RESEARCH ARTICLE

Digital twin maturity levels: a theoretical framework for defining capabilities and goals in the life and environmental sciences [version 1; peer review: 3 approved with reservations]

Brett Metcalfe [1,2], Hendriek C. Boshuizen [3,4], Jandirk Bulens [5], Jasper J. Koehorst [1,6]

[1]Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Wageningen, Gelderland, The Netherlands
[2]Department of Earth Sciences, VU Amsterdam, Amsterdam, The Netherlands
[3]National Institute for Public Health and the Environment (RIVM), Bilthovenn, The Netherlands
[4]Division of Human Nutrition, Wageningen University & Research, Wageningen, The Netherlands
[5]Wageningen Environmental Research, Wageningen University & Research, Wageningen, Gelderland, The Netherlands
[6]UNLOCK Large Scale Infrastructure for Microbial Communities, Wageningen University & Research, Wageningen, Gelderland, The Netherlands

v1 **First published:** 10 Aug 2023, **12**:961
https://doi.org/10.12688/f1000research.137262.1
Latest published: 10 Aug 2023, **12**:961
https://doi.org/10.12688/f1000research.137262.1

Abstract

Background: Digital twins (DT) are the coupling of a real-world physical asset to a virtual representation to provide insight and actionable knowledge. The benefits of DT are considered to include improvements in reproducibility, reliability of interventions, increased productivity, as well as increased time for innovation. For instance, a DT could be used to boost agricultural productivity whilst also meeting various targets (e.g., biodiversity, sustainability). Or a DT could be used to monitor a cell culture, predict interactions, and make subtle adjustments to maintain the environment allowing researchers to conduct other work. Yet in developing DT two fundamental questions emerge: 'What will the DT capabilities be?' (i.e., the range of features and possible actions) and 'What will the DT do?' (i.e., which capabilities will it utilise).
Methods: Here we discuss a theoretical framework for DTs developed during Wageningen University & Research's Investment Programme on DTs that aims to answer these questions. Focusing on the Life and Environmental Sciences to help developers and stakeholders to agree on the capabilities, purpose, and goal of a DT. As well as identifying iterative design stages that may help set interim development goals such as a minimum viable product.
Results: This framework defines a DT as sitting at one of five maturity, or capability, levels associated with specific types of DT: a status, an informative, a predictive, an optimisation, and an autonomous twin.
Conclusions: The aim of DTs is to make better, data-driven, decisions

Open Peer Review

Approval Status ? ? ?

	1	2	3
version 1 10 Aug 2023	? view	? view	? view

1. **Koen Smit**, HU University of Applied Sciences, Ultrecht, Norway
2. **Mezzour Ghita**, HASSAN II University, Casablanca, Morocco
3. **Pratik Maheshwari**, Indian Institute of Management, Jammu, India

Any reports and responses or comments on the article can be found at the end of the article.

yet there can be a gulf between expectations of what a Digital Twin will do and the reality. The five maturity levels outlined here can be used to first identify and communicate about the type of Digital Twin required for a particular project prior to DT development. Bridging the gap between what project leads, developers, and stakeholders envision the end-product will be.

Keywords
Digital Twins, Maturity levels, Capabilities, Digital Twins for Life Sciences

Corresponding author: Brett Metcalfe (brett.metcalfe@wur.nl)

Author roles: Metcalfe B: Conceptualization, Investigation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Boshuizen HC**: Conceptualization, Funding Acquisition, Investigation, Supervision, Writing – Review & Editing; **Bulens J**: Conceptualization, Funding Acquisition, Investigation, Supervision, Writing – Review & Editing; **Koehorst JJ**: Conceptualization, Funding Acquisition, Investigation, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The authors disclosed receipt of the following financial support for the research, authorship, and publication of this article: The Wageningen University & Research Investment Programme "Digital Twins" for the Digital Twin Platform Methodology research that was (partly) subsidised by the Dutch Ministry of Agriculture, Nature, and Food quality (project number KB-41-004-002). J.J.K acknowledges the Dutch Research Council (NWO), and Wageningen University & Research for their financial contribution to the UNLOCK initiative (NWO: 184.035.007). This project has received funding under the European Union's HORIZON-INFRA-2022-TECH-01 research and innovation programme (grant agreement n° 101094287 BIOINDUSTRY 4.0).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Metcalfe B, Boshuizen HC, Bulens J and Koehorst JJ. **Digital twin maturity levels: a theoretical framework for defining capabilities and goals in the life and environmental sciences [version 1; peer review: 3 approved with reservations]** F1000Research 2023, **12**:961 https://doi.org/10.12688/f1000research.137262.1

First published: 10 Aug 2023, **12**:961 https://doi.org/10.12688/f1000research.137262.1

Introduction

Digital revolution

With the advent of concepts such as the Internet of Things (IoT) computers can now more readily 'sense' their environment in (near) real-time using a variety of embedded sensors all networked together.[1,2] Many of these IoT devices are everyday items,[1] others are small increasing their portability and ensuring that where there is the possibility of a connection to the internet, or another network, observations can be made.[2] One application of such interconnected sensors is the acquisition of data from an asset that maybe too big (*e.g.,* a city), remote (*e.g.,* streetlights), or inaccessible (*e.g.,* inside an air conditioning system) to study in detail with traditional data collection methods. Significantly, applying these interconnected sensors to a problem allows for the collection or generation of a continuous stream of data independent of analysts allowing for higher resolution, continuous datasets to be acquired. These developments are considered to have significantly furthered our progress to the next, fourth, iteration in our industrial development.[3,4]

The fourth industrial revolution (Industry 4.0 or 4IR) should see the transition from the digital revolution[4] - represented by the switch from mechanical and analogue systems to automated digital electronics - to industrial processes dominated by cyber-physical systems, smart sensors, and cognitive computing. Conceptually the fourth industrial revolution is built upon expanding the digital technology revolution[5,6] that occurred during the third industrial revolution through increased interaction and integration[7,8] between the virtual and physical (*i.e.,* "real world") environments. Such interaction and integration can take many forms (Figure 1), such as (digital) models that simulate processes that occur in the physical world; digital shadows or digital threads that collect, store, and provide information about a physical object; or digital generators, devices that construct physical objects (*e.g.,* 3D printers).

Digital twins

Associated with the coupling of the virtual and physical into cyber-physical systems is the potential formation of digital twins[9–11] with commercial[12–15] and academic[16,17] interest piqued regarding the possible application of digital twins to various complex processes and systems (*e.g.,* Ref. 18). Digital twins can be differentiated from other 'digital' variants (Figure 1) by the connection between the virtual and physical realms having a continuous automated two-way flow of information[16] (Figure 2). Through this communication, a virtual representation simulates how a process or physical asset operates throughout its lifecycle[19–21] from which decisions can be made that control the real-world process. Such simulations can be simple, for example keeping up to date readings of sensors that register the status of the physical asset. Or the virtual simulations can be anticipatory, using prior knowledge, time-series, and/or physics-based models to determine the likely future outcome based upon the current status of the physical asset. The feasibility of creating digital twins has existed for a time in the manufacturing and engineering sectors, where the intricacies of components and their complex processes are known in detail or can be intuited from blueprints.

From such information it is possible to place sensors that can generate data of sufficient quality and resolution to allow for the development of Digital twins that can monitor, control, and optimize processes. Or remove geographic and temporal constraints, so that operators can be trained remotely or decision maker can view historic, current and future states of the physical asset whilst not needing to be 'on location' or proximal to it.[22,23]

Digital twins in use

Beyond the factory floor the digital revolution is taking place in industries related to the Life and Environmental Sciences.[17] Such as Utilities,[18] the Agricultural sector,[24–28] Horticulture,[29] Apiculture,[30] Smart farming,[24] Animal farming,[31] as well as the Healthcare sector,[32,33] and the Pharmaceutical Industries[34] amongst others. Other twins aim to provide information on the built-environment (*e.g.,* planning, environmental impact, air quality, etc.), advice on sustainable water management [1], or to help protect and restore (European) biodiversity [2]. More ambitious digital twins aim to provide a digital twin of the Oceans [3] and our Planet [4] to assist in the green transition.[35]

Such an interest in the possible application of Digital Twins to the Life and Environmental sciences has been ignited by the emergence of new techniques and new technology that can circumvent the complexity of real-world phenomenon. New techniques include those that can derive patterns and connections without the explicit requirement to know the process behind such patterns (*e.g.,* machine learning) or those that enable a model to be kept synchronised with observations (*e.g.,* data assimilation). Alongside the development[36] of smaller more portable interconnected technology and devices,[37] loaded with sensors[38–40] that can acquire sufficient data to feed into the virtual half of the twin.

