MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF TUNIS EL MANAR

FACULTY OF SCIENCES OF TUNIS

PhD THESIS

to obtain the title of

PhD IN COMPUTER SCIENCE

Defended by

Karim Houidi

En titled

Explainable Federated Deep Learning Models on Digital Twins for Smart Agriculture

Prepared at

LIPAH research Lab AND LIASD research Lab

In front of the jury composed of:

Pr. Hella KAFFEL BEN AYED	FST,University of Tunis El Manar	President
Pr. Akram HAKIRI	ISSATM,University of Carthage	PhD Supervisor
Pr.Nedra Mellouli-Nauwynck	LIASD,University Paris8	PhD Supervisor

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Faced with the challenges of population growth and global warming, many countries of the two rivers of the Mediterranean Sea are facing a severe water shortage in the 21st century. The efficient management of drinking and irrigation water in arid and semi-arid regions today poses a major problem at the top of global development priorities. Worsened by the war in Ukraine, food security in southern Mediterranean countries will depend on their ability to better understand climate change and to adapt water consumption in all its uses and in particular for smart irrigation management. Agriculture, the main sector of water consumption, uses most of the available water reserves, representing nearly 40.03% of available resources. Rain-fed agriculture is constrained by the arid climate, the irregularity of rainfall and the degradation of soil quality due to erosion.

The origins of agriculture can be traced back to the early history of human civilisation, remaining one of the world's most important industries even today. The announcement of 4th agricultural revolution or "Agriculture 4.0" in academia and industry has brought the promise of digitization, technological advancement and increased efficiency. Digital twins (DT) are being increasingly adopted by several disciplines, including the manufacturing (Kritzinger et al., 2018), automotive (Caputo et al., 2019) and energy (Sivalingam et al., 2018) sectors, for addressing multidisciplinary problems.

The following of this paper is as it follows. Chapter 2 presents an overview of DTs in the scientific literature, whereas Chapter 3 introduces our architecture for smart factories. Chapter 4 introduces the available frameworks for DTs, by presenting their technical features. Chapter 5 discusses how to realize our architecture through available frameworks.

Chapter 2

Literature review

This systematic literature review specifically focuses on works related to business and engineering aspects of DT technology. According to (Cook et al. 1997), a systematic review differs from traditional general review in that a duplication of the distinct and objective process is possible. As DT technologies are progressively developed for a wider range of industries to tackle extensive corporate functions such as strategic planning, it is essential that the technical, engineering PLM and business aspects of DT technology be reviewed to investigate the collective insights on theoretical analysis of existing studies.

2.1 Methodology in research selection

A systematic literature search is based on the PRISMA method, covering most of the peer-reviewed interdisciplinary research papers, where a broad sum of studies on DT and other related literature can be identified using the systematic review methodology. Articles collected were further refined through a the 4 steps of which are detailed in the flow chart, as depicted in figure2.1.

2.2 Digital Twin overview

Digital Twins are systems designed to work with rich data provided by both the insight of the models and implicit knowledge of objects, products of their behavior, only requiring the availability of raw data about their physical counterparts. To achieve this, Digital Twins must accurately simulate, analyze, and predict real-world events and situations, via the collection of

