.22	Eszter Lukács (IEEE Client Services Manager)	N
Complex network systems: introduction and open challenges	Seminar	2	0.4	17.11.22	SSM	Y
Cybercrime and Information Warfare: National and International Actors	Seminar	2	0.4	18.11.22	Proff. Romano, Natella	Y
Privacy and Data Protection	Seminar	2	0.4	22.11.22	Proff. Romano, Natella	Y
IoT Data Analysis	Course	-	4	9-27.01.23	Dr. Raffaele Della Corte	Y
Statistical Data Analysis for Science and Engineering Research	Course	-	-	6-16.02.23	Prof. Roberto Pietrantuon o	N
Sensors Webinar | 5G-Enabled IoT and Digital	Seminar	2	0.4	24.05.23	Sensors	Y

Twins: Cybersecurity and Resilience						
IEEE 2023 Smart World Congress	Seminar	19	1.9	28-31.08.23	IEEE	N
Semantic Artifacts and Multimedia Knowledge Graphs for bio-data integration	Course	-	2	12.09.23-3.10.23	Dr. Cristiano Russo	Y

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	0	1.3	5	0	6.3
Bimonth 2	4	0	6	0	10
Bimonth 3	0	0	6	0.16	6.16
Bimonth 4	0	0.4	6	0	6.4
Bimonth 5	0	1.9	6	0	7.9
Bimonth 6	2	0	7	1	10
Total (2nd year)	**6**	**3.6**	**53**	**1.16**	**63.76**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

Total (1st year)	**28.8**	**5.4**	**27**	**0.6**	**62**
Total (2nd year)	**6**	**3.6**	**53**	**1.16**	**63.76**
Total (1-2 years)	**34.8**	**9**	**80**	**1.76**	**125,76**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

Research activity:

During my second year, my research has been focused mainly on the architectural concerns of the Digital Twins (DTs), that are delaying the widespread implementation and adoption of this extremely beneficial technology. In fact, the lack of understanding of the Digital Twin concept, its differences from the mere simulation and the Cyber-Physical System (CPS) notion brought to a multitude of definitions strongly dependent on the DT application context. Indeed, as pointed out in a recent literature review [1], the absence of a Digital Twin definition may partially reflect the great interests of the communities defining the term. However, rather than creating yet another DT definition, I am relying on a recent generic one: "a live digital coupling of the state of a physical asset or process to a virtual representation with a functional output" [2] because of the absence of assumptions such as the purpose/use of the Digital Twin, the nature of the physical entity that is digitally replicated or the sector in which it will be used.

Starting from this definition and from the analysis of literature, five main DT requirements have been identified:

- *RQ1 – physical asset/process virtual representation*: a Digital Twin will provide a **comprehensive virtual representation** across the entire lifecycle of an entity with economic, social or commercial value [3].
- *RQ2 – state twinning*: the condition the physical asset/process is in **at a specific time** must be digitally replicated in its virtual counterpart.
- *RQ3 – liveness*: the state information is available in a timeframe that is close enough to the underlying event.
- *RQ4 – digital coupling*: the transmission mechanism between data sources and data consumption methods uses a digital carrier medium.
- *RQ5 – functional output:* information/alert/command generated by the Digital Twin must be transmitted to a system or human observer that is actionable to deliver value.

Even if my six-layered architectural proposal (done in the first year) with Physical, Data, Digital Twin, Service, Communication and Security layer seemed to fulfill these requirements and maps the ISO 23247 standard that provides a framework for DTs only in the manufacturing domain, what is currently missing is a major focus on ***data***. In fact, in a DT context, data represent the so-called Digital Twin fuel: the construction of the virtual representation is aided by real-world historical and operational data, which then feed the digital models to obtain a faithful replica; furthermore, the data produced by the DT can generate new information that can be shared with systems or humans and executed in the real world. This means that in a Digital Twin architecture, data represent a pillar and thus solutions for their management, storage, analysis need to be considered with respect to their heterogeneity (e.g., Smart City context).