---

[1] https://www.wur.nl/nl/onderzoek-resultaten/onderzoeksinstituten/environmental-research/show-wenr/digital-twin-voor-waterbeheer.htm
[2] https://biodt.eu/news/biodt-digital-twin-prototype-help-protect-and-restore-biodiversity
[3] https://www.ocean-twin.eu/
[4] https://digital-strategy.ec.europa.eu/en/policies/destination-earth

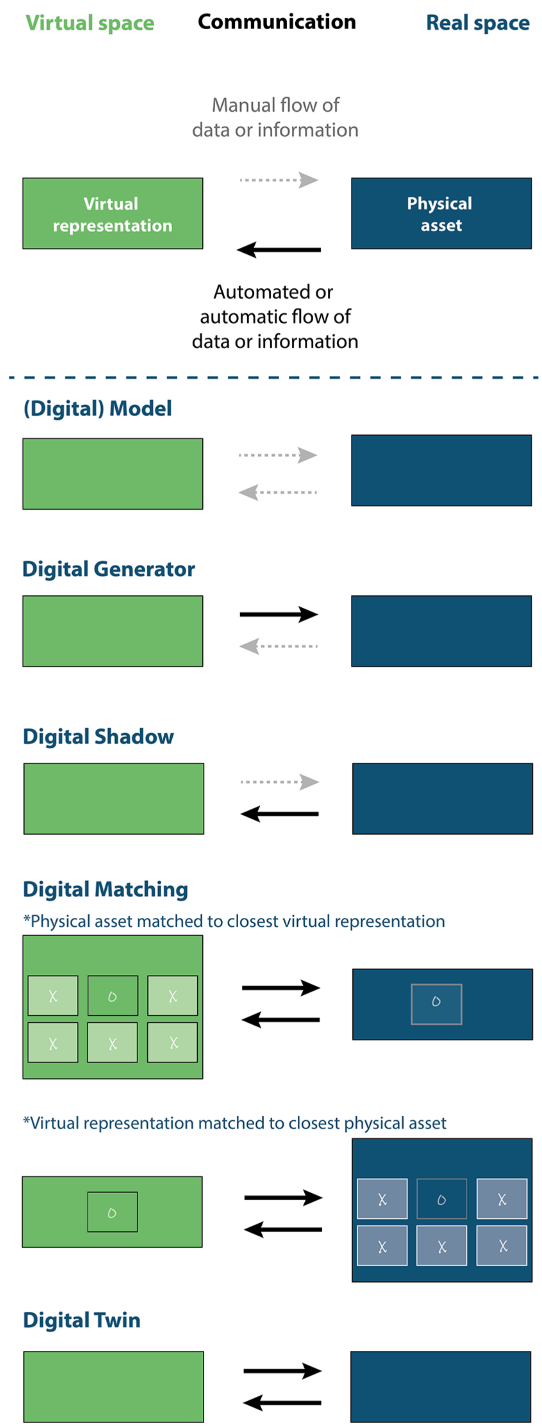

Figure 1. The current 'Digital' landscape. Variants of digital technologies characterized by how information flows between the real (blue) and virtual (green) worlds. Examples of these varieties include 3D printers and barcodes here referred to as a digital generator and digital shadow respectively. Solid black lines denote an automated flow of information, whereas grey lines denote a 'manual' flow of information. Based upon Refs. 7, 8.

Yet given their integration into various aspects of our daily lives' digital twins for the life and environmental sciences may have the potential to cause considerable socioeconomic upheaval. Such upheaval may occur through changing our interaction with ourselves with the development of personalised healthcare[38,40–42] and (personal) environmental monitoring. Or *via* adapting environments (*e.g.*, greenhouses) and organisms to meet our needs *e.g.*, optimisation of

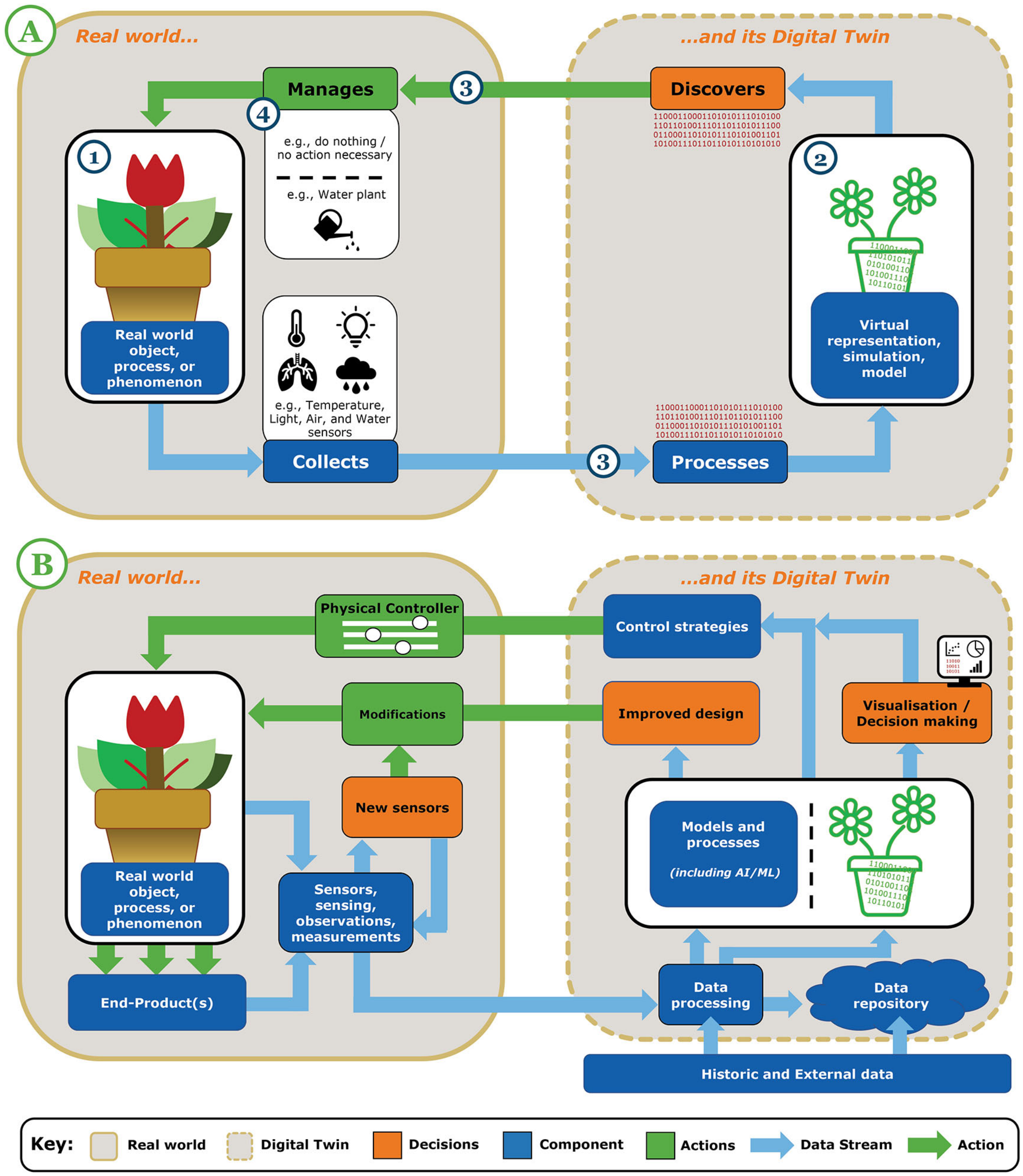

Figure 2. Concept and components of a digital twin. (A) Conceptual components of a digital twin, having a (1) physical asset; (2) virtual representation; (3) two-way communication; and (4) feedback or actionable knowledge. (B) In reality the various components maybe more complex. The terms used in (A) to describe the various stages, *i.e.,* 'Collects', 'Processes', 'Discovers', and 'Manages', are from Ref. 12.

agricultural practices. However, for all the possible benefits and opportunities of utilising digital twins for life science applications there are many barriers to success such as the 'blueprints' of a natural process or phenomenon being only partially known, too complex, technically unfeasible or unethical to be obtained. There is also the potential for such socioeconomic upheaval to exacerbate rather than curtail existing problems, such as those relating to social injustices or healthcare[43–46] or ensuring continuation of poor practices (*e.g.,* in agriculture[47]).

Aims and objectives

Given the socioeconomic opportunities and consequences applying of digital twins to life and environmental science questions[17] it is important to consider a framework that is communicable to a range of stakeholders that not only conveys what a digital twin is but what it should do. Ill-defined definitions can lead to ambiguous projects goals; be counter-productive for product development; create too high expectations; and limit stakeholder participation. Defining the capabilities of the digital twin and which of these capabilities will be used in each situation will be part of the design process for developing digital twins. Clarification over what a digital twin will do, ultimately defines the test of whether it is feasible, prudent or necessary to develop a digital twin for a given problem.[16,48,49] Likewise, a well-defined definition can help develop short-term goals to make a project achievable. Either by providing sufficient scope to a project to prevent scope-creep or through curtailing over ambition by allowing a project to be divided into a series of distinct iterations. Iterative design allows for a period of reflectance between design stages to assess the projects continued needs, requirements, direction, and goals. Furthermore, additional benefits includes having different iterations ready to apply to different problems. Here such a framework is outlined. First by discussing the definition of a digital twin (Section 'Definition and maturity'), then the conceptual framework behind maturity levels that represent the functionality and capabilities of a digital twin (Sections 'Definition and maturity' and 'Maturity levels'). Before outlining the maturity framework (Section 'Maturity levels') and its associated Digital Twin types (Section 'Types of digital twin') that is applicable to a range of applications in the life and environmental sciences.

Definition and maturity

Digital twin definition

There is no explicit formal definition of a digital twin leading to some divergence between various (scientific) disciplines (*e.g.,* Ref. 50) about what a digital twin actually is. To further exacerbate the problem studies have shown that in some papers what is referred to as a digital twin has not always been clarified[51] leading others to attempt to catalogue the variation in descriptions[3,7,16,24,52–54] including identifying what can or cannot be considered a digital twin.[7,8]

From a practical perspective here, it is considered that a definition of a digital twin is valid if it meets the following criteria of being a virtual representation, replica, or simulation of an actual, or possible, physical asset in (near) continuous two-way communication with its counterpart to play an active role in planning, decision, and control processes through providing feedback. The digital twin will be primarily data-driven and capable of, or have the capacity for, predicting future trends or developments over a period of time. Here feedback need not be an automated process but simply incorporation of the results of the twin in a decision-making process. There are four commonalities between the various types of digital twins: (1) a physical asset; (2) a virtual representation; (3) two-way communication; and (4) some form of feedback (Figure 2). In the following, let us expand upon these various components.

Physical asset

A digital twin must be comprised of a physical asset in physical space and a virtual replica, copy, analogy, or representation in virtual space. There is no exclusivity with respect to twinning, so the same physical asset can be twinned by multiple digital twins.[55] Likewise, the same sensors could be used for multiple twins. In some descriptions of digital twins whether they are a part or the whole asset has been differentiated by adding the following terms 'component', 'asset', 'system or unit', or 'process' to the term digital twins.[15,56] Where a 'component' is part of an asset and a 'system', 'unit', and 'process' refer to a greater and greater collection of assets. Therefore, in digital twinning the physical asset need not be the 'whole' instead it can be a singular component of a larger system (*e.g.,* a motor in a factory; the digestive tract of an animal; *etc*) or alternatively a group of individuals (*e.g.,* a fleet of delivery vehicles; a herd of cows; a field; *etc.*) as long as it can conceivably or does physically exist. Similarly, digital twins that describe a prototype of a physical asset, explicitly one that can exist, have been referred to as a Digital Twin Prototype to differentiate it from those with a real-world physical asset. In the life sciences the digital twin prototype can possibly occur in the biotech industry representing organisms in a natural state or as genetically modified organisms, where the cells may be modelled first prior to their creation.

