CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

The philosophical foundations of digital twinning

David J Wagg[1,2] (iD), Christopher Burr[1] (iD), Jason Shepherd[3], Zack Xuereb Conti[1] (iD), Mark Enzer[4] and Steven Niederer[1,5] (iD)

[1]The Alan Turing Institute, London, NW1 2DB, United Kingdom
[2]Department of Mechanical Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom
[3]Fujitsu Services Limited, Lovelace Road, Bracknell, RG12 8SN, United Kingdom
[4]Mott MacDonald, 8-10 Sydenham Road, Croydon, CR0 2EE, United Kingdom
[5]Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom
Corresponding author: David J Wagg; Emails: david.wagg@sheffield.ac.uk, dwagg@turing.ac.uk

Abstract

Digital twins are a new paradigm for our time, offering the possibility of interconnected virtual representations of the real world. The concept is very versatile and has been adopted by multiple communities of practice, policymakers, researchers, and innovators. A significant part of the digital twin paradigm is about interconnecting digital objects, many of which have previously not been combined. As a result, members of the newly forming digital twin community are often talking at cross-purposes, based on different starting points, assumptions, and cultural practices. These differences are due to the philosophical world-view adopted within specific communities. In this paper, we explore the philosophical context which underpins the digital twin concept. We offer the building blocks for a philosophical framework for digital twins, consisting of 21 principles that are intended to help facilitate their further development. Specifically, we argue that the philosophy of digital twins is fundamentally holistic and emergentist. We further argue that in order to enable emergent behaviors, digital twins should be designed to reconstruct the behavior of a physical twin by "dynamically assembling" multiple digital "components". We also argue that digital twins naturally include aspects relating to the philosophy of artificial intelligence, including learning and exploitation of knowledge. We discuss the following four questions (i) What is the distinction between a model and a digital twin? (ii) What previously unseen results can we expect from a digital twin? (iii) How can emergent behaviours be predicted? (iv) How can we assess the existence and uniqueness of digital twin outputs?

Impact Statement

Creating digital twins (or the process of digital twinning) is a concept of growing importance in a wide range of industries and technology sectors. Digital twins can be used as a method to obtain value from data and as deployment platforms for AI and data-science techniques such as machine learning and statistical analysis. In many applications, digital twins offer the means to integrate together multiple previously separate components in order to achieve a specified objective(s). This type of integration of digital components is based on a fundamentally holistic philosophy. This paper presents a conceptual framework for digital twins that considers how such a holistic integration can be achieved, including current questions of interest, and challenges for future research.

Check for updates

1. Introduction

A digital twin is a virtual representation of a natural, engineered, or social system[1] (called the *physical twin*) that enables a two-way coupling between the digital and physical domains, using some form of network-based connectivity (that is bidirectional flow of information typically across the Internet). The digital twin evolves over time and is constructed from digitized information such as recorded data and the output of computational models.

The origins of digital twinning are usually attributed to the work of NASA during the Apollo program, where physical duplicates were used to help mission control support their astronauts respond to a critical failure with their oxygen tanks and engine (Rosen et al., 2015). However, the term "digital twin" itself first appears in work relating to product lifecycle management (see Grieves 2019 and discussion therein). The idea has received considerable attention since then in a wide range of areas including product design, manufacturing, civil infrastructure, medicine, asset management, health/condition monitoring, energy networks, space structures, and nuclear fusion.[2]

Digital twins have been promoted as a way to accelerate our ability to understand engineering (and other) systems at previously unmatched levels of performance. This vision and aspiration were captured in the quote from Eric Tuegel and his coauthors (in the context of structural life prediction) in 2011 who stated that:

> "The digital twin is a reengineering of structural life prediction and management. Is this science fiction? It is certainly an audacious goal that will require significant scientific and technical developments. But even if only a portion of this vision is realized, the improvements in structural life prediction will be substantial"—Tuegel et al. (2011).

This is certainly a very exciting prospect. However, it is important to always maintain a healthy level of skepticism when dealing with such claims. The aspiration for digital twins, particularly from commercial vendors, seems to imply that the new technology will somehow capture and contain "the best of everything" in some optimal way (e.g. efficient assemblage and federation between models, data, machine learning methods, processes, controls, decision, etc.). In addition, it is often implied that digital twins will somehow overcome the fundamental challenges and limitations related to modeling that we already have (e.g. limited computational resources, epistemic gaps), enabling benefits such as improved fidelity, trust, and insight. But how exactly might that happen? When such questions are not satisfactorily answered, the conclusion for some is that the whole idea is over-hyped, skepticism can become cynicism, and genuine scientific and technological progress can become stalled.

We believe that part of the underlying issue stems from the fact that the concept of a digital twin is so versatile and universally applicable, and so it is open to a very wide range of interpretations—as evidenced by recent reviews (Korenhof et al., 2021). Those interpretations come from a large number of different research and practitioner communities, which themselves have very wide-ranging cultures and practices. However, a significant part of the digital twin paradigm is about interconnecting these previously unconnected domains. For example, building socio-technical digital twins is a major ambition in this field (Okita et al., 2019; Wang et al., 2020; Zhang et al., 2021a; Savage et al., 2022; Yossef Ravid and

---

[1] See Committee on Foundational Research Gaps and Future Directions for Digital Twins et al. (2024) for a more in-depth discussion of these different types of systems, and their incorporation into a comprehensive definition.