Since the Digital Twin leverages on data to build up simulation and other services that will affect the physical system, I faced the introduction of data in the Digital Twin architecture from two different perspectives:

1. Data are obtained from heterogeneous and often untrustworthy data sources, thus I investigated the possibility to use Distributed Ledger Technologies (DLTs) to secure data used by the DT when both in transit and at rest and ensure the trustworthiness of sources. In particular, I conducted a Systematic Literature Review (SLR) D and in contrast to the existing ones, I answered to 7 Research Questions (RQs) whose aim is to understands the benefits and challenges of including DLTs in a DTs environment, but also what are the available implementations, their functional and non-functional requirements, eventually architectural models and evaluation metrics. Moreover, I also investigated the usage of different kinds of DLTs, such as blockchain and tangle.
2. A Digital Twin system should provide data integration from multiple data sources supporting data storage, data transformation and analytics, data access and sharing in compliance with data protection requirements. According to recent studies (*e.g.*, [3-4]), to realize the original idea of the Digital Twin and thus to fulfill the *RQ1*, the concept of DT has to be integrated with the *Data Space* one. A Data Space is an umbrella term for all the approaches that support data sharing in collaborative environments with respect to the European Data Strategy. This means that data spaces should enable the access to and privacy-compliant data usage through the creation of suitable ecosystems [3-5].

Considering the state-of-the-art architectural models of International Data Spaces (IDS), I started to extend the first architectural proposal. More in detail, the original Data Layer has been splitted in four different layers:

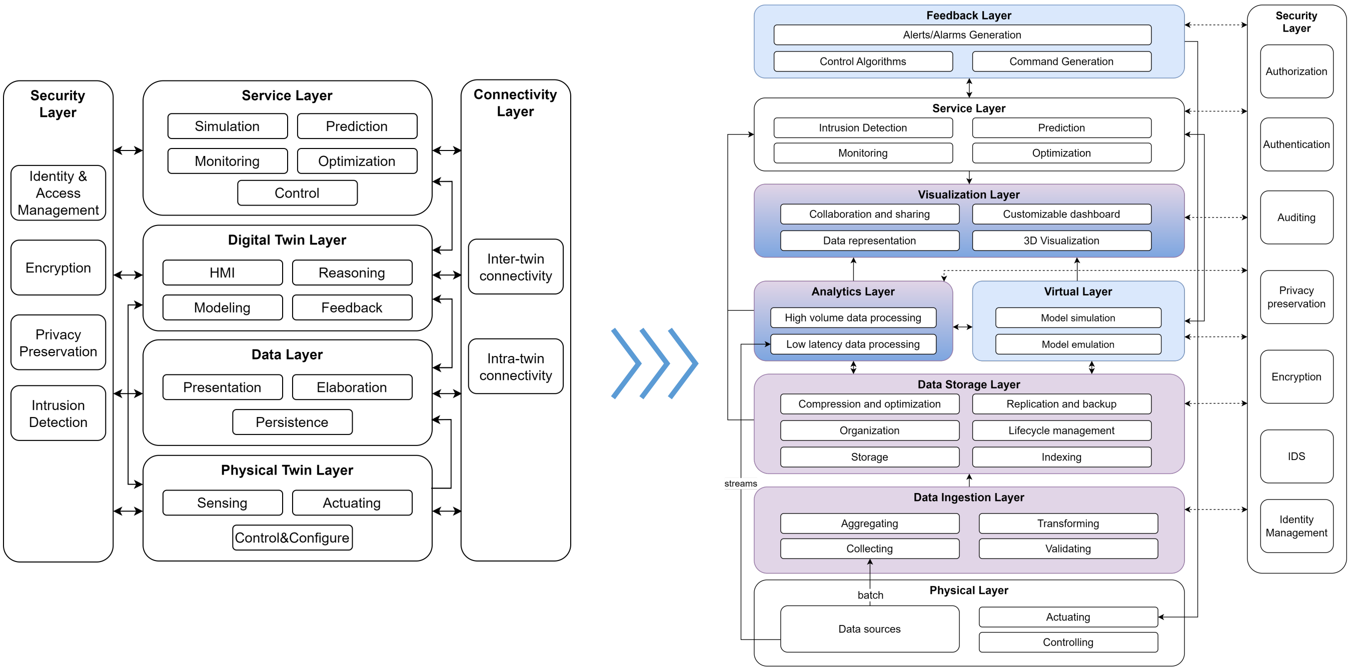

Figure 1. The extension of Digital Twin architectural model.

1. the *Data Ingestion Layer* responsible for collection and processing raw data from various sources that usually come in different formats and structures;
2. the *Data Storage Layer* that aims at securely store the ingested data;
3. the *Analytics Layer* in charge of handling and processing data streams, i.e., DT operational data, and static and historical ones, such as information provided by domain-knowledge experts;
4. the *Visualization Layer* extends the general HMI functionality of the Digital Twin, so it is useful for data representation and 3D visualization.