Virtual representation

In the virtual environment, the physical asset is twinned by a virtual representation that should in theory provide the user with relevant information they could obtain were they to go out and physically inspect the asset in the real-world. Assuming that measurements, time, money, and distance are not limiting factors. Of course, in reality the twin utilises a subset of measurement components that either directly, or indirectly (*e.g.,* proxies) capture various states of the measured system. The virtual representation can be broadly split into a visualisation component (what the user sees) and a processing component (how the data is transformed, used, or incorporated into models). Some definitions of digital twins implicitly state that in order to have maximum fidelity with the real-world asset by necessity the visualisation component of virtual representation must be a 3D model of the asset. Here, however, the requirement of the digital twin to have a 3D visualisation is omitted as ultimately the visualisation should depend upon the (i) type of data; (ii) the goal of the visualisation; and (iii) the needs, wants, and requirements of the user. Hence, the visualisation components can range from simple iconography, a dashboard, graphs and graphics, a 3D model, augmented reality or virtual reality. Whereas the processing components can range from just the data, a data-model, physics-based model, artificial intelligence, or some combination of all of these. Ultimately the form the virtual representation takes will be dependent upon the knowledge of the targeted user group and the purpose of the digital twin.

Two-way communication

Between the virtual-physical and physical-virtual spaces two-way communication should occur (Figure 1). Of course, the nature and method of this two-way communication will vary. Communication between the physical and virtual worlds may be via networked sensors, data upload, or manual data-entry (*e.g.,* patient results). Communication between the virtual and physical worlds could be in the form a simple visualisation (*e.g.,* icon forms or composite interactive dashboard) or more complex feedbacks and actions (*e.g.,* automation).

Likewise, the frequency of communication will also vary being dependent upon the use case. For instance, high frequency communication may be required where rapid changes occur but just as equally when trying to pinpoint subtle variations. It may be prudent to first identify the physical assets' own unique trends and rhythms prior to defining the communication interval to ensure the right frequency of communication occurs. For example, if one was aiming to twin a plant then consideration must be made regarding the frequency of communication so as to sufficiently capture the various internal and external natural trends and rhythms (*e.g.,* diurnal or seasonal cycles). In other instances, there may be variables or components of the Digital Twin that are updated only infrequently. For example, when representing a complete city, the volume of traffic might be updated continuously whereas road layout itself is semi-rigid - being fixed but requiring updates at a vastly different time scale and often performed manually by cartographers.

Actionable knowledge

Finally, there must be an initiation of a decision-making process and where required feedback; with the Digital Twin providing the source of actionable knowledge to aid in the decision-making process. Actionable knowledge refers to the coupling of Science and Management.[57] That is using directed action to bring about intended consequences (*i.e.,* management) whose actions and consequences are defined by the generalisations provided by empirical research to describe or explain a particular phenomenon (*i.e.,* science). Actionable knowledge will be knowledge that can provide either the specific conditions or actions necessary to ensure or lead to a particular set of intended consequences beyond the settings in which it was first created, and which can be applied to both individuals as well as to groups.

Digital twin

In summary, a digital twin is composed of a virtual representation or replica of an actual, or possible, physical asset with these two counterparts in continuous communication so as to play an active role in planning, decision, and control processes. Following coupling between the physical and virtual twins both twins are collectively referred to as a Digital Twin Instance.[11,16]

Going beyond a definition

As can be seen from the proceeding section a definition does not inform the reader of what a digital twin will do for a given problem.[58] A concept, or definition, of Digital Twins will only provide a 'foundation level' knowledge, *i.e.,* knowing the definition of a digital twin serves to differentiate it from other types of digital technology (Figure 1) and what components are key (Figure 2A) but not much else. This is because for each criterion (*i.e.,* virtual asset; two-way communication; feedback; and a decision-making process) there can be a range of potential outcomes (Figure 2B) giving an equally wide variety of possible digital twin end products, or more aptly 'phenotypes' (*i.e.,* observable characteristics or traits). How the various components of a Digital Twin are expressed and formulated must be discussed during the development phase. The resultant Digital Twin will depend upon the end goal of the digital twin, as well as any legal and technological practicalities, and the various decisions made in the design phase.

Solving the question of what a particularly digital twin will do and its capabilities necessitates communication between developers, stakeholders, and users. Yet, communication between stakeholders requires a common frame of reference. Hence, in order to solve the issue of what a digital twin will do, several initiatives have placed digital twins onto a hierarchical framework (Figure 3), describing various types, capabilities, and the evolution[59–63] or developmental stage[26] of digital twins.[64] These have been referred to as digital twin maturity levels,[59,65,66] a series of digital twin integration levels,[7] a continuum,[60] a spectrum,[65,67] archetypes,[68] and/or a classification[62] or taxonomy of digital twins,[69] all of which seek to define the functionality of the various digital twin levels (Figure 3).

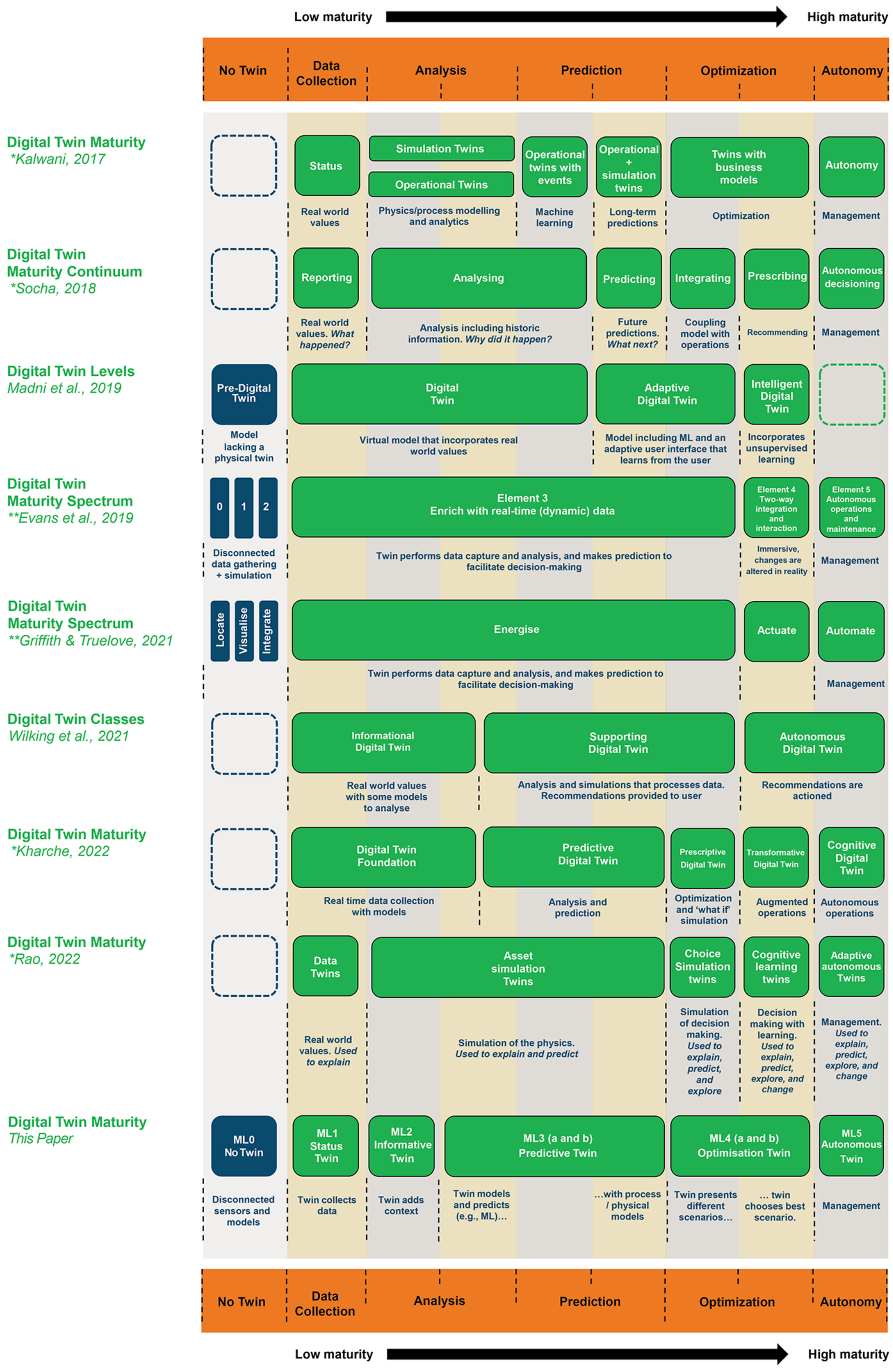

Figure 3. Synthesis of digital twin maturity levels. Comparison of the various levels of maturity[59–63,65,70,71] placed upon a relative scale where the sizes of the boxes reflects the coarseness of the granularity. Blue boxes represent levels that do not represent a Digital Twin. Whereas, dashed unfilled boxes represent levels not specifically defined by the schema. References with a * are from online (business) blogs, ** from technical reports, and others from academic journals.

Digital twin maturity

Previous work

As such investigating previous work is a useful starting point for defining a digital twin maturity model that is applicable to Life Science problems. However, several previous studies have expanded their models to incorporate other digital instances (Figure 1). For example, Kritzinger *et al.*'s[7] digital twin integration levels put digital twins on a continuum involving digital models and digital shadows (Figure 1). For brevity only previous work differentiating between different types of digital twins will be discussed here. Where the different types of digital twin reflect differences in functionality of the digital twin. As opposed to the type of asset being twinned (*i.e.,* Refs.15, 56).