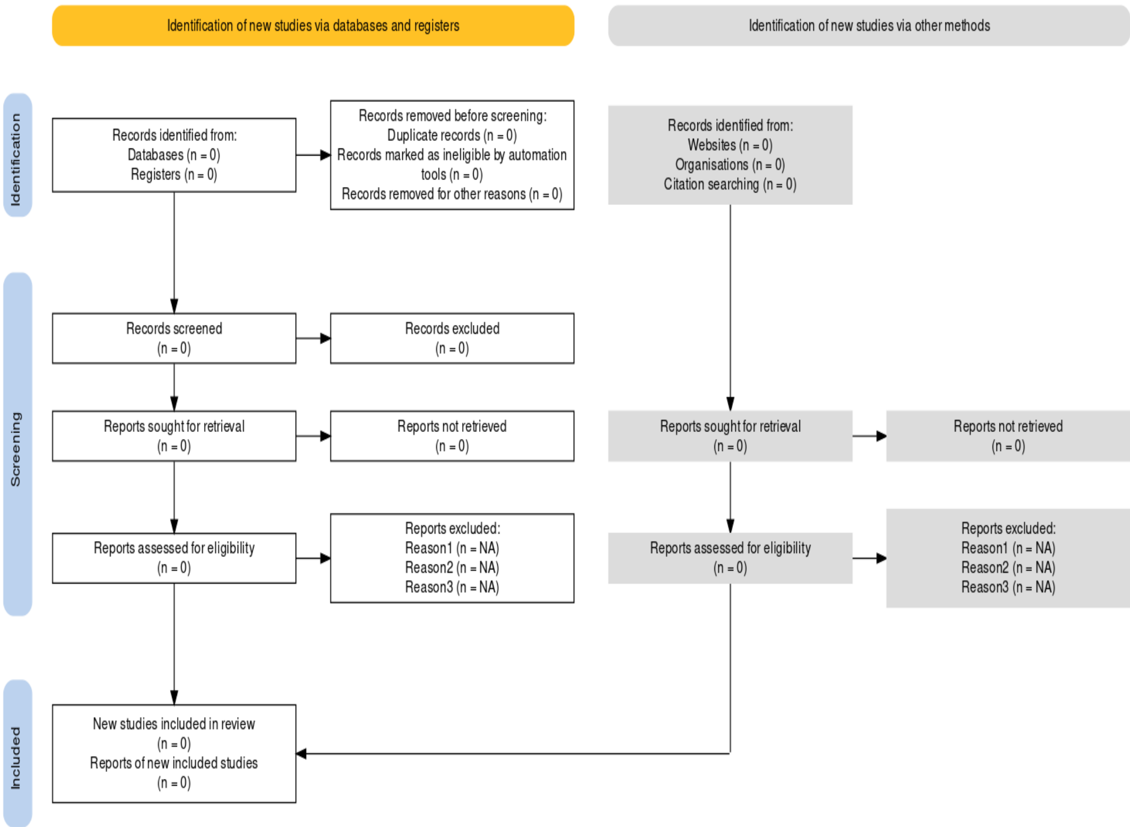


Figure 2.1: Systematic review flow diagram

real-time physical data, mirroring it into its virtual space. With the rapid development of virtual technology and data acquisition technology, digital twin (DT) technology was proposed and gradually become one of the key research directions of intelligent manufacturing.This section gives a brief overview of the definition of DT and presents the definition from the perspective of its characteristics and functions.

2.2.1 Definition of Digital Twin

Based on our literature survey we present the following definition of digital twin: A digital twin is defined as a virtual representation of a physical asset enabled through data and simulators for real-time prediction, monitoring, control and optimization of the asset for improved decision making throughout the life cycle of the asset [6].