Finally, leveraging on results obtained from one or more services based on DT models execution, I introduced the *Feedback Layer* that eventually executes control algorithms and generates alerts/alarms and or command according to DT purposes.

It is possible to note that the splitting of the data flow in real-time streams and batches allows to facilitate the integration of DLTs and in particular blockchains. In fact, even if from the literature review that I conducted I found out that there is still a trend to store heave information on-chain, blockchains can be easily introduced storing real-time data in distributed file systems (*e.g.*, IPFS) and their hash on-chain coping with scalability and performance issues.

From a technological point of view, there are pre-built solutions such as Siemens Mindsphere or 3D Experience of Dassault Systèmes. During this year, I tested different open-source tools to demonstrate the validity and applicability of the proposed architecture:

- *Eclipse Ditto* is an open-source framework that can be used to create and manage DTs. It provides a set of features and services for physical devices modeling, managing device data and interaction with IoT devices. However, it is resource intensive, especially in terms of memory and processing power.
- *Cloud2Edge* that integrates two other Eclipse frameworks, i.e., Eclipse Hono and Eclipse Ditto.
- *MiniCPS,* simulation engine for security purposes used during the period abroad.

During the last period, I am investigating the possibility to use FIWARE technology to set-up a Digital Twin platform integrating the Data Space concept. In fact, FIWARE is an open-source initiative defining a universal set of standards for *context data management* which facilitate the development of smart solutions. As shown in Fig. 2, the FIWARE architecture is based on *context information captured from the real world* and managed

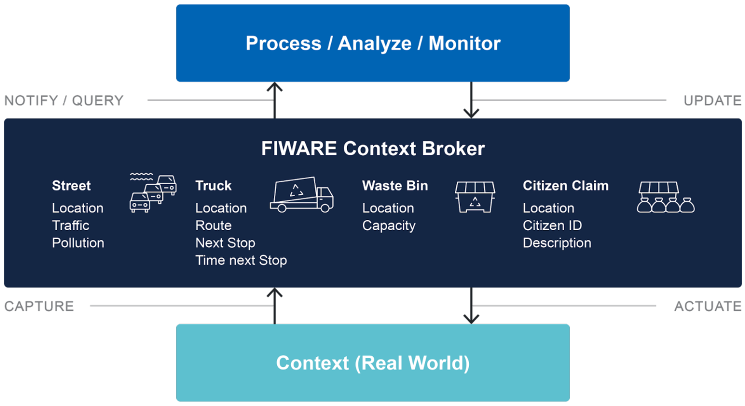

Figure 2. FIWARE-based system architecture.

in the FIWARE layer called the Context Broker. The data managed by the Context Broker are used for processing, analyzing and monitoring of data. Finally, it is used to actuate what is happening in the real world. FIWARE Context Broker Techology has been chosen as new CEF (Connecting Europe Facility) Building Block by all European member states. Thanks to Smart Data Models initiative, FIWARE can be a solution to develop a platform for DT with respect to the proposed architecture [6].

3.1 References:

[1] Boyes, H., Watson, T., "Digital twins: An analysis framework and open issues", *Computers in Industry*, vol. 143, 2022.

[2] Catapult, H. V. "Untangling the requirements of a digital twin." *Univ. Sheff. Adv. Manuf. Res. Cent. (AMRC)*, 2021.

[3] Usländer, T.; Baumann, M.; Boschert, S.; Rosen, R.; Sauer, O.; Stojanovic, L.; Wehrstedt, J.C. Symbiotic Evolution of Digital Twin Systems and Dataspaces. *Automation*, 2022.

[4] Volz F, Sutschet G, Stojanovic L, Usländer T. On the Role of Digital Twins in Data Spaces. *Sensors*, 2023.

[5] Nagel, L.; Lycklama, D. "Design Principles for Data Spaces—Position Paper". Available online: https://zenodo.org/record/5105744

[6] J. Conde, A. Munoz-Arcentales, Á. Alonso, S. López-Pernas and J. Salvachúa, "Modeling Digital Twin Data and Architecture: A Building Guide With FIWARE as Enabling Technology," in *IEEE Internet Computing*, May-June 2022.