[2] For those readers that might be interested in the history, development and applications of digital twins there are multiple detailed descriptions of these (and many other) topic areas in the growing number of review papers on the topic of digital twins including Ríos et al. (2015); Negri et al. (2017); Kritzinger et al. (2018); Cimino et al. (2019); Enders and Hoßbach (2019); Boje et al. (2020); Errandonea et al. (2020); Jones et al. (2020); Liu et al. (2020); Melesse et al. (2020); Wagg et al. (2020); Wanasinghe et al. (2020); He and Bai (2021); Huang et al. (2021); Jiang et al. (2021); Lo et al. (2021); Semeraro et al. (2021); Shahat et al. (2021); Korenhof et al. (2021); Botín-Sanabria et al. (2022); Purcell and Neubauer (2022); Singh et al. (2022); Somers et al. (2022); Tao et al. (2022); Jafari et al. (2023); Liu et al. (2023); Sepasgozar et al. (2023); Thelen et al. (2023); Dale et al. (2023); Xames and Topcu (2024); Katsoulakis et al. (2024); Piras et al. (2024); Wang et al. (2024); Jin et al. (2024).

Aharon-Gutman, 2022). As a result, when conversations happen, people are often talking at cross-purposes, because they have different starting points, cultural assumptions, biases, and motivations.

Therefore, in this paper, we seek to understand the philosophical context (or foundations) that underpin the concept of a digital twin. We will argue that the philosophy of digital twins is fundamentally based on the idea of *holism*. In short, certain properties represented by a digital twin apply only to wholes formed out of assemblies of more basic parts, such that it is conceptually incoherent to apply these properties to lower-level components of the system. Furthermore, a key (and related) aspiration for digital twins is that they can capture emergent phenomena. However, as we will describe, these are not the only philosophies that relate to digital twinning. Little has been previously published on the philosophy of digital twins, specifically, with some notable exceptions, such as the work of Korenhof et al. (2021), who have proposed that digital twins are "steering representations"—something we discuss later on.

It is important to note that, because holism and emergentism are typically at odds with the reductionist paradigm used in the majority of current mainstream science and engineering models, the approach is neither well established nor well understood. To support our argument, we will review several aspects related to the philosophy of modeling. In the second half of the paper, we present a set of *principles for digital twinning* along with some examples that are intended to support the development of a more complete philosophical framework—an exercise we leave for further research. Some readers may wish to jump ahead to these principles prior to reading the first half of the paper, to get an idea of our destination. Others may wish to review the foundational context set out in the first half. We leave this as a choice for the reader.

The paper is structured as follows. In Section 2, we present the philosophical context for the concept of digital twins. The concept of modeling will be key to this discussion, and we will explore related philosophical topics and issues that will help the reader understand the motivation behind the principles presented in Section 4. In Section 3, we then turn to consider the types of complexity that occur in engineering systems, and how this might be represented in a digital twin. With the context set, in Section 4, we then introduce a series of principles that we argue could form a conceptual framework for digital twins, or the process of *digital twinning*. This framework is then used to suggest answers to four key questions relating to digital twins. Finally, in Section 5, we conclude and suggest several open questions and future directions for research.

2. Philosophical issues and concepts for digital twins

Although there has been much discussion on the potential definitions relating to digital twins (Semeraro et al., 2021; Committee on Foundational Research Gaps and Future Directions for Digital Twins et al., 2024), one area that has not received much attention is the philosophical underpinnings of digital twinning as a distinct concept and practice.[3] In this section, we seek to address this gap.

We begin by reviewing key topics in the philosophy of modeling as applied to a wide range of scientific and engineering domains, before turning to related concerns (e.g. epistemic issues). Models are very important for digital twins because they are one of the key *components* that make up a digital twin.[4] However, the relationship between models and digital twins can also be a common source of confusion (e.g. what distinguishes a digital twin from a mere model or simulation).[5] Therefore, it is helpful to first unpack some key issues that will help us in our later discussions.

---

[3] Though see Korenhof et al. (2021) for a notable exception.

[4] Throughout this paper, we will deal primarily with *computational* models, because of our focus on digital twins. However, there is a significant number of different model types that are relevant in the context of digital twins (e.g. statistical models, data-driven & machine learning (ML) models, hybrid models, finite element models, computational fluid dynamics models, agent-based models, behavioural models, discrete event simulation models, process models, analytical models, empirical models, black-box/white-box models, Markov models, graphical models, multi-physics models, geospatial models, fuzzy logic models, neurosymbolic models, etc.). These types of models will overlap with one another due to a common set of "family resemblances", such as their "ability to represent", the fact that they are "idealised" in some sense (e.g. toy models), or because they can aid in increased knowledge and understanding.

[5] We will discuss this in more detail in Section 4.2.
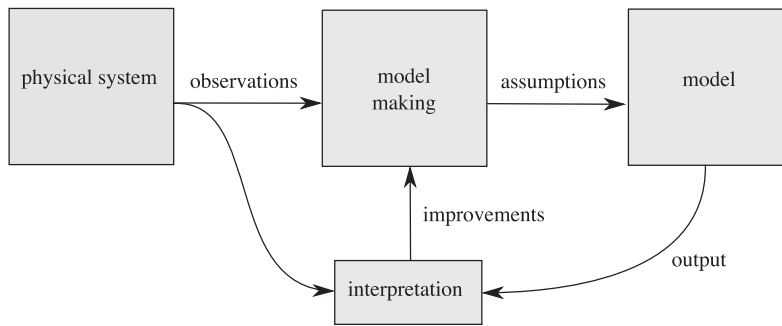
Figure 1. *Schematic diagram showing the typical method of making a model of a physical system. The physical system can be a process or a material object.*

2.1. Some key concepts in the philosophy of modeling

The history of philosophical approaches to modeling as a concept and practice goes back as least as far as ancient periods of Greek, Chinese, and Indian thought (Curd et al., 2014). It is not our goal to rehearse or review this entire history.[6] Rather, our goal is simply to understand key parts of the philosophical context of digital twinning in order to motivate the presentation of the principles for digital twinning set out in Section 4. Therefore, we identify and discuss some *key concepts* that underpin how a model of a physical system is typically made in a science and engineering context—with an emphasis on drawing out certain implicit assumptions. To support this goal, an idealized example of such a model-making process is shown schematically in Figure 1.

The process starts with a series of "observations" (or measurements) of some physical system. This presumes that there is already a physical system in existence, which may not be the case in engineering when we are asked to design something not previously built. Some discussion on this is given in Wagg et al. (2020), but for now, we assume that a physical system is available for observation. These observations are then used to construct a model based on a set of "assumptions". Next, the "output(s)" of the model (e.g. predictions) are then interpreted, and in many cases, this leads to "improvements" being made to the model (e.g. optimizing the performance of predictions), and the process is repeated as often as deemed necessary.

The idealized process of model making (or abstraction) shown in Figure 1 allows us to locate and separate some key philosophical concerns: *natural laws*, *determinism*, *reductionism*, and *holism & emergence*.

In science and engineering, a "physical system" is typically assumed to be mechanistic in nature. This mechanistic worldview is characterized by a belief in a set of "natural laws" that govern physical systems. And, science seeks to discover these laws or rules, relies upon them to apply "universally," and then utilizes them within subsequent processes of model-making. The value of such a belief is articulated well by Descartes, who said that:

> "…reliable rules which are easy to apply, and such that if one follows them exactly, one will never take what is false to be true or fruitlessly expend one's mental efforts, but will gradually and constantly increase one's knowledge till one arrives at a true understanding of everything within one's capacity."—René Descartes: Rules for the Direction of the Mind (see reprint: Descartes, 1985, first published 1701).

---

[6] Readers are referred to Oberkampf and Roy (2010) and Curd et al. (2014) for comprehensive overviews of modelling in the scientific and engineering domains. Note also that the word "model" in this context is used to mean a "representation", e.g. an abstraction, of the physical system on interest. It could be any type of model used in the context of engineering practice depending on the context. However, a point we return to later is that models can be made of both *processes* and *material objects*.

Natural laws, such as those of physics, biology, and chemistry, provide a foundational set of "rules" describing how real-world systems operate. Because digital twins aim to represent these real-world systems in a virtual environment, an understanding of these laws allows researchers and developers to ensure that their models accurately reflect these fundamental rules (e.g. incorporating Newton's laws of motion into a digital twin of a spaceship).

"Determinism" connects to this concept of natural laws insofar as it is assumed that if natural laws are strictly followed, then future states of a system can be precisely predicted[7]. And, *if* the physical world is deterministic, then digital twins should, in theory, be able to predict the future states of a system given accurate input data. These beliefs (among others) are the first set of assumptions that cascade through the typical model-making process and affect our "observations"—assumptions that we will later challenge.

A belief in a law-governed, deterministic world, however, does not preclude the possibility of a complex world. Physical systems can be highly complex (e.g. economies or ecologies)[8], and so it is common to (conceptually) decompose (or reduce) physical systems into "simpler" parts to be more effectively represented and studied. By studying this reduced version, it is assumed, useful information about the complete system can be obtained (Heylighen et al., 2007).

For the purpose of this discussion, we can consider two forms of reductionism. The first is *component-based reductionism* which involves dividing the physical system into separate physical (e.g. geometric or process) components, and if required dividing these components into smaller and/or simpler parts, as required. The second is *physics-based reductionism*, which is to simplify, approximate, or even neglect entirely some part of the physics. Physics-based reductionism can be applied to the whole system or sub-components of the whole system after component-based reductionism has been carried out. In Figure 1, these reductions are encoded as a further set of assumptions. The reductionist philosophical approach has come to dominate scientific and engineering practices over time, and is associated with a Newtonian world-view—Newton himself said:

> "We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances. To this purpose, the philosophers say that Nature does nothing in vain, and more is in vain when less will serve; for Nature is pleased with simplicity and affects not the pomp of *superfluous causes.*" — Isaac Newton, Principia: The Mathematical Principles of Natural Philosophy (Newton, 1686, emphasis added).

Here the idea of avoiding "superfluous causes" and the idea that "Nature is pleased with simplicity" has been taken as an argument for reduction to enable simplification.

Classical mechanics has been built on these beliefs and assumptions, with huge success, and the plethora of law-governed, deterministic, and reduced models are typically defined with a high degree of mathematical rigor. But experience also tells us that much about the physical world still cannot be predicted or modeled to high levels of precision. Reductionist models, by definition, cannot capture the *holistic* behavior of the physical system, and they also struggle to account for *emergent* behaviors.

Roughly, "emergence" refers to "phenomena that arise from and depend on some more basic phenomena yet are simultaneously autonomous from that base" (Bedau and Humphreys, 2008, p. 1). However, as is evident from the ongoing debates in the philosophy of science, we should not presume that a single type of emergence exists, nor that a single definition will always and everywhere suffice[9]. This is also true for the emergence of digital twins, given the myriad uses and phenomena that digital twins can

---

[7] This idea is generally attributed to Laplace (see reprint: Laplace, 2012, first published 1795).

[8] The distinction between complicated and complex systems will be discussed in Section 3.

[9] For instance, we could distinguish "weak" forms of emergence that reflect epistemological limitations from "strong" forms that seek to defend ontological commitments. Or, we could distinguish synchronic from diachronic forms of emergence, based on whether the high-order phenomena or properties emerge over time or not. See the edited collection by Bedau and Humphreys (2008) for further discussion of these debates.

represent and affect. However, a key aspect of emergence that will be important later on, is that complex systems often exhibit properties or behaviors that cannot be fully *explained* or *predicted* by understanding the individual components of the system alone. Often such interacting systems contain intricate hierarchies or interdependencies, and emergence (e.g. self-organization) can happen within a part, or across the entire system (see e.g. Bedau and Humphreys 2008). The existence of such behaviors supports the idea of *holism*, which we can describe as the assumption that systems should be viewed as wholes, not just as collections of parts. In terms of model-making, this approach requires studying (and modeling) the interdependencies and interactions within a system, rather than isolating and analyzing individual components—in contrast to the reductionist approach.

But emergence and holism also place pressure on the notions of determinism and natural laws, as emergent behaviors cannot typically be anticipated or predicted. This raises the question of whether such a barrier is metaphysical or epistemic in nature (that is something about the fundamental nature of the physical system or our knowledge of the system). Let us look at the epistemic issues first.

2.2. The role of knowledge in model making

To start, two distinctions can be drawn between *objectivism* and *subjectivism*, and between *epistemic uncertainty* and *aleatory uncertainty*.

Very briefly, *objectivism* holds that the "model making" process in Figure 1 is an attempt to objectively describe and represent reality. Models, then, are seen as tools that aim to accurately capture the underlying structures, laws, and relationships that exist in the natural world, in a manner that is independent of human perception or interpretation. In contrast, *subjectivism* holds that models are information constructs that reflect or encode human perspectives, interpretations, and choices. As such, the process of observing a physical system is contingent on factors such as the goals of the study.[10] The distinction between aleatory and epistemic uncertainty relates to this first distinction.

On the one hand, *aleatory uncertainty* (sometimes known as "stochastic uncertainty") refers to any uncertainty that arises from inherent randomness or variability in a system (see e.g. Hughes and Hase (2010)). Such uncertainty (or chance) is considered irreducible because it is linked to the natural variability of phenomena. An objectivist, therefore, would take this uncertainty to be real and to exist independently of our knowledge or beliefs.

On the other hand, *epistemic uncertainty* refers to uncertainty that arises from a lack of knowledge or incomplete information about a system. The missing knowledge (or *model error/inadequacy*) can arise due to many factors (e.g. inadequate measurement or sensors, incomplete theories). In the context of subjectivism, uncertainty is, therefore, not just a feature of the external world but crucially depends on the observer's knowledge and perspective. For instance, if we are uncertain about the parameters of a model due to insufficient data (that is observations), this uncertainty is epistemic in nature and could in principle be reduced by gathering more data.

In practice, the distinction between objectivism and subjectivism, and between aleatory/epistemic uncertainty is not always as clear-cut as presented here.[11] Some uncertainties may exhibit characteristics of both aleatory and epistemic uncertainty—but saying more about this is beyond the scope of the current

---

[10] This relates to the idea of *relativism* and also *perspectivism* (Giere, 2019) in the sense that the "truth" (or experience) is affected by and/or not the same for all observers. For example, as was famously shown to be the case in the early 20th century physics—see for example Greiner (1994); Rovelli (2016).

[11] Nor have we said anything of how the the subjectivist or objectivist would handle cases that are common in digital twinning, such as "what-if" forms of modelling and simulation (e.g. scenario testing) where there is no reality, because the digital twin is "representing" some future state. In such cases, the objectivist could argue that although the various future scenarios have not occurred, the underlying laws and principles governing the system (including those latent in data-driven approaches, such as machine learning) are objective and can be applied to predict future states. Insofar as their model captures these laws or principles, therefore, it serves as a formal or empirical extension of known facts or laws. In contrast, the subjectivist could claim that modelling already involves myriad subjective choices, such as the selection of parameters, features or variables, and future (or counterfactual) scenario testing only serves to further exacerbate these sources of uncertainty (e.g. multiple plausible futures or scenarios). Our thanks to an anonymous reviewer for raising this point to allow us to clarify.

discussion. A related issue in both the philosophy of science, but also moral philosophy and the social sciences is the impact of human bias on the myriad value judgments that intersect with science and engineering (e.g. model selection and design choices).

Firstly, human interpretations are problematic, and it is difficult for us to be objective when constructing models and interpreting the results. Humans tend to adopt worldviews that suit them (e.g. techniques or methods that are familiar to a specific community of practice), and we are all subject to cognitive biases (e.g. confirmation bias). In addition, poor research practices can mean that models are not properly validated, calibrated, or tested once they are built, leading to claims that many published research results may in fact be false (Ioannidis, 2005; Marques, 2021). Furthermore, groups and communities are subject to negative group dynamic effects, such as those that may stem from a lack of diversity or inclusivity, combined with entrenchment and groupthink that can exacerbate negative views of other groups and their associated philosophies. For example, those working in the "hard sciences" often fail to understand the approach and values of those working in social sciences or humanities and vice versa (see reprint: Snow, 2012, first published 1964). There are multiple other types of *philosophical tribalism* and dogmatic behavior[12] that can impede the adoption of useful model-making practices. For example, researchers and practitioners are often philosophically aligned to either *quantitative* or *qualitative* methodologies, where in many circumstances, *mixed methods* (e.g. a combination of quantitative and qualitative methodologies, see e.g., Varga 2018) would be more beneficial. This will be an important point for digital twins, where both quantitative and qualitative functions are often required.

So far, these initial philosophical concepts have addressed only the top row of Figure 1. However, the schematic is cyclical in nature, with improvements feeding back in after the interpretation of a model's output(s). Therefore, let us focus on this stage next.

2.3. *Defining a purpose for a model*

In 1982, British statistician George Box published the now-famous adage, *"All models are wrong, some are useful."* Box's central point was that no (statistical) model can ever be "correct" in the sense that there is a "perfect" match with the physical system. But Box's statement also emphasizes the idea of model usefulness (or utility[13]). We can capture the idea that models can have a *useful purpose*, even though they can never be perfect, with the concept of *utility*. But we can also ask, "useful for what?"

An obvious response would be that the primary purpose of a model is to gain (or enhance, extend, and/or clarify) knowledge. Such a view seems to be captured by the idea of "model-dependent realism," expressed by Hawking & Mlodinow's Grand Design (Hawking and Mlodinow, 2010), which claims that the value of models lies in their ability to make accurate predictions and provide useful explanations, not in any claim to absolute truth (that is realism). As Hawking & Mlodinow state:

> "According to model-dependent realism, it is pointless to ask whether a model is real, only if it agrees with observation."—Hawking and Mlodinow (2010) .

This pragmatic perspective has a rich history in both philosophy and philosophy of science[14]. A common emphasis of pragmatic perspectives is how knowledge gained from scientific models ultimately leads to additional explanatory capability or *insight*—noting, again, the cyclical nature of Figure 1.[15]

---

[12] French philosopher, Simone Weil captured this sentiment with the quotation: "The villagers seldom leave the village; many scientists have limited and poorly cultivated minds apart from their specialty"—Weil (1968).

[13] A concept developed in economic and game theoretic models in the 20th Century (Heylighen et al., 2007) and also extended to other contexts including, for example, structural health monitoring (Hughes et al., 2021).

[14] See (Dennett, 1991) for a highly accessible discussion of the central issues that also relates to the topic of emergence.

[15] Once additional knowledge is gained, it can be used (exploited) to create further utility (or practical value), for example by supporting decisions. We will say more about this broader notion of practical value (e.g. as it relates to social or ethical considerations of digital twins) in section 5.2

Both insight and utility depend on a final key concept, *trust.* As Harlow Shapley, the American astronomer said,

> "No one trusts a model except the man who wrote it; everyone trusts an observation except the man who made it."

The first part of this quote acknowledges that models are often, and rightfully, viewed skeptically by those who did not create them. For instance, whereas the creators of a model may be intimately familiar with its strengths and limitations, leading to greater trust or confidence in the model's utility and ability to generate insights, others may be more critical or cautious about accepting its predictions or conclusions due to a lack of familiarity with the assumptions that went into its development. But as the second part of the quotation emphasizes, the creators of the model who actually made the initial observations will also be more aware of potential errors, biases, or limitations in their methodology, and thus may be more cautious about the reliability of their own observations.

With Shapley's quote in mind, if we cycle through the process depicted in Figure 1 we can begin to see how the various assumptions and values we have identified so far would become increasingly concealed, leading to greater uncertainty and potential sources of distrust. It is perfectly possible, for instance, for people to be working on a model for which they did not do any of the original model makings, and therefore be unaware of the philosophical context used in developing the original version of the model, or the encoded assumptions within the model that are inherited by successive generations of practitioners (e.g. a choice to use one "standardized" measurement scale over another). And, in many domains (engineering being one) the separation of practitioners from the model-making process (and the associated assumptions) is increasingly common as modeling becomes more frequently integrated into sophisticated software tools or packages.

To take just one example of how this separation could affect notions such as utility, insight, or trust, consider the choice about how to delineate and demarcate the boundaries of the modeled system when dealing with complex physical systems.

In his book, "Seeing Like a State", James Scott provides a compelling account of how such choices about delineation or demarcation—choices that are inseparable from the process of modeling—can have significant societal impacts. Scott's account deals with non-computational models, and focuses on the emergence of scientific agriculture (Scott, 2020) in 18th and 19th Germany. However, it is still instructive for our paper, because Scott provides evidence of how foresters began to view forests primarily in terms of their commercial timber yield, leading to a model of forests that were orderly, regimented, predictable, and controllable, but also fundamentally mono-cultural and antithetical to the inherent diversity of the real forests. To put it another way, the *utility* of this scientific (but non-computational) model, optimized the production of a single commodity (timber) while ignoring other aspects of the forest's ecosystem, such as undergrowth, soil health and nutrient cycles, wildlife habitats, and impacts on local communities who used trees for additional purposes (e.g. sap)—a myopic set of insights.

Initially, this approach led to increased timber yields and was seen as a success. But, because so much of the real-world (or physical system) fell outside of the scope of the model, over time these "simplified" forests became more vulnerable to diseases, pests, and environmental stresses due to how they were managed (or controlled). In summarizing his example, Scott provides a persuasive argument against modeling that is fundamentally reductionistic:

> "The metaphorical value of this brief account of scientific production forestry is that it illustrates the dangers of *dismembering* an exceptionally complex and poorly understood set of relations and process in order to *isolate* a single element of instrumental value."— (Scott, 2020, p.21, emphasis added).

While some may wish to dismiss such concerns as falling beyond the scope of science, and instead falling within the jurisdiction of ethicists or policy-makers, this would be a naive position to hold.

In contrast, the claim we wish to defend here is that *utility, trust, and insight* are three key *generic* requirements (or properties) of models that should also extend to digital twins[16]. This should not be read as a dismissal of other important scientific values like fidelity, parsimony, cost, or optimality. Rather, we would argue that these characteristics will depend on the *specific context* of the model (or digital twin) (that is in contrast with the previous *generic* requirements). Parsimony, however, requires a further comment as it relates to one of the key principles to be discussed later in Section 4.

Essentially the parsimony principle for models means that a simpler model with fewer parameters is regarded as better than more complex models with more parameters, assuming that both models fit the observations similarly well. However, in recent years, and particularly in research related to living systems, cognitive science, and AI, there is a growing amount of evidence that does not favor parsimony. For example:

> "AI researchers were beginning to suspect—reluctantly, for it violated the scientific canon of parsimony—that intelligence might very well be based on the ability to use large amounts of diverse knowledge in different ways,"—Pamela McCorduck, (McCorduck, 2004).

See also discussions in Marsh and Hau (1996); Huelsenbeck et al. (2008); Hastie et al. (2009) (for example) relating to nonparsimonious models[17].

In summary, this section has explored a range of philosophical topics and concepts related to an idealized process of modeling. This will serve as an important context for the principles we set out in Section 4. However, the discussion has been based on an idealized model, with an emphasis on scientific representation. As such, it is important that we provide a more realistic account of the true complexity of such a process, as well as addressing some of the specifics of engineering practice.

3. Complexity in engineering systems

> "Engineering is the art of modeling materials we do not wholly understand, into shapes, we cannot precisely analyze, to withstand forces we cannot properly assess, in such a way that the public has no reason to suspect the extent of our ignorance"—Dr. A. R. Dykes, from the British Institution of Structural Engineers President's Address, 1978.

By the mid 20th Century, many researchers were identifying that complex behaviors occurred that could not be modeled using the reductionist paradigm. This included a diverse range of applications such as the long-established field of life sciences (see e.g. Weaver (1948)), and new areas like information theory (Shannon, 1948), cybernetics (see e.g. Ashby (1956)), operations research (see e.g. Churchman et al. (1957)), and artificial intelligence (see e.g. Turing (1950)). Within these fields, many researchers were interested in studying emergent behaviors, and so were attracted to holistic philosophies for building models. All of these topics would become large fields of research in their own right, but they also all contributed to the three important current fields of study. The first is *complexity science*[18], which primarily focuses on emergent and adaptive behaviors (see e.g. Waldrop 1993; Mitchell 2009; Jensen 2022). Second is *systems research* which is focused on managing large-scale socio-technical systems (see for example Meadows (2008)) and the related field of *systems engineering*. Lastly is *artificial intelligence* (see Russell and Norvig (2010)), which has been very influential in developing new approaches to learning from data, knowledge modeling, and intelligent agents (amongst many other things).

---

[16] These have some similarity with "purpose, trust and function" from the Gemini Principles (Bolton et al., 2018).

[17] Note also the recent advance in hyperdimensional computing Thomas et al. (2021).

[18] Complexity science (including complex adaptive networks) has primarily been developed through the study of life and information sciences and sociology with an emphasis on the interaction that occurs between "agents" in networks or other interactive frameworks (Waldrop, 1993; Mitchell, 2009; Jensen, 2022). There are a range of emergent behaviours, for example, self-organisation (Gershenson, 2007).

When considering complex system behavior, it is possible to distinguish between different categories of a system based on linear versus nonlinear, ordered versus disordered, deterministic versus non-deterministic[19], reduced versus holistic, etc, and combinations of these categories. Here we will adopt the broad distinction that *complex* relates to a system that can have emergent behavior (i.e. "more is different") whereas *complicated* relates to a system that is not "simple" but does not have interacting components that could lead to emergent behaviors.[20] These distinctions will be important when we discuss the principles related to digital twins in Section 4. We also note here that complexity techniques are already being promoted for digital twins of cities see for example Rozenblat and Fernández-Villacanas 2023; Caldarelli et al. 2023.

3.1. Types of complexity in engineering systems

Engineers are expected to design, build, commission, operate, maintain, manage, and decommission a huge range of different systems. The quote from A. R. Dykes at the start of this section gives a sense of the engineering process. Multiple categories of complex and uncertain factors (in this case materials, shapes, forces, and public expectations) need to be brought together to achieve the required task. Table 1 lists some of the types of complex (and/or complicated) phenomena that can arise in, or influence, physical systems.

It is typical for engineering applications to have multiple types of complexity contained within them from the list in Table 1. For example, geometric complexity and joints are used extensively in a wide range of manufactured products, as are sophisticated materials, such as composites. These different aspects of the manufactured product are often designed, modeled, and tested separately before being integrated into the final version of the product.

As the format of Table 1 Indicates, our usual method for dealing with mixed complexity is to separate it and consider each type independently. Usually, this is mapped onto our siloed (e.g. reduced) set of divisions within subject areas (and education system). Roles and specialisms are also then aligned with these divisions, creating teams of experts in each separate topic area.

Furthermore, unlike scientific inquiry, where the focus is on understanding and explaining the behavior we observe (as in complexity science), engineering is often required to create something new or deal with a socio-technical system that is highly complex/uncertain and is changing over time[21]. In order to try to address some of the related challenges, the field of *systems engineering* has developed some methodologies based on a holistic approach, which we discuss next.

3.2. Systems engineering

Systems engineering was developed during the 20th Century alongside the related other fields of systems research and complexity already described above[22] (Schlager, 1956). The underlying philosophy of systems engineering is that of holism, and the field has now matured into an established methodology for managing complex engineering projects (see e.g. Walden et al. 2015; Hirshorn et al. 2017). At the heart of current-day systems engineering is the role of *processes*, to enable the design, implementation, and management of the engineering application or project.

Systems engineering processes have evolved from being document-based to being *model-based* (Estefan et al., 2007), as technologies have improved to allow information to be captured with more automation and presented graphically. This approach underpins the diagrammatic approach to *enterprise architecture* (Dandashi et al., 2006) and could be regarded as a predecessor to digital twinning. Indeed, digital

---

[19] Note that the use of non-deterministic can be interpreted in different ways by different communities. There is also *indeterminism* which is generally taken to mean non-causal e.g. not caused in a deterministic way.

[20] The interested reader is directed to discussion in Ladyman et al. (2013), Van Beurden et al. (2013) and Grieves and Vickers (2017) for further comments on the notion of complexity.

[21] Added to which there is also the complexity of cooperation (Axelrod, 1997).

[22] It has also incorporated multiple other influences that we have not described, most notably aspects of management research.

Table 1. *Examples of complex (and/or complicated) phenomena that can influence physical systems*

Type	Examples (not an exhaustive list)
Environmental	Temperature, pressure humidity & climatic effects; physical location
Geometric	Multiple components of varied shape & geometries; joints
Material	The physical & chemical properties of matter; wear, ageing & damage
Behavioural	Mechanistic behavior of solids & fluids; vibrations & time-dependence
Operational	Control & feedback; updates; faults; failures; connectivity
Computational	Time & memory requirements; processing resources; data sizes
Processes	Design; decisions & interventions; sequencing & workflow; behaviour
Organizational	Structure & hierarchies; practices & organisation culture; rewards & incentives
Social	Attitudes; motivations; culture; education level; religion; beliefs; gender etc.

twins that enable planning and design may be considered examples of model-based systems engineering as they facilitate the exchange of information, alignment of design, and management of programmatic complexity in the same way as now-traditional systems engineering documentation processes do.

The ethos of systems engineering is to give a framework that enables multiple uncertainties and complexities to be managed simultaneously, and for the technical processes to be aligned with the decision, management, and wider related business processes. It is important to make a clear distinction between working with "engineered systems" and the practice of engineering in complex systems. Confusingly, both can be called systems engineering, but the key distinction is that engineered systems can be controlled/optimized whereas complex systems typically cannot.

The systems engineering community has given a considerable amount of time and thought to the philosophical and pragmatic frameworks needed to deal with complex/complicated engineering applications. For example, in recent papers (Watson, 2019; Watson et al., 2019) 3 hypotheses for systems engineering were articulated:

H1. If a solution exists for a specific context, then there exists at least one ideal systems engineering solution for that specific context.

H2. System complexity is greater than or equal to the ideal system complexity necessary to fulfil all system outputs.

H3. Key stakeholders' preferences can be represented mathematically.

We shall discuss these hypotheses further in Section 4.2, in the context of digital twins. But to mention just briefly, H1 relates to the concepts of *existence & uniqueness.* H2 is related to the idea of *counter-parsimony*, by which we mean choosing *not* the simplest model that fits the data, but the model with sufficient complexity[23]. And, H3 is anticipating the stakeholders preference for quantitative solutions.

Other important concepts that are emphasized in systems engineering are the idea of the *lifecycle* of a system, requirements analysis, and hierarchies of systems that lead to systems-of-systems (see e.g Adams and Meyers (2011))[24].

Although the subject borrows from and integrates, several of the concepts and methodologies from systems research and complexity science, it should be noted that some researchers have been critical of the systems theory ethos. For example, Micheal Grieves expresses reservations about treating everything as a process;

---

[23] This has an interesting connection to the concept of *requisite complexity* in cybernetic systems—see Beer (1985).

[24] Note there is also the extension of hierarchies to more complex structures such as holarchies—see for example Calabrese et al. (2010); Cardin et al. (2018) and references therein.

"We like to think that what we do in our organizations is process. Under systems theory, the process is a deterministic way of linking inputs to outputs. In a systems view of the world, we have inputs, processes, and outputs. For any given set of inputs, we get a well-specified and consistent set of outputs. It is all very neat and well defined." (Grieves, 2005, page 19).

Grieves argues instead that not everything can be made a deterministic process, and that engineers need to make extensive use of *practices* as well, with results that lead to *satisficing*[25] instead of optimization (Grieves, 2005).

The broader point is that engineering contains some form of "art" (alluded to in the A. R. Dykes quote above) typically encoded in the form of attributes like *engineering judgments* and *design choices*[26]. As much as many practitioners would like, these creative activities cannot be entirely turned into repeatable processes. It is interesting to note that some in the social science community, who have adapted systems thinking, have extended the concepts to include *dialogue* and create an *architecture of evolution*—see for example Christakis (2006)[27].

Using more philosophical arguments, Weinbaum (2015) describes systems theories as based on a "black box dogma" with unresolved clarity on issues relating to the role of feedback, evolutionary adaption, and causality.

In response to the criticisms, it is certainly true that the systems engineering approach favors defining multiple processes with associated inputs and outputs, and that in itself could be an over-constraining structural format for some applications. It is also true that the role of reductionism and deterministic modeling was strongly used in some of the early systems research fields, and some of that thinking has been inherited by the modern version of the field. Finally, creative activities cannot always be turned into processes, and we should recognize that.[28]

As pragmatists, engineers often have little concern for this type of philosophical subtlety, but it should be borne in mind when these approaches are used in digital twins. Despite the limitations, systems engineering offers some useful tools for constructing digital twins, and the connections have already begun to be discussed in the literature—e.g. by Heber and Groll (2017); Schluse et al. (2018); Madni et al. (2019); Jinzhi et al. (2022); Michael et al. (2022); Olsson and Axelsson (2023). However, we are seeking to simulate emergent behaviors, of the kind discussed in Section 2, which we will now return to.

3.3. Emergent behaviours

In the context of digital twins, the basic idea is to join components together to *reconstruct* the dynamic behavior of the combined system. The simplest case is joining two components.

In engineering, we make extensive use of numerical simulation tools that essentially reduce (e.g. break up) complex geometries and behaviors into an assemblage of simpler elements for which the behavior can be defined. These techniques, such as the finite element method, have evolved into sophisticated tools that are widely used to simulate the behavior of complex/complicated systems that cannot be captured using simpler modeling techniques (see e.g. Crisfield 1997). The outputs from element-based methods are, in fact, emergent behaviors. This usually relates to field quantities such as stress, displacement, flow rate, or temperature, which are approximated as a form of "self-organization" between the elements, acting within

---

[25] A decision-making strategy (or heuristic) in which an agent selects the first option that meets some pre-defined criteria or threshold, regardless of whether it is the optimal one. See also regret minimisation.

[26] There is a similar quotation from Ove Arup, who is quoted to have said that: "Engineering problems are under-defined, there are many solutions, good, bad and indifferent. The art is to arrive at a good solution. This is a creative activity, involving imagination, intuition and deliberate choice."

[27] Note that we are not distinguishing here between the different facets of systems theory such as "hard" and "soft" systems thinking. For a more in-depth discussion of these topics, the interested reader can find more information in Checkland (1999).

[28] In fact, creative process such as design have been explored using qualitative research methodologies, such as *activity theory*, see e.g. Barthelmess and Anderson (2002); Cash et al. (2015); Lu et al. (2018).

the overall element-based model. Essentially, the overall behavior arises from local interactions between the multiple elements.

In addition to self-organization, there are other types of emergent behavior, and multiple authors have described how the various types might be categorized—see for example Ashby (1956); Holland (2007); Frei and Serugendo (2012); Fernández et al. (2014); Holland (2018); Tadić (2019); Jensen (2022) and references therein. Broadly speaking, the types of emergent behaviors range from relatively simple types, such as self-organization and synchronization, (Jensen, 2022), through to *evolutionary* forms of emergence (Kauffman, 2000). The ability to make predictions for emergent behaviors is a significant capability that is seen as a very desirable functionality (Gershenson, 2013), including for digital twins. We will return to discuss how digital twins might be expected to produce such behaviors, especially for very complicated applications in Section 4.3. Next, we consider the role of artificial intelligence for digital twins.

3.4. *Artificial intelligence*

The quest for artificial intelligence (AI) (as described, for example, by Nilsson (2009)) is multi-faceted, and has been driven by several different motivations. Those motivations include inspiration from human intelligence and other biological examples, the desire to create intelligent machines, and the application of AI to solve complex applied problems. There are multiple other facets, implementations, and deployments of AI, which we leave to the interested reader to explore—see for example Minsky (1988); Nilsson (2009); Russell and Norvig (2010); Haenlein and Kaplan (2019); Marcus (2020) and references therein.

Russell and Norvig (2010) use the unifying theme of *intelligent agents* in their comprehensive textbook on artificial intelligence. A current topic of interest is deep reinforcement learning, where agents are used (for example) to solve *sequential decision-making problems*, such as autonomous driving vehicles (see e.g. Kiran et al. (2021)). Sequential decision-making problems are also highly relevant to digital twins, which by their nature are time evolving, and are also required to support a sequence of decision-making tasks.

Importantly for the digital twin paradigm, the AI work on agent-based methods has enabled more sophisticated multi-agent methods than previously developed either in complexity science or systems engineering (although there is now some cross-over between these topics (Vrabič et al., 2021). For example, techniques such as multi-agent reinforcement learning where the agents take actions and receive feedback in a highly adaptive manner (Graesser and Keng, 2019; Kiran et al., 2021).

In very general terms, it could be said that symbolic AI (such as logical reasoning) was the earliest to mature, but despite the success of some aspects, such as expert systems (Krishnamoorthy and Rajeev, 2018), it has more recently been overtaken by sub-symbolic AI which has become the dominant force in AI in recent years, particularly deep learning (LeCun et al., 2015; Goodfellow et al., 2016) and most recently large language models (Teubner et al., 2023). In the past few years, some AI experts have been pointing out the limitations of connectionism, (Marcus, 2018), and there is a revised interest in the possibility of combining the two approaches in the form of neurosymbolic AI[29] (Belle, 2022).

For the purposes of our discussion, we note the following points regarding AI for digital twins. Firstly, both learning and reasoning are highly desirable functions that we often want to build into our digital twin applications, meaning that AI techniques are very important in this respect. In addition, digital twins can be viewed as a *method of deployment* for AI and it is associated techniques[30]. There are multiple examples of this type of deployment—see for example DebRoy et al. (2017); Farhat et al. (2020); Kapteyn et al. (2020); Ritto and Rochinha (2021); Tripura et al. (2023); Siyaev et al. (2023)—and this is a topic we will return to later on. Finally, just like digital twins, AI still has no formally agreed overarching definition. In

---

[29] This is a overly simplistic summary, but readers who are interested can find more detail in the associated references. See also recent work on hyperdimensional computing Thomas et al. (2021).

[30] Digital twins can be considered as a "method of deployment" of other technologies too. More broadly digital twins are also a "method of integration" for a range of technologies and also a form of "technological convergence" Bainbridge and Roco (2016).

large part, this is because of the *philosophical breadth* of the topic—something which hopefully is described by the preceding discussion[31].

3.5. Other methods

Lastly in this section, we would like to mention that there are multiple other communities of researchers and practitioners that have developed sophisticated methods for modeling highly complex and uncertain applications. Some overlap with AI and other fields mentioned above, and others have developed their own areas of endeavor. For example (with just a few selected references) dynamical systems theory (Kuznetsov, 2004; Strogatz, 2019), data assimilation (Evensen et al., 2009; Kutz, 2013), Bayesian statistics (Barber, 2012; Särkkä, 2013; Gelman et al., 2014; Kruschke, 2014), data mining (Hastie et al., 2009; Han and Kamber, 2022), game theory (Jones, 2000), ensemble modeling (Zhou, 2019), spatiotemporal modeling (Banerjee et al., 2014), agent-based modeling (Abar et al., 2017; Zhang et al., 2021b), statistical relational learning (Getoor and Taskar, 2007; Belle, 2022), asymptotic theory (Van der Vaart, 2000), time series analysis (Hamilton, 2020), adaptive & nonlinear control (Åström and Wittenmark, 1995; Fradkov et al., 1999; Barlow, 2002; Wagg and Neild, 2015), information theory (MacKay, 2003), network science, (Baker, 2013), and optimization methods (Boyd and Vandenberghe, 2004) to name just a few.

4. Towards a philosophical framework for digital twins

> "It ought to be remembered that there is nothing more difficult to take in hand, more perilous to conduct, or more uncertain in its success than to take the lead in the introduction of a new order of things."—Niccol Machiavelli, The Prince, 1532.

As we discussed in Section 1, the ambitions for digital twins are set very high across a very wide spectrum of possible applications. In practice, we need to manage these high expectations and make clear what are the possibilities and limitations to using digital twins. To this end, in this section we develop the foundations for a *philosophical framework* within which we can build, evaluate, and better understand specific instances of digital twins.

As noted by Machiavelli, introducing something new is fraught with potential difficulties, and we argue that a firm philosophical foundation is an essential part of the process. However, it is important to note that we are not the first to attempt this goal. For instance, Korenhof et al. (2021) reviewed and critically analyzed the dominant conceptualizations of digital twins in the academic literature. In doing so, they raise the question, "If a digital twin is expected to actively intervene in a physical entity, is it really only a representation?." Their answer is that DTs should be treated as "steering representations" that are used to "direct a physical entity towards certain goals by means of multiple representations." Their proposal has considerable merit, and