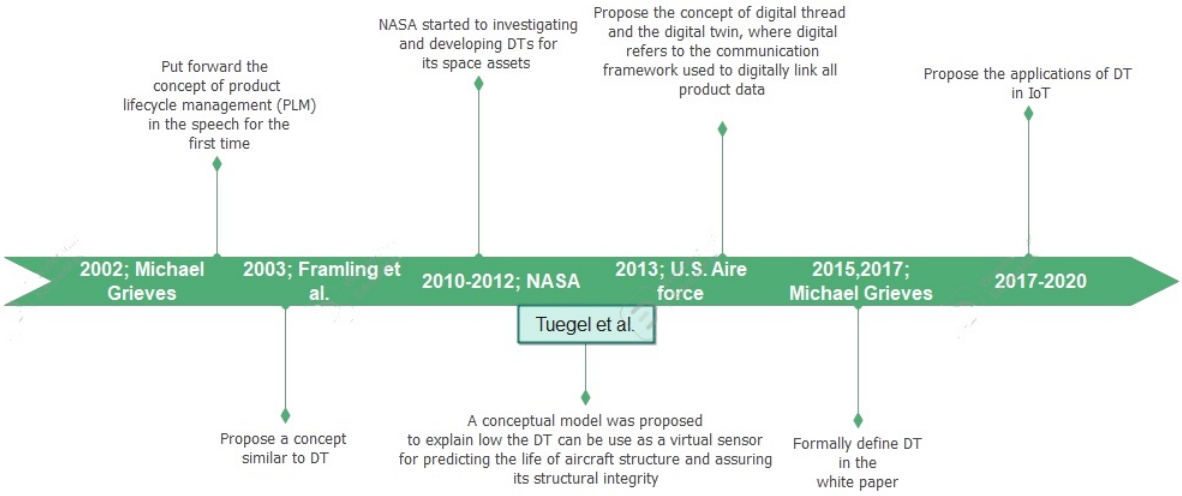


Figure 2.2: Timeline of DT

References	Definitions	Key word
[1]	A digital twin is a virtual representation of a physical product or process, used to understand and predict the physical counterpart's performance characteristics. Digital twins are used throughout the product lifecycle to simulate, predict, and optimize the product and production system before investing in physical prototypes and assets.	Digital Twin,Virtual model
[10]	The DT is a set of virtual information that fully describes a potential or actual physical production from the micro atomic level to the macro geometrical level. At its optimum, any information that could be inspected from a physical manufactured product can be obtained from its DT	Digital Twin, Cyber Physical Systems (CPS)
[7]	Faster optimization algorithms, increased computer power and amount of available data, can leverage the area of simulation toward real-time control and optimization of products and production systems – a concept often referred to as a Digital Twin	Simulation, Computerized
[8]	digital twin is an integrated multi-physics, multiscale, probabilistic simulation of a complex product and uses the best available physical models, sensor updates, etc., to mirror the life of its corresponding twin	Digital twin, Product lifecycle Design, Manufacturing, Service, Big data, Cyber and physical convergence
[5]	An integrated multiphysics, multiscale, probabilistic simulation of an as-built system, enabled by digital thread, that uses the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its corresponding physical twin	Integreted System
[2]	the Digital Twin itself refers to a comprehen-	Digital Twin, opera-

Table 2.1: Definition of DT

Chapter 3

Digital twin types

You can use digital twin models to represent everything from individual components to entire systems. While every type of digital twin fundamentally does the same thing – virtually modeling a real-world object or system – their purposes and scope greatly vary from one to another. The four primary types of digital twins are:

3.1 Component twins

Component twins are digital representations of an individual part of a system or product, such as a gear or screw. Rather than simply modeling all the individual parts of a product, though, component twins are typically used to model integral parts, such as those under particular stress or heat. By digitally modeling these parts and subjecting them to dynamic simulations, designers and engineers can see how the parts can be improved to ensure their integrity in likely scenarios.

3.2 Asset twins

Asset twins, also called product twins, are virtual representations of a physical product rather than its individual parts. While asset twins can technically be composed of numerous component twins, their purpose is to understand how their various parts operate together within a single real-world product. For example, a wind turbine might have an associated asset twin used to monitor its performance and identify possible parts failure due to common wear and tear.

3.3 System twins

System twins, also called unit twins, are virtual representations of systems of products working together. While asset twins model real-world products comprised of many parts, system twins model these individual products as components of a larger system. Understanding how assets interact with one another offers the opportunity to improve how they relate to one another, increasing productivity and efficiency as a result.

3.4 Process twins

Process twins are digital representations of systems working together. For example, while a system twin might model a manufacturing line, a process twin could model the entire factory all the way down to the employees operating the machines on the factory floor.