Research products:

J1. De Benedictis, A., Flammini, F., Mazzocca, N., Somma, A., Vitale, F., "A Digital Twin Architecture for Anomaly Detection in the Industrial Internet of Things", *IEEE Transactions on Industrial Informatics* (TII). Status: published.

J2. De Donato, L., Dirnfeld, R., Somma, A., De Benedictis, A., Flammini, F., Marrone, S., Saman Azari, M., Vittorini, V., "Towards AI-Assisted Digital Twins for Smart Railways: Preliminary Guideline and Reference Architecture", *Journal of Reliable Intelligent Environments*. Status: published.

J3. De Donato, L., Dirnfeld, R., Somma, A., Flammini, F., Marrone, S., Saman Azari, M., Vittorini, V., "Integrating AI and DTs: Challenges and Opportunities in Railway Maintenance Application and Beyond", *Journal of Simulation: Transactions of the Society for Modeling and Simulation International*. Status: accepted.

J4. Somma, A., De Benedictis, A., Esposito, C., Mazzocca, N. "The convergence of Digital Twins and Distributed Ledger Technologies: A systematic literature review and an architectural proposal", *Journal of Computer Network Applications*. Status: under 1st stage of review.

J5. Somma, A., De Benedictis, A., Urciuolo, F., Mazzocca, N., Netti, P., "Digital Twins applied to bioengineering: tissue-on-chips", to be submitted.

C1. Somma, A., Casola, V., Cavalli, A. R., De Benedictis, A., Mallouli, W., Valdés, V. E., "A Cyber Digital Twin Framework to Support Cyber-Physical Systems Security", in *IEEE 2023 Smart World Congress*.

C2. Somma, A., De Benedictis A., Longo, A., Martella, A., Martella, C., "Digital Twin Space: integration of Digital Twin and Data Space concepts", *2023 IEEE International Conference on Big Data*, to be submitted.

Conferences and seminars attended

The 2023 IEEE International Conference on Digital Twin held in the IEEE Smart World Congress, August 28th – 31st 2023, University of Portsmouth where I presented my work "A Cyber Digital Twin Framework to Support Cyber-Physical Systems Security".

Training campus on "FIWARE" technologies, University of Naples Parthenope, June 5th – 9th 2023.

First plenary "DYNABIC" meeting in Barcelona held by BEWARE, June 18th-20th 2023.

Periods abroad and/or in international research institutions

From February 1st to July 31st 2023 (**six months**), I spent a research period abroad at Montimage EURL under the supervision of Research Engineer Wissam Mallouli. I took part to DYNABIC European Project whose goal is the increasing of resilience and business continuity capabilities of Critical Infrastructures (CIs) using the concept of Digital Twins. In fact, since IoT, Cloud, AI are more and more adopted as integral part of systems, the door to new cyber-physical attack vectors been opened. The main research objectives can be classified as follows:

- *Research objective 1*: enable operators of CIs to predict, quantitatively, assess, and mitigate in real-time business continuity risks and cascading effects in interconnected CIs.
- *Research objective 2:* methods and tools for disaster preparedness and the prevention of business continuity risks in cross-organization and cross-domain incidents and attacks.
- *Research Objective 3:* dynamic autonomous adaptation of critical infrastructures to meet resilience goals with personalized assistance in human tasks.

I investigated the usage of the so-called *Cyber Digital Twins* (CDTs), that I started to study during my first year of PhD. More in detail, at the beginning, I conducted a state-of-the-art analysis to understand what a CDT is and when it can be used and I found out that first, a CDT is a cybersecurity-oriented virtual replica of an asset that can be used for cybersecurity purposes, *e.g.*, to simulate potential cyber-attacks and test security measures. In fact, according to the literature analysis, CDT security objectives are [part of my work aforementioned – C1]: i) security testing; ii) cybersecurity training; iii) attack/intrusion/anomaly detection; iv) security controls selection; v) Cyber-Threat Intelligence (CTI) generation; vi) diagnostics.

From an experimental point-of-view, starting from my six-layered architecture, I extended and customized it to obtain a comprehensive CDT architectural model for CPS systems and support multiple cybersecurity objectives. Finally, I proved the effectiveness of my proposal through a proof-of-concept (PoC), leveraging on the MiniCPS simulator. This architectural model (with some enhancement) will be part of Deliverable D2.1 "DYNABIC Framework Use cases, Requirements and Architecture" (SEN) of DYNABIC project.