A quick review of this subset of the literature (Figure 3) suggests that the types of digital twin are generally differentiated by the level of insight that the twin provides about a problem or toward a goal.[71] A digital twin having better insight can be seen as having increased capabilities alongside an increased capacity to perform the actions itself. It is our opinion that Digital twins can be broadly categorised into those that provide 'Data about', 'Knowledge of', 'Suggestions for' and 'Control of' a physical component, asset, process, or phenomenon. These can be split into data collection, analysis, prediction, optimisation, and autonomy (Figure 3). Differences between previous studies largely reflects the 'coarseness' of the various digital twin sub-divisions. For instance, 'twins with business models' reflecting a stage where the digital twin focuses upon optimising a process[59] can itself be sub-divided into various substages: an initial substage coupling the twin to business operations, a subsequent substage recommending various decisions,[60] and finally making those decisions autonomously.[63] Yet despite differences between the assorted maturity models there is a general consensus that to reach the foundational level of a digital twin requires an availability of data[71] and/or a connection between the physical and virtual worlds.[65,70] The lowest level of a digital twin is generally reserved for a twin that simply relays the collected data, or real-time status, and the highest level when the twin has gained autonomy to control.[72] With many prior maturity models referencing automobile automation or the SAE's levels of driving automation for self-driving cars[73] as a comparable frame of reference (*e.g.,* Ref. 59). Between these two end members (data collection and automation) there are levels devoted to analysis (including modelling, simulating), prediction, and optimising (Figure 3).

Conceptual framework

For the framework outlined here we continue the use of the term maturity[59,65,66] given how it denotes both a natural, and sequential, progression toward more complex states (*e.g.,* see the data-driven manufacturing maturity model[74]). Although it is entirely feasible to also refer to it in terms of functionality or capability. Likewise, our maturity levels are also based on the presumption that the ultimate goal of digital twins is to streamline decision making processes through the automation of sequential tasks. Whether or not the ultimate goal of an autonomous twin is obtainable is another matter (see, Discussion) and there is some debate regarding digital twins and cyber-physical systems (*e.g.,* Ref. 53). In our framework the various maturity levels can be conceptualised as a series of progressive stages[74] that add functionality towards the goal of automation.[73] At low digital twin maturity levels, the user or operator of the digital twin carries out more of the data capture, decision making, process and control operations. As a digital twin matures processes once controlled by human operators become subsumed by the virtual twin, until the process is fully automated. These levels focus upon providing insight upon a particular aspect of the object under study, with aspects that address the what's, the how's, and the why's. What questions relate to conveying the basic research elements, (raw) data and information ('what is the time', 'what value does X have'). How questions relate to understanding the relations of these basic research elements, *e.g.,* through models, from which predictions can be inferred. Why questions relate to identifying cause and effect as well as whether something can or should be done. Why can be sub-divided into 'appreciation of why' and 'why'. This sub-division is based upon the idea that there are instances where the reasoning or knowledge behind why something occurs may not be known ('why') but an answer or explanation can still be formulated ('appreciation of why').

To explain the conceptualisation let us consider a simple example, that of traffic in a city. If we want to know the amount of traffic on the road, we can begin by asking a simple question, "What is the number of vehicles on the road?". Such a question relates solely to the measurement of vehicles on the road (*i.e.,* "6 vehicles" or "5 cars, 1 bus, …") and it is up to the receiver of that data to provide the context. By changing our question to "What is the current volume of traffic?" context and meaning are added to the answer (*i.e.,* "low", "high"). But it might be prudent to know "How will traffic change over the course of a day, week, or month?" and therefore, understanding the variables necessary for prediction and modelling becomes necessary. Knowing the amount of traffic can be used to provide insight for a practical use, for example selecting the optimum route: "What routes are there from location to A to destination B?". Finally, such knowledge could be used to address "What is the best route and best time to leave from A to get to B?"

The basis of this conceptualisation is the knowledge pyramid (or knowledge hierarchy; information hierarchy; knowledge management pyramid),[75–77] a common feature in information and knowledge management.[77] The knowledge pyramid is represented by a hierarchy commonly presented as a pyramid proceeding from the base of data to information, knowledge, understanding, and ultimately the pinnacle representing wisdom. To transition from data to information requires understanding relations; information to knowledge requires understanding patterns; and knowledge to wisdom requires understanding principles. The stages of the pyramid hierarchy can be mapped to the conceptualisation's different questions, the data and information relating to 'knowing what', knowledge to 'knowing how', and wisdom to 'knowing why'. Understanding is linked to 'appreciating why', the why might not be fully understood but cause and effect might be. The mapping between the hierarchy and these concepts has been argued by Zeleny[76] who stated that to "manage wisely implies knowing why to do something; to manage effectively implies knowing what to do; to manage efficiently implies knowing how to do it (and to 'muddle through' implies nothing and having 'lots of data' around)" [5].

Maturity levels

From low to high the five maturity levels outlined here deal with 'knowing what' (data); 'knowing what' (information); 'knowing how' (knowledge); 'appreciating why' (understanding split into: understanding and insight); and 'knowing why' (wisdom).

At the data level the digital twin focuses upon data collection and generation of values that represent the various properties of the physical asset, presenting the user with its current state. The information level focuses on structuring data into a coherent and organised format, adding context, and ensuring that current and historic data are presented to the user. Knowledge involves the processing of this structured data using previous experience, either through identifying underlying pattern and trends or physics-based process models. Understanding is the exploration of 'if/then' situations through scenario testing and interactions, understanding is further divided into insight which focuses upon optimisation and developing new relationships. Finally, the wisdom level is reached, at this point the digital twin is fully autonomous providing feedback and control through understanding.

The maturity levels can be visualised as a matrix (Figure 4) in which the columns represent the various types of digital twin and the rows the concept, these can be represented as questions and aspects of the knowledge pyramid. The cells of the matrix are divided into tasks performed by a digital twin (green boxes in Figure 4C) and not by a digital twin (blue boxes in Figure 4C).

Types of digital twin

These maturity levels (ML) can be represented by distinct types of digital twin (*e.g.,* Ref. 59) which can be referred to from low to high as: status, informative, predictive, optimisation, and autonomous digital twins. The five types (Figure 4) of digital twin are defined as follows:

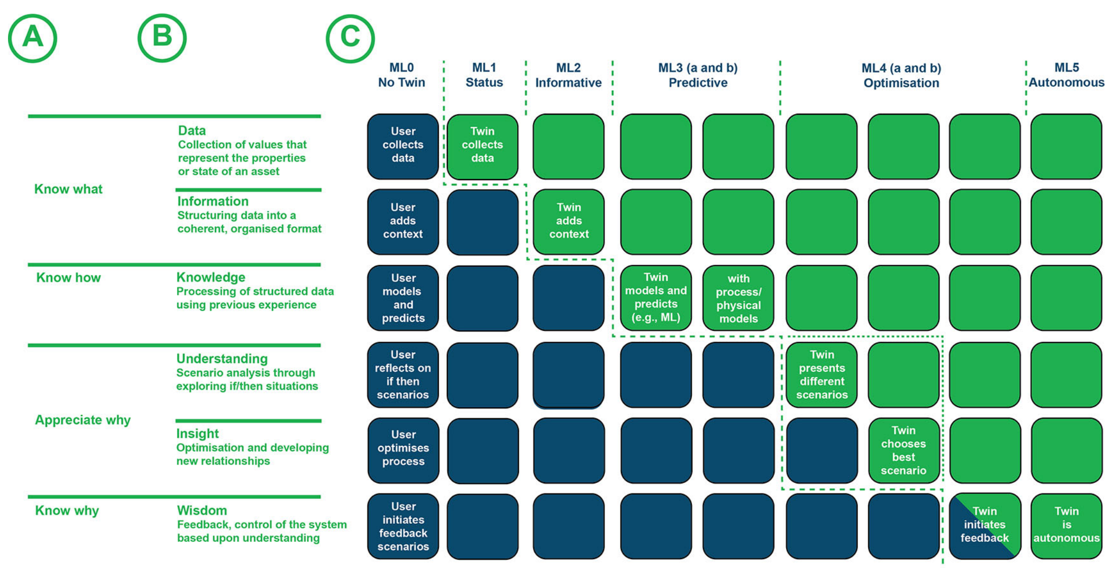

Figure 4. Digital twin maturity levels. The maturity index is composed of five levels that address the level category of a twin. The maturity levels can be visualised as a matrix in which the columns represent the various types of digital twin, and the rows are the concepts which can be represented as either (A) questions or (B) aspects of the knowledge pyramid. (C) The cells of the matrix are divided into tasks performed by (Green) and not by (Blue) the digital twin.

[5]Note that since publication of Zeleney's[76] work the levels in the knowledge pyramid that relate to 'what' and 'how' have switched. At the time data was how and knowledge was what, now data is what and knowledge is how. Here the quote is used in full as per the original unmodified.

- Digital twin maturity level 0 (ML0): No Twin. Measurements are disparate or non-existent, sensors are either not present or if they are, they lack any (inter) connected networking. The process maybe described by a process or physics model however, these are not linked to the real-world data and/or apply fixed boundary conditions.

- Digital twin maturity level 1 (ML1): Status. Focus upon real-time data capture, collection and visualization of information.

- Digital twin maturity level 2 (ML2): Informative. Real-time data capture and collection and incorporation of historical data or normative data (benchmarks) and visualization of the information.

- Digital twin maturity level 3 (ML3): Predictive. Real-time data capture and collection with historic data coupled to a machine learning or physical process-based model.

- Digital twin maturity level 4 (ML4): Optimisation. Real-time data capture and collection with historic data coupled to a machine learning or physical process-based model. The Virtual Twin allows for scenarios of different 'if then/what if' scenarios to be explored. Helping the operator to deduce the outcome of particular decisions.

- Digital twin maturity level 5 (ML5): Autonomous. Real-time data capture and collection with historic data coupled to a machine learning or physical process-based model. The Virtual Twin deduces the optimal scenario and enacts controls that lead to it occurring.

Discussion

Digital twin development

Digital twins are being developed in a variety of disciplines both commercially and academically.[17,18,24–34] Therefore, these projects may encompass a wide range of use cases and stakeholders who may not even have a cursory knowledge of what a digital twin is, let alone an understanding of what it will do. Likewise, there won't be a 'one-size-fits-all' approach to developing digital twins (as outlined in the 'Digital twin definition' and 'Going beyond a definition' sections above). Therefore, as previously noted, development of a digital twin boils down to two fundamental methodological questions. The first question relates to the full range of the digital twins features and possible actions, *i.e.,* 'What will the digital twins capabilities be?'. The second aims to address which of the capabilities identified in the initial question the digital twin will it utilise, *i.e.* 'What will the digital twin do?'. The answer to these questions can be considerably broad, especially depending upon the intricacies and peculiarities of the use case (*e.g.* note the range of expressions in Ref. 17). Critically it is important that in the process of describing what a digital twin will do that we end up with less of a wish list and more of a list of requirements for development. The levels of digital twin maturity proposed here aims to serve as a means for effective communication between project leaders, project owners, project participants, developers, and stakeholders about the intention of the project. Allowing developers to outline to interested parties what the goal and purpose of the digital twin will be. The maturity model should also act as a springboard in defining what is feasible with current resources; and, to know where a digital twin is in its development phase.

Example: Twinning a house plant

As a simple example, let us consider how the maturity levels could be applied to twinning something simple such as a house plant (Figure 2). The first step in digital twin development will be to perform an initial scoping exercise that will preliminary determine what type of digital twin (Figure 4) suits the intended objective.[59] This exercise should identify the asset under study; an intended goal; prior knowledge; measurable and required variables; define processes; determine the expected values of key variables; define the intended users and stakeholders; and consider if and what output and (anticipated) feedback the digital twin will produce. For a house plant (our object) there are several parameters that can be directly affected by an individual (our stakeholder) to promote growth (our goal). These include providing sufficient light, water, and temperature, all of which can be measured to a greater or less degree of accuracy by commercially available sensors. Once the initial scoping exercise has been performed, a digital twin developer should use the answers obtained to sketch out the rough divisions between the various maturity levels. Using coarser divisions such as data collection and communication (ML 1-2); analytics, models, and predictive tools (ML3); and 'use cases' (ML 4-5) can help make preliminary design decisions. Following which, it is prudent to next identify previous or preliminary work that already fulfils or can feed into one or more of the maturity levels. Applying the maturity model follows an iterative design methodology with each level, or iteration, being a cycle of development, testing, analysis, and reflection. The result of each learning cycle can then be used in subsequent design iterations, either another iteration at the same maturity level or progression onward.

Assuming that the process starts from the initial maturity level (ML1), design choices should address what sensors will be used and how these sensors will communicate their data to the virtual twin. For example, questions to be addressed include 'What is the nature of the communication?' *e.g.*, WiFi, LoRa, (*etc.*), 'How frequent will these sensors communicate?' or 'Will data arrive as a stream or be sent in batches?' amongst others. This level adds low level knowledge of a process, such as insight into range and distribution of obtainable values and/or how such variables change spatially or temporally. Such knowledge may help develop understanding of a process. Returning to our example, a digital twin for a house plant at ML1 provides real-time data capture of the amount of light, the soil moisture, and the room temperature. Values can be collected by an IoT sensor and presented on a display next to the plant, or as values on a simple web server.

Having obtained and communicated data in ML1, at the next maturity level, ML2, such data will be given context. Here 'context' is conveying known information about the subject matter, either from prior knowledge or previous data. Defining which context will be added requires communicating with experts, users, or a combination of both to identify key parameters. For instance, for our house plant, context could be added through a knowledge rule that translates values, such as the light into 'low', the soil moisture into 'dry', or the temperature into 'cold'. One could colour code the values on the display or add simple iconography to indicate (un)suitable conditions for the plant, allowing someone to identify problems and come up with their own actions. Such as moving the plant to a sunnier or shadier position, to water it, or to adjust the thermostat.

Progressing to the next level requires the twin to shift from solely focusing on past and present values to predicting future values. Hence, ML3, involves adding predictive capacity. For our example that could be knowing the future value of soil moisture and when it will drop below a specific level that requires watering. Generally, predictions will be made via a model, though the exact type (*e.g.*, mechanistic, process, physics based, machine learning, *etc.*) will depend the nature of the digital twin (*e.g.*, the use case; availability of data; *etc.*). Though developers can utilise the previous levels to identify required inputs and outputs as well as those variables that are of importance, those that vary, and which could be conceivably parameterised. For instance, for our house plant the digital twin could incorporate a mechanistic model that calculates the length of day and the angle of the Sun by giving the latitude, date, and time which can then be used to constrain the upper or lower light levels depending on weather. Whereas, predicting soil moisture and/or temperature could make use of a time series machine learning algorithm to make predictions without direct inference of the cause of the variation. However, crucially the model(s) must allow for the updating of various states. Techniques such as data assimilation can be used to combine new observations and model forecasts to update the prediction.

The final levels (ML4 and ML5) transition from relaying knowledge (ML1 to ML3) to creating and using new knowledge, with these levels being split between optimisation (ML4) and automation (ML5). How either of these levels will be ultimately implemented by a developer will be highly dependent upon the needs and requirements of users, experts, and stakeholders as well as taking into account various practicalities. The distinction between ML3 and ML4 is the addition of potential corrective actions and their outcomes. In our house plant example, optimisation could focus on resolving questions such as: 'When is the optimal moment to water to maintain constant soil moisture levels?' or'Is it better to water when the room is cooler?'. How those questions and their answers are presented will again depend on project specifics, for example they could be embedded in a user interface or developers could allow a user to test them out in a sandbox environment. Once it is possible to optimise a process, then a natural progression is the possible addition of some degree of automation, *i.e.*, bypassing the decision maker to quickly and consistently provide a corrective action. Therefore, at the final maturity level, ML5, the process is autonomous. For our example, as the moisture levels drops the digital twin would react by watering the plant to sustain an optimal amount of soil moisture.

Choosing the right twin

Whilst our example discusses the full spectrum of maturity levels projects developing digital twins should be aiming for and settling upon the right maturity level for a particular use case rather than reaching for higher maturity levels. This might seem far from aspirational however, the goal should be to have the best fit of digital twin to a particular goal or problem. Nor is it the intention of the maturity levels presented here is to produce a ranking system. As it is more than likely that specific digital twin use cases will have an upper maturity level limit that reflects the legal and ethical considerations as well as technological feasibility. For instance, for a given problem the highest level of a Digital Twin is dependent upon consensus on the expected values a set of variables should have and the certainty of relevant knowledge (Figure 5 based upon concepts in Refs. 59, 78–83). In situations where there is a lack of both knowledge and certainty of the values, a Status Digital Twin could provide the necessary foundational knowledge. Where there is consensus on the values and high degree of certainty regarding our knowledge of a process, then an Autonomous Digital Twin might be more prudent.

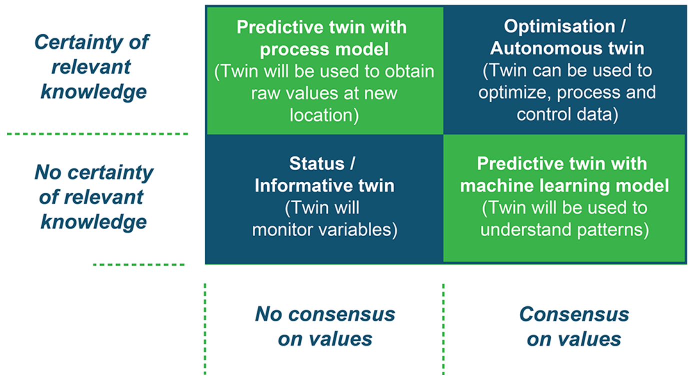

Figure 5. Selecting the right twin. Which twin is for you? Comparing certainty of knowledge verses consensus of values. Deciding on which digital twin is required for a given problem also necessitates a critical look upon known knowns, known unknowns and unknown unknowns. In this example the initial digital twin is dependent upon two variables, whether consensus is reached on the values a variable should have allowing for knowledge rules to be produced and the certainty of relevant knowledge. Based upon ideas and concepts relating to digital twins and uncertainty communication described elsewhere.[59,78–83]

Technological readiness

Our house plant example shows that the capabilities of a digital twin will however, be constrained by the availability and feasibility of technology. Automating the watering of a house plant is far more feasible than automating the movement of a plant to track light. Therefore, in this case the more automated levels will focus upon solutions that are feasible. However, it should be noted that the maturity levels outlined here do not directly refer to the technological readiness of Digital Twins or their constituent components. Technological readiness being defined by the technical progress or maturity of a technology under evaluation, a technology is ranked along a technology readiness level (TRL) scaled from research is beginning (= 1) to full scale deployment (= 9). For digital twins the feasibility of obtaining each of the maturity levels defined here will depend on the individual project, its intended goal, and thus the technology associated with it. In some instances, it is conceivable that the technology associated with higher maturity levels (*e.g.,* machine learning models, control systems, *etc.*) are more ready than the technology associated with lower maturity levels (*e.g.,* automated data capture and ingestion). Where automated or networked sensors do not yet exist for example, lower maturity levels that rely on data collection may not be obtainable but higher maturity levels such as modelling (using other or legacy datasets) may not be so constrained. Therefore, a mapping between maturity and technological readiness remains highly dependent on the topic of the digital twin. Although, it is the authors opinion that the maturity levels outlined here and the TRL are complimentary. The TRL will reflect the state of technical innovation[7,54,84] and key enabling technologies.[85,86] By using the maturity index to first identify and communicate about the type of digital twin required for a project and its goals, developers can use the TRL at the next phase of project development to identify technological feasibility and/or bottlenecks.

Conclusion

As part of Industry 4.0, physical processes, assets, or phenomenon will be in the future linked to a virtual representation referred to as digital twins. Digital twins are physical assets linked by a two-way flow of information to a virtual representation that is intended to provide actionable knowledge. These characteristics can be expressed in different ways making knowing the type, format, and capabilities of a digital twin somewhat ambiguous. To ensure fruitful dialogue between stakeholders and digital twin developers here we presented a model of the capabilities along a path to a fully autonomous digital twin. This maturity index is composed of five levels that address 'knowing what' (data); 'knowing what' (information); 'knowing how' (knowledge); 'appreciating why' (understanding and insight); and 'knowing why' (wisdom). Each represented by a distinct type of digital twin: status; informative; predictive; optimisation; autonomous. These types of digital twin share commonalities and the natural progression between the levels can be used to define the human and computer tasks. At low maturity levels, such as the status digital twin, the operator will perform many of the complex tasks. Proceeding from this low maturity level denoted by low complexity and few autonomous attributes the types of digital twin become increasingly complex and automated.

Author contributions

All authors contributed to the conceptualisation and writing.

Data availability

No data are associated with this article.

References

1. IBM: *What is the internet of things?* IBM.
Reference Source

2. Holdowsky J, Mahto M, Raynor M, *et al.*: **Inside the Internet of Things (IoT) A primer on the technologies building the IoT.** *A primer on the technologies building the IoT.* Deloitte University Press; 2015.

3. Negri E, Fumagalli L, Macchi M: **A Review of the Roles of Digital Twin in CPS-based Production Systems.** *Procedia Manuf.* 2017; **11**: 939–948. 2351-9789.
Publisher Full Text

4. Maynard AD: **Navigating the fourth industrial revolution.** *Nature Nanotech.* 2015; **10**(12): 1005–1006. Number: 12 Publisher: Nature Publishing Group. 1748-3395.
PubMed Abstract | Publisher Full Text

5. Morisson A, Pattinson M: **Industry 4.0.: A policy brief from the policy learning platform on research and innovation. Technical report, Interreg Europe Policy Learning Platform, Lille, France.** 2019.

6. Nolan A: **The next production revolution: Key issues and policy proposals.** OECD, editor. *The Next Production Revolution: Implications for Governments and Business.* Paris; OECD Publishing; 2017; pp. 25–70. 978-92-64-27115-9.
Reference Source

7. Kritzinger W, Karner M, Traar G, *et al.*: **Digital Twin in manufacturing: A categorical literature review and classification.** *IFAC-PapersOnLine.* 2018; **51**(11): 1016–1022. 24058963.
Publisher Full Text

8. Tekinerdogan B, Verdouw C: **Systems architecture design pattern catalog for developing digital twins.** *Sensors.* 2020; **20**(18): 5103.
PubMed Abstract | Publisher Full Text | Free Full Text

9. Främling K, Holmström J, Ala-Risku T, *et al.*: **Product agents for handling information about physical objects. Technical Report TKO-B 153/03, Helsinki University of Technology Laboratory of Information Processing Science, Helsinki, Finland.** 2003.
Reference Source

10. Shafto M, Conroy M, Doyle R, *et al.*: **Appendix n: Ta11 modeling, simulation, and information technology and processing.** *NASA Space Technology Roadmaps and Priorities: Restoring NASA's Technological Edge and Paving the Way for a New Era in Space.* Washington, DC: The National Academies Press; 2012; pp. 282–293. 978-0-309-25362-8.
Publisher Full Text

11. Grieves M, Vickers J: **Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems.** Kahlen F-J, Flumerfelt S, Alves A, editors, *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches.* Cham: Springer International Publishing; 2017; pp. 85–113. 978-3-319-38756-7.
Publisher Full Text

12. Boje C, Guerriero A, Kubicki S, *et al.*: **Towards a semantic construction digital twin: Directions for future research.** *Autom. Constr.* 2020; **114**: 103179.
Publisher Full Text

13. Zhang X, Zhu W: **Application framework of digital twin-driven product smart manufacturing system: A case study of aeroengine blade manufacturing.** *Int. J. Adv. Robot. Syst.* 2019; **16**(5): 1729881419880663. 1729-8806. Publisher: SAGE Publications.
Publisher Full Text

14. Bamberg A, Urbas L, Bröcker S, *et al.*: **The digital twin – your ingenious companion for process engineering and smart production.** *Chem. Eng. Technol.* 2021; **44**(6): 954–961. 1521-4125.
Publisher Full Text

15. Melesse TY, Di Pasquale V, Riemma S: **Digital twin models in industrial operations: State-of-the-art and future research directions.** *IET Collob. Intell. Manuf.* 2021; **3**(1): 37–47. 2516-8398.
Publisher Full Text

16. Jones D, Snider C, Nassehi A, *et al.*: **Characterising the Digital Twin: A systematic literature review.** *CIRP J. Manuf. Sci. Technol.* May 2020; **29**: 36–52. 1755-5817.
Publisher Full Text

17. Knibbe WJ, Afman L, Boersma S, *et al.*: **Digital twins in the green life sciences.** *NJAS: Impact in Agricultural and Life Sciences.* 2022; **94**(1): 249–279. 2768-5241.
Publisher Full Text

18. Pedersen AN, Borup M, Brink-Kjær A, *et al.*: **Living and Prototyping Digital Twins for Urban Water Systems: Towards Multi-Purpose Value Creation Using Models and Sensors.** *Water.* 2021; **13**(5): 592. 2073-4441. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
Publisher Full Text

19. Verdouw CN, Kruize JW: **Digital twins in farm management: illustrations from the FIWARE accelerators SmartAgriFood and Fractals PA17 – The International Tri-Conference for Precision Agriculture in 2017, Hamilton (2017).** *Precision Agriculture Association New Zealand, Hamilton, New Zealand.* 2017; page 154.
Publisher Full Text

20. Verdouw CN, Beulens AJM, Reijers HA, *et al.*: **A control model for object virtualization in supply chain management.** *Comput. Ind.* 2015; **68**: 116–131. 01663615.
Publisher Full Text

21. Grieves M: **Digital twin: manufacturing excellence through virtual factory replication. White Paper, Florida Institute of Technology, Florida, US.** 2014.
Reference Source

22. Verdouw CN, Beulens AJM, van der Vorst JGAJ: **Virtualisation of floricultural supply chains: A review from an internet of things perspective.** *Comput. Electron. Agric.* 2013; **99**: 160–175. 0168-1699.
Publisher Full Text

23. Verdouw CN, Wolfert J, Beulens AJM, *et al.*: **Virtualization of food supply chains with the internet of things.** *J. Food Eng.* 2016; **176**: 128–136. 02608774.
Publisher Full Text

24. Verdouw C, Tekinerdogan B, Beulens A, *et al.*: **Digital twins in smart farming.** *Agric. Syst.* 2021; **189**: 103046. 0308-521X.
Publisher Full Text

25. Bronson K, Knezevic I: **Big data in food and agriculture.** *Big Data Soc.* 2016; **3**(1): 2053951716648174. 2053-9517. Publisher: SAGE Publications Ltd.
Publisher Full Text

26. Pylianidis C, Osinga S, Athanasiadis IN: **Introducing digital twins to agriculture.** *Comput. Electron. Agric.* 2021; **184**: 105942. 0168-1699.
Publisher Full Text

27. Purcell W, Neubauer T: **Digital twins in agriculture: A state-of-the-art review.** *Smart Agricultural Technology.* 2023; **3**: 100094. 2772-3755.
Publisher Full Text

28. Pylianidis C, Snow V, Overweg H, *et al.*: **Simulation-assisted machine learning for operational digital twins.** *Environ. Model Softw.* 2022; **148**: 105274. 1364-8152.
Publisher Full Text

29. Ariesen-Verschuur N, Verdouw C, Tekinerdogan B: **Digital twins in greenhouse horticulture: A review.** *Comput. Electron. Agric.* 2022; **199**: 107183. 0168-1699.
Publisher Full Text | Reference Source

30. Johannsen C, Senger D, Kluss T: **A digital twin of the social-ecological system urban beekeeping.** Kamilaris A, Wohlgemuth V, Karatzas K, *et al.*, editors. *Advances and New Trends in Environmental Informatics.* Cham: Springer International Publishing; 2021; pages 193–207. 978-3-030-61969-5.

31. Neethirajan S, Kemp B: **Digital Twins in Livestock Farming.** *Animals.* 2021; **11**(4): 1008. 2076-2615. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Liu Y, Zhang L, Yang Y, *et al.*: **A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin.** *IEEE Access.* 2019; **7**: 49088–49101. 2169-3536. Conference Name: IEEE Access.
Publisher Full Text

33. Duggal R, Brindle I, Bagenal J: **Digital healthcare: regulating the revolution.** *BMJ.* January 2018; **360**: k6. 0959-8138, 1756-1833. Publisher: British Medical Journal Publishing Group Section: Editorial.
Publisher Full Text

34. Finelli LA, Narasimhan V: **Leading a Digital Transformation in the Pharmaceutical Industry: Reimagining the Way We Work in**

Global Drug Development. *Clinical Pharmacology & Therapeutics.* 2020; **108**(4): 756–761. 1532-6535.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Bauer P, Stevens B, Hazeleger W: **A digital twin of Earth for the green transition.** *Nat. Clim. Chang.* February 2021; **11**(2): 80–83. 1758-6798. Number: 2 Publisher: Nature Publishing Group.
Publisher Full Text

36. Porter ME, Heppelmann JE: **How Smart, Connected Products Are Transforming Competition.** *Harv. Bus. Rev.* 2014; **92**(11): 65–88.
Reference Source

37. Ray T, Choi J, Reeder J, *et al.*: **Soft, skin-interfaced wearable systems for sports science and analytics.** *Curr. Opin. Biomed. Eng.* 2019; **9**: 47–56. 2468-4511.
Publisher Full Text

38. Teo JX, Davila S, Yang C, *et al.*: **Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging.** *Commun. Biol.* 2019; **2**(1): 310–61. 2399-3642. Number: 1 Publisher: Nature Publishing Group.
PubMed Abstract | Publisher Full Text | Free Full Text

39. Nag A, Afsrimanesh N, Mukhopadhyay SC: **Impedimetric microsensors for biomedical applications.** *Curr. Opin. Biomed. Eng.* 2019; **9**: 1–7. 2468-4511.
Publisher Full Text

40. Marsch LA: **Digital health data-driven approaches to understand human behavior.** *Neuropsychopharmacology.* January 2021; **46**(1): 191–196. 1740-634X. Number: 1 Publisher: Nature Publishing Group.
PubMed Abstract | Publisher Full Text | Free Full Text

41. Fagherazzi G: **Deep Digital Phenotyping and Digital Twins for Precision Health: Time to Dig Deeper.** *J. Med. Internet Res.* 2020; **22**(3): e16770.
PubMed Abstract | Publisher Full Text | Free Full Text

42. Lydon-Staley DM, Barnett I, Satterthwaite TD, *et al.*: **Digital phenotyping for psychiatry: accommodating data and theory with network science methodologies.** *Curr. Opin. Biomed. Eng.* March 2019; **9**: 8–13. 2468-4511.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Wilbanks JT, Topol EJ: **Stop the privatization of health data.** *Nature News.* 2016; **535**(7612): 345–348. Section: Comment.
Publisher Full Text

44. Zarocostas J: **How to fight an infodemic.** *Lancet.* 2020; **395**(10225): 676. 0140-6736, 1474-547X. Publisher: Elsevier.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Editorial: **More accountability for big-data algorithms.** *Nature News.* September 2016; **537**(7621): 449. Section: Editorial.
PubMed Abstract | Publisher Full Text

46. Bruynseels K, Santoni de Sio F, van den Hoven J: **Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm.** *Front. Genet.* 2018; **9**: 1664-8021. Publisher: Frontiers.
PubMed Abstract | Publisher Full Text | Free Full Text

47. Bronson K: **A digital "revolution" in agriculture?: Critically viewing digital innovations through a regenerative food systems lens.** Duncan J, Carolan M, Wiskerke JSC, editors. *Routledge Handbook of Sustainable and Regenerative Food Systems.* Routledge; August 2020; pp. 336–349. 978-0-429-46682-3. Publication Title: Routledge Handbook of Sustainable and Regenerative Food Systems.
Publisher Full Text

48. Muse ED, Topol EJ: **Digital orthodoxy of human data collection.** *Lancet.* 2019; **394**(10198): 556. 0140-6736, 1474-547X. Publisher: Elsevier.
Publisher Full Text

49. Huckvale K, Venkatesh S, Christensen H: **Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety.** *NPJ Digit. Med.* September 2019; **2**(1): 11–88. 2398-6352. Number: 1 Publisher: Nature Publishing Group.
PubMed Abstract | Publisher Full Text | Free Full Text

50. Torkamani A, Andersen KG, Steinhubl SR, *et al.*: **High-definition medicine.** *Cell.* 2017; **170**(5): 828–843. 0092-8674, 1097-4172. Publisher: Elsevier.
PubMed Abstract | Publisher Full Text | Free Full Text

51. Barricelli BR, Casiraghi E, Fogli D: **A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications.** *IEEE Access.* 2019; **7**: 167653–167671. 2169-3536.
Publisher Full Text

52. Wright L, Davidson S: **How to tell the difference between a model and a digital twin.** *Adv. Model. Simul. Eng. Sci.* March 2020; **7**(1): 13. 2213-7467.
Publisher Full Text

53. Boyes H, Watson T: **Digital twins: An analysis framework and open issues.** *Comput. Ind.* 2022; **143**: 103763. 0166-3615.
Publisher Full Text

54. Liu M, Fang S, Dong H, Xu C: **Review of digital twin about concepts, technologies, and industrial applications.** *J. Manuf. Syst.* 2021; **58**: 346–361. 0278-6125.
Publisher Full Text

55. Walsh J, Tolle D: **ASSESS-Theme-Positioning-Paper-Digital-Twins-1.10.pdf. ASSESS Theme Positioning Paper Version 1-10, ASSESS Initiative.** 2020.
Reference Source

56. Manca L, Grugni R, Mirzazadeh R: **Digital twin: A digital copy of reality that enables you to simulate and find answers in a risk free and secure environment. Technical report, Engineering (eng.it).** 2019.
Reference Source

57. Argyris C: **Actionable Knowledge: Design Causality in the Service of Consequential Theory.** *J. Appl. Behav. Sci.* December 1996; **32**(4): 390–406. 0021-8863. Publisher: SAGE Publications Inc.
Publisher Full Text

58. Broo DG, Schooling J: **Digital twins in infrastructure: definitions, current practices, challenges and strategies.** *Int. J. Constr. Manag.* August 2021; **23**: pp. 1254–1263. 1562-3599, 2331-2327.
Publisher Full Text

59. Kalwani S: **The Evolution of Digital Twins for Asset Operators.** 2017.
Reference Source

60. Socha D: **The Digital Twin Maturity Continuum.** September 2018.
Reference Source

61. Madni A, Madni C, Lucero S: **Leveraging Digital Twin Technology in Model-Based Systems Engineering.** *Systems.* January 2019; **7**(1): 7. 2079-8954.
Publisher Full Text

62. Wilking F, Schleich B, Wartzack S: **Digital twins - definitions, classes and business scenarios for different industry sectors.** *Proceedings of the Design Society.* 2021; **1**: 1293–1302. 2732-527X.
Publisher Full Text

63. Rao AS: **Evolution of digital twins open data science conference.** May 2022.
Reference Source

64. Singh M, Fuenmayor E, Hinchy EP, *et al.*: **Digital twin: Origin to future.** *Applied System Innovation.* 2021; **4**(2): 36. 2571-5577. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.
Publisher Full Text

65. Evans S, Savian C, Burns A, *et al.*: *Digital twins for the built environment: An introduction to the opportunities, benefits, challenges and risks. Technical report.* Hertfordshire, UK: The Institution of Engineering and Technology (IET); 2019.
Reference Source

66. Korovin G: **Digital twins in the industry: Maturity, functions, effects.** Kumar V, Leng J, Akberdina V, *et al.*, editors. *Digital Transformation in Industry, Lecture Notes in Information Systems and Organisation.* Cham: Springer International Publishing; 2022; pp. 1–12. 978-3-030-94617-3.
Publisher Full Text

67. Bolton A, Butler L, Dabson I, *et al.*: *The Gemini Principles: Guiding values for the national digital twin and information management framework. Technical Report CDBB_REP_006.* Cambridge, UK: Centre for Digital Built Britain; 2018.
Publisher Full Text

68. van der Valk H, Haße H, Möller F, *et al.*: **Archetypes of digital twins.** *Bus. Inf. Syst. Eng.* 2022; **64**(3): 375–391. 1867-0202.
Publisher Full Text

69. Pai M, Parle D, Smith C: *A Digital Twin Taxonomy for Industry Implementations. Technical report.* Bangalore, India: Wipro Limited; December 2019.
Reference Source

70. Griffith C, Truelove M: **Framework for spatially enabled digital twins: Information paper. Information Paper, CSIRO Data 61, Australia.** 2021.

71. Kharche V: **Digital Twin Maturity.** January 2022.
Reference Source

72. Assad Neto A, Ribeiro da Silva E, Deschamps F, *et al.*: **Digital twins in manufacturing: An assessment of key features.** *Procedia CIRP.* 2021; **97**: 178–183. 2212-8271.
Publisher Full Text

73. On-Road Automated Driving (ORAD) committee: *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Ground Vehicle Standard J3016_202104.* USA: SAE International; 2021.
Publisher Full Text

74. Weber C, Königsberger J, Kassner L, *et al.*: **M2ddm – a maturity model for data-driven manufacturing.** *Procedia CIRP.* 2017; **63**: 173–178. 2212-8271.
Publisher Full Text

75. Ackoff RL: **From Data to Wisdom Presidential Address to ISGSR, June 1988.** *J. Appl. Syst. Anal.* 1989; **16**: 3–9.

76. Zeleny M: **Management support systems: Towards integrated knowledge management.** *Hum. Syst. Manag.* 1987; **7**(1): 59–70. 0167-2533. Publisher: IOS Press.
Publisher Full Text

77. Rowley J: **The wisdom hierarchy: representations of the DIKW hierarchy.** *J. Inf. Sci.* April 2007; **33**(2): 163–180. 0165-5515. Publisher: SAGE Publications Ltd.
Publisher Full Text

78. Bergsma P: **Besluiten onder onzekerheid Hoe gebruiken bestuurders van de Nederlandse waterschappen statistische informatie over onzekerheid bij beslissingen over wateroverlast? Master's thesis, Wageningen University and Research, Wageningen, The Netherlands.** 2016.
Reference Source

79. Wehrens R, ten Broeke G, Stigter H, *et al.*: *Digital Twins for decision making; The Uncertain half of the Twin. Wageningen Environmental Research WENR 3071.* Wageningen, The Netherlands: Wageningen Environmental Research; 2021.
Publisher Full Text

80. Boogerd A: *Van Droge Kennis naar Natte NatuurDe Interactie tussen Natuurwetenschap en Beleid over Verdroging.* Utrecht: Uitgeverij Eburon; 2005. 9059720768/9789059720763.
Reference Source

81. Turnhout E, Hisschemöller M, Eijsackers H: **Science in Wadden Sea policy: from accommodation to advocacy.** *Environ. Sci. Pol.* 2008; **11**(3): 227–239. 1462-9011.
Publisher Full Text

82. Petersen AC: *Simulating Nature: A philosophical study of computer-simulation uncertainties and their role in climate science and policy advice.* 1st ed. Apeldoorn – Antwerpen: VU University Amsterdam PhD Proefschrift. Het Spinhuis; 2006. 978 90 5589 280 8.
Reference Source

83. Huijs S: **Het vertellen van verhalen: een politieke manier van omgaan met onzekerheid.** van Asselt M, Petersen A, editors. *Niet bang voor onzekerheid.* Vol. 1. Den Haag, NL: Boom Juridische uitgevers; 2003; pp. 105–118. 978-90-5931-287-6.

84. Reimsbach-Kounatze C: **Benefits and challenges of digitalising production.** OECD, editor. *The Next Production Revolution: Implications for Governments and Business.* Paris: OECD Publishing; 2017; pp. 73–117. 978-92-64-27115-9.
Reference Source

85. Commission of the European Communities: *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: "Preparing for our future: developing a common strategy for key enabling technologies in the EU" - {COM(2009) 512}. Commission Staff Working Document SEC(2009)1257.* Brussels: Commission of the European Communities; 2009.

86. Butter M, Fischer N, Gijsbers G, *et al.*: *Horizon 2020: Key enabling technologies (kets), booster for european leadership in the manufacturing sector. Study for the ITRE Committee IP/A/ITRE/2013-01 PE 536.282, European Parliament. Directorate-General for Internal Policies.* Brussels: Policy Department A: Economic and Scientific Policy; 2014.

Open Peer Review

Current Peer Review Status: ? ? ?

Version 1

Reviewer Report 16 October 2023

https://doi.org/10.5256/f1000research.150414.r208826

?

Pratik Maheshwari
Indian Institute of Management, Jammu, India

This paper introduces an intriguing topic, "Digital twin maturity levels: a theoretical framework for defining capabilities and goals in the life and environmental sciences." However, the paper's implications and contributions to the present state of the art are somewhat limited. Several observations and recommendations are outlined below:

1. The introduction provides a clear overview of the research area. However, it would be beneficial to more explicitly emphasize the significance and relevance of digital twin maturity levels in the context of life and environmental sciences.

2. The literature review is comprehensive and well-organized. It effectively establishes the groundwork for the proposed framework. However, to enhance the paper's currency, consider incorporating more recent research findings, especially those published within the last two years.

3. The research objectives and framework are well-defined and aligned with the research problem. Nevertheless, it would be helpful to provide clarification on how this framework differentiates itself from existing ones, where applicable.

4. While the paper briefly outlines the methodology employed in developing the framework, enhancing its credibility could be achieved by providing more comprehensive details regarding research methods and data sources.

5. The proposed digital twin maturity levels framework is presented clearly and logically. Nonetheless, supplementing it with practical examples or case studies that illustrate its application within the life and environmental sciences would enhance its comprehensibility.

6. The discussion section offers valuable insights into the implications of different maturity levels. However, a more detailed analysis of potential challenges and limitations in

implementing these levels would provide greater depth to the paper.

7. It is advisable to highlight practical applications and real-world scenarios where this framework can prove beneficial in the life and environmental sciences. Additionally, discussing how it can address specific challenges within these fields would be insightful.

8. The paper's writing requires improvement; there are areas where language could be clarified to enhance understanding. Consider rephrasing sentences or providing additional explanations as needed.

9. The figures and tables are indeed helpful, but ensuring they are labeled and explained with utmost clarity is essential. If feasible, consider incorporating additional visual aids to further enhance reader comprehension.

10. The conclusion effectively summarizes the paper's key findings and contributions. However, it would be advantageous to suggest potential directions for future research or extensions of the framework, thereby providing a more forward-looking perspective.

11. It is recommended to meticulously review the references to ensure completeness and accuracy in citations. Additionally, verify that they conform to the chosen citation style consistently.

12. Lastly, emphasize the paper's contribution to the digital twin field and its potential impact on the life and environmental sciences. A concise summary of these aspects can enhance the paper's overall significance

Is the work clearly and accurately presented and does it cite the current literature?
No

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
No

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Supply chain Digital Twin

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 16 October 2023

https://doi.org/10.5256/f1000research.150414.r208820

?

Mezzour Ghita

National and High School of Electricity and Mechanic (ENSEM), HASSAN II University, Casablanca, Morocco

The paper proposes a five maturity level framework that discuss all the aspects for proper integration of digital twins especially for life science application. The authors draw an initial macroscopic view what it can be defined as a digital twin and gives all relevant description of the element constituting the context for digital twins. The presented work describes several contributions that can help in the practical development and implementation of DT in the field. However some modifications are needed to enhance technical aspects of the work.

The authors in their introduction claims that the framework is intended to help integration of DT for life and environmental sciences this cannot be apprehended through out the paper except for the use case that gives just an overview of how the framework can be exploited. It would be relevant to highlight what industry 4.0 technologies are already in use in the field and where they fail including for digital twins. Authors can get inspiration from some patent developed in the agricultural field and also from work on industry 5.0 that deals same obejctives. The aim would be to define a proper field related problematic that value the proposed framework in this context. Some interesting papers are referred by the authors and patents can give some additional value to this.

"There is no explicit formal definition of a digital twin" actually there is work in progress for the definition of a digital twin standard by ISO including formal definition of the context. Its true that in the field of interest discussed here it was not defined yet formally but in general the norm would serve as a reference in this case.

The framework presents maturity levels through a knowledge pyramid perspective but at the end no explicit process is defined of how this levels can be apprehend its left to the user to decide what questions, measures, criteria can help to achieve this evaluation. The use case give some insights but no explicit procedure is defined which at the end makes it difficult for technical evaluation of the framework and its feasibility. To deal with this the authors proposed to include TRL but again no explicit description is provided for replication. TRL can also help to include some additional relevant points that hinder efficient and large scale inclusion of DT in some fields as instance communication, costs and all project specific management aspects. This point can be highlighted as a following research direction of the study.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital Twins, Smart manufacturing, Advanced Driver Assistance Systems

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 08 September 2023

https://doi.org/10.5256/f1000research.150414.r198894

?

Koen Smit
HU University of Applied Sciences, Ultrecht, Norway

Dear authors,

Thank you for proposing your contribution titled "*Digital twin maturity levels: a theoretical framework for defining capabilities and goals in the life and environmental sciences*" for publication in F1000 Research.

The paper aims to answer two questions about which capabilities are tied to a DT and how the capabilities will be utilized by the DT. The paper proposes a theoretical view on five maturity stages for DT capabilities. The specific focus of the contribution is interesting as the life sciences domain can potentially benefit greatly from digital twin technology and the body of knowledge should converge on practices to help structure the design of this complex technology to be used successfully in practice. Based on my analysis, I have the following considerations for the authors to improve the contribution:

- Abstract/background should focus more on life sciences as the claim that DT capabilities must be explored is not true for, e.g., the manufacturing industry, where DT's are used for quite some time now. Consider pointing out the relevance for the specific domain better.

- I have some difficulty with the concept of 'new sensors' on top of sensing (in blue) in figure 2 on page 5. While I like the notion that on a detailed level, digital twins are more complex due to all concepts mentioned in figure 2b, the orange part should be explained better or left out. Adding to this, consider adding to figure 2 that, depending on the level of required capabilities, complexity increases as it now presents a situation in which all DTs should be enhanced with AI/ML, which is not always required.

- One of the main statements in the contribution about why such a framework should be researched is as follows: "*Ill-defined definitions can lead to ambiguous projects goals; be counterproductive*

- *for product development; create too high expectations; and limit stakeholder participation*." Please consider adding strong sources for this as there are many contributions that focus on these issues and readers should be informed about this knowledge.

- Another point that could be added is that big bang approaches in developing DT solutions often fail because of complexity and therefore fail to meet expectations. Iterative development is mentioned in the contribution, so I agree that this helps in expectation management and has a higher chance of successful implementation in practice. Again, there are, in the context of DT, but surely outside of that as well (general IS-research), a lot of contributions that point towards this universal problem and helps make the point of avoiding big bang approaches. Consider discussing this, accompanied with appropriate sources, in the aims and objectives-section of the paper.

- The contribution presents multiple other works from the DT body of knowledge that focus on (the quantification of) maturity levels and models. How this maturity model adds to the already available knowledge should be discussed more extensively in the beginning of the contribution.

- I find figure 3 very comprehensive and must complement the authors on the informative value it brings to the paper, though consider enlarging the figure a bit to improve readability.

- Are the results of the subset of contributions in the DT body of knowledge complete? The authors do not explain how these papers were selected. For example, why is the following reference, which is very relevant in the context of this study, not included (disclaimer: I am not a (co)-author of this specific reference):

    - Uhlenkamp, J. F., Hauge, J. B., Broda, E., Lütjen, M., Freitag, M., & Thoben, K. D. (2022). Digital twins: A maturity model for their classification and evaluation. *IEEE Access*, *10*, 69605-69635.

- My previous comment also points out a larger problem with the contribution, regarding its methodological quality. While the contribution is theoretically scoped, I would expect details about how the selection/filtering in the DT body of knowledge is performed to derive the presented levels and characteristics in figure 3, 4 and 5. This is a major flaw in the contribution that should be improved upon.

- Figure 3s most left column is categorized as 'no twin'. While I know that multiple papers refer to this lowest stage of maturity in a similar manner, consider delving into what it can be referred to as, because the current label is hardly informative for practitioners. The textual information presented in this column does not really help in informing the reader. Later, the focus of the discussion seems to point towards more disconnected sensors and models of both the physical and virtual world, so disconnected or non-linked. If that is the largest caveat for this level, why not refer to it as such in the model?

- The example presented in page 12 contains a paragraph in which a 'knowledge rule' is used to explain logic that is required to translate data into contextual information or categories. The term knowledge rule is new to me, as far as I am aware such logic is often referred to as business rules. Consider replacing the concept of knowledge rules or explain more specifically what they are (if different from business rules/logic).

- The authors state the their opinion is that the Technological Readiness Level is important to take into account in the context of utilizing DT technology and establishing maturity levels. The levels presented in the contribution are complementary to the TRL, but how that translates to the model exactly is kept in the dark, unfortunately. Consider illustrating or describing how they complement each other as it is stated that TRL should be taken into account. At least provide an example, e.g., in conjunction with the example provided earlier in the sections.

- Even if the contribution is theoretically scoped, I would expect the authors to discuss possible limitations of the study and its results in a dedicated discussion section. Unfortunately, such a section is not present. Also, future research directions in this (and other) domain(s) are important notions to base further research upon. However, these are not presented at the end of the paper. Consider adding these sections as they contribute towards proper academic discussion and follow-up.

Minor spelling/grammar:

- Page 3: real-world phenomenon should be: phenomena (plural)

- Please consider carefully checking the entire paper for punctuation accuracy as, for example, I found some comma's to be incorrectly used in sentences.

References

1. Uhlenkamp J, Hauge J, Broda E, Lutjen M, et al.: Digital Twins: A Maturity Model for Their Classification and Evaluation. *IEEE Access*. 2022; **10**: 69605-69635 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
No

Are sufficient details of methods and analysis provided to allow replication by others?
No

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Digital Twin Technology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.