3.5 Integration Levels

The following integration levels are in ascending order, meaning digital models are the least integrated ones and digital twins are the most integrated, as proposed in [3].

- **Digital model**: In its basic concept, the digital model will not integrate any automatic information flow from the physical world to the virtual world. This means that the virtual and physical world are not automatically connected, so any change must be reflected through manual modification.

- **Digital shadow**: The digital shadow will integrate unidirectional automatic information flow from the physical world to the virtual world [3]. This is best represented by a system where sensors measure information from the physical model and transfer signals to the virtual model. Regardless of whether information flows in a polling or interrupt method, as long as it is automatic, the integration level can be determined as a digital shadow.

- **Digital twin**: A fully integrated twin where the virtual and physical world interact in a bidirectional fashion. This means that information flows automatically to and from each world. In this case, information flowing from the virtual world will be useful to perform changes in the physical model or to instruct actuators to perform an operation.

Conversely, data from the physical twin may influence the virtual twin automatically in such a way that the virtual twin accurately represents the current state and the evolution of its physical counterpart.

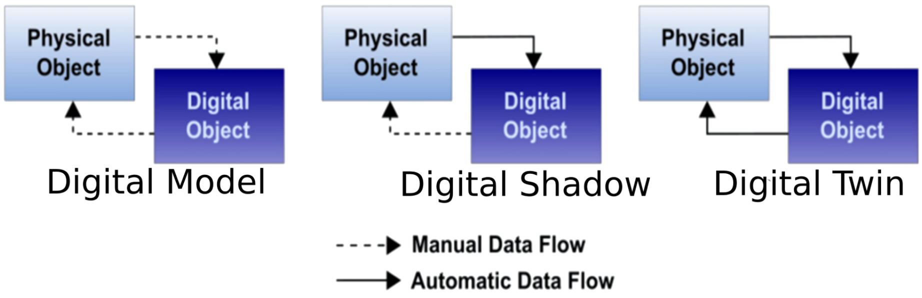
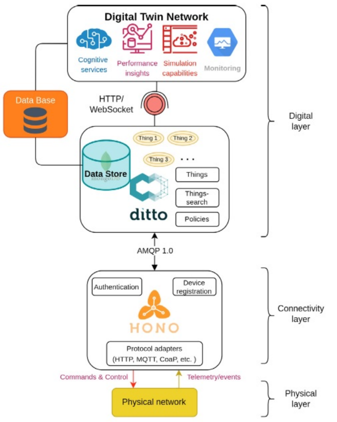


Figure 3.1: Digital twin types

Chapter 4

Digital Twin Architectures

A high-level architecture for the digital twin of wireless systems is introduced in [4] in three layers: Physical Interaction Layer, Twin Layer and Services Layer. Two aspects are considered to discuss different details regarding digital twins for wireless systems. Wireless for twins considers the efficient use of wireless resources to enable efficient signaling of twins over a wireless link. On the other hand, twins for wireless discusses the role of digital twins in the implementation of wireless systems.
[9] proposes a four-layer architecture for DTN including a physical network layer, a data lake layer, a DT layer, and a network application layer. The data lake layer collects, stores and preprocesses the collected data to provide knowledge and relationship extraction to the DT layer for building the DTN model. The DT layer includes: Physical entity modeling that complements the modeling of individual network components, Requirements modeling that is used to develop a variety of scenario models such as network resource prediction, anomaly detection, automatic operation through AI algorithms, The twin management and control center that manages the DT layer to satisfy the needed requirements.
In this section we describe the DTN architecture for the IIoT as an effort to enable intelligent management and more innovation for such networks. As depicted in 4.1, the physical IoT network interacts with the connectivity layer where Eclipse Hono (Eclipse (2017a)) is used. This latter is designed for simplified IoT device connectivity by eliminating the protocol silos of the different devices. In fact, Hono contains for each supported protocol (e.g., HTTP, MQTT or CoaP) a microservice called Protocol Adapter which maps the connection protocol of the device to Hono's APIs (Application Programming Interface). A device registration API is included to make Hono aware of devices that can or will connect to the service. In addition, authentication API is used to verify the identity of the devices willing to connect to Hono.