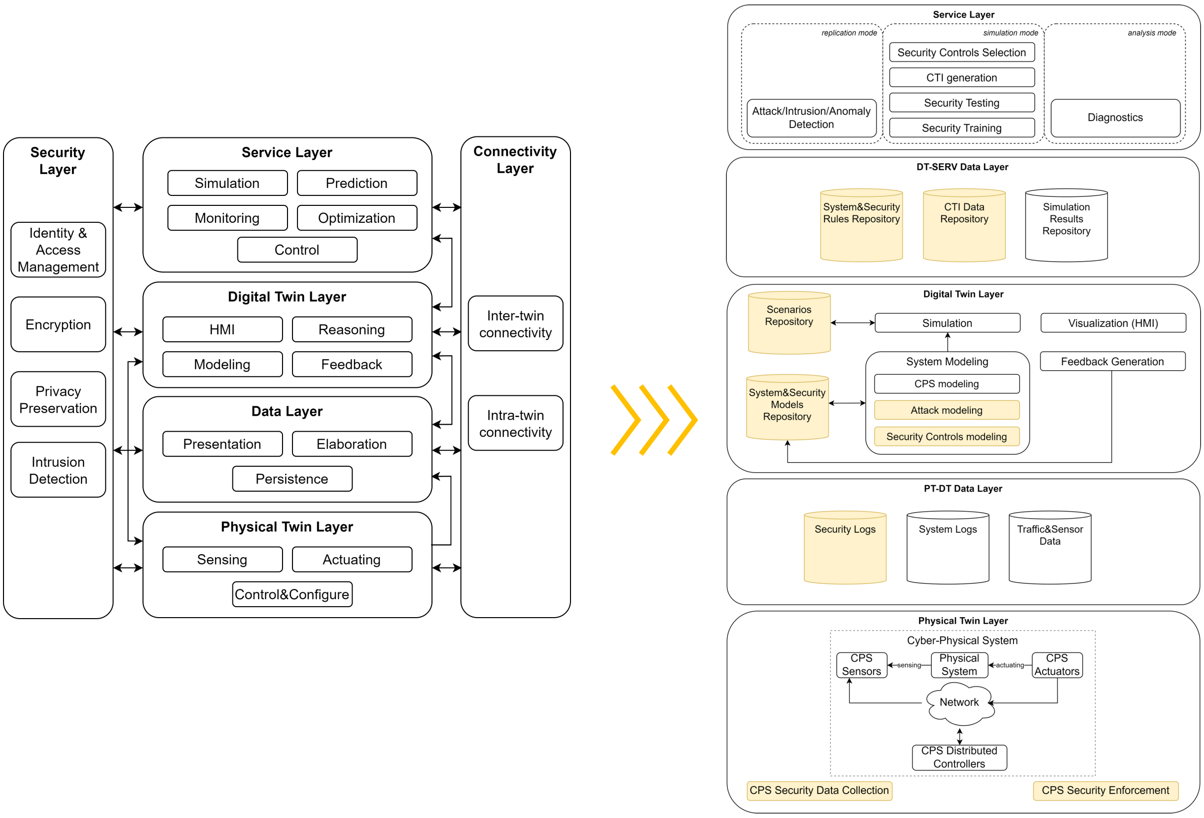

Figure 3. Cyber Digital Twin architectural model.

7. Tutorship

Co-supervisor of the thesis of students: Grazia Napolitano, Antonio Landolfi, Prisco Santillo, Gennaro Napolano, Federica Esposito.

4h Seminar for Risk Assessment MSc course.

Support for Computer System Design (CSD), Architetture dei Sistemi Digitali (ASDi) and Calcolatori Elettronici I exams.

8. Plan for year three

For the next year, my plan is to bring together the main aspects and results of my research activities. In fact, my idea is to complete the definition of a general architectural model for Digital Twin implementation, extending the Digital Twin requirements, and define a DTS – Digital Twin Space as the integration of DTs and Data Spaces. Moreover, I plan to map the high-level functional architecture onto a technological one and develop a comprehensive platform for DT implementation, leveraging on complex technologies such as FIWARE. Finally, to validate the proposal, I will use the architecture in at least one case study, i.e., Urban Digital Twins (UDTs) done in collaboration with Prof. Longo from University of Salento in the context of PNRR project (MUR-PNRR High-Performance Computing, Big Data e Quantum Computing Research Centre, CN_00000013).