Digital Twin Network using Eclipse Ditto + Hono.

Figure 4.1: Digital Twin Network

Eclipse Hono communicates with back-end business applications using Advanced Message Queuing Protocol (AMQP) 1.0. In our case it is connected to Eclipse Ditto (Eclipse (2017b)) which is an open-source frame-work for creating digital IoT twins. It focuses on back-end scenarios by providing web APIs in order to simplify working with already connected devices and things from customer apps or other back-end software. Ditto ensures that access to twins can only be done by authorized parties via the policies service. Eclipse Ditto structures the data sent by devices via Hono into Digital IoT Twins. A digital twin of an IoT device in Ditto is represented as a Thing consisting of attributes and features. The Things service is responsible for persisting things and features in Ditto.

Two channels are available for use in Ditto, a twin channel that connects to the digital representation of a Thing and a live channel that routes a command/message towards an actual connected device. Ditto uses MongoDB as a database and only stores the latest state of Things. Furthermore, Ditto provides an HTTP API and a WebSocket to applications to facilitate retrieving Things-related information.

The Digital Twin Network interacts with Ditto via this HTTP API or a WebSocket. The DTN stores things states in the database and updates them wherever a modification to the Things is made in Ditto. This is done because Ditto only stores the latest states of the Things in its data store (mongoDB). This interaction allows storing historical data of the Digital IoT Twins during the entire network lifecycle.

Chapter 5

Conclusion

The emergent use of DT applications across a number of domains is on the rise and, combined with enabling technologies such as big data, ML, advanced modeling,simulation and advanced communication interfaces, it enables insights on their physical twin's operation in a way that is useful and actionable by the designer or the operatorinsight leads to a data-driven decision making, which in some domains is the main advantage of DTs.
DT technology is still in its early stages and reaching its full potential will require addressing significant limitations and challenges for a modern DT implementation, such as: costs, information complexity and maintenance, a lack of standards and regulations and issues related to cybersecurity and communications.

Bibliography

[1] Digital twin | siemens.

[2] Stefan Boschert, Christoph Heinrich, and Roland Rosen. NEXT GENERATION DIGITAL TWIN.

[3] Diego M. Botín-Sanabria, Adriana-Simona Mihaita, Rodrigo E. Peimbert-García, Mauricio A. Ramírez-Moreno, Ricardo A. Ramírez-Mendoza, and Jorge De J. Lozoya-Santos. Digital twin technology challenges and applications: A comprehensive review. 14(6):1335.

[4] Latif U. Khan, Walid Saad, Dusit Niyato, Zhu Han, and Choong Seon Hong. Digital-twin-enabled 6g: Vision, architectural trends, and future directions.

[5] Zheng Liu, Norbert Meyendorf, and Nezih Mrad. The role of data fusion in predictive maintenance using digital twin. page 020023.

[6] Adil Rasheed, Omer San, and Trond Kvamsdal. Digital twin: Values, challenges and enablers.

[7] Rikard Söderberg, Kristina Wärmefjord, Johan S. Carlson, and Lars Lindkvist. Toward a digital twin for real-time geometry assurance in individualized production. 66(1):137–140.

[8] Fei Tao, Jiangfeng Cheng, Qinglin Qi, Meng Zhang, He Zhang, and Fangyuan Sui. Digital twin-driven product design, manufacturing and service with big data. 94(9):3563–3576.

[9] Ke Zhang, Jiayu Cao, Sabita Maharjan, and Yan Zhang. Digital twin empowered content caching in social-aware vehicular edge networks. 9(1):239–251.

[10] Yu Zheng, Sen Yang, and Huanchong Cheng. An application framework of digital twin and its case study. 10(3):1141–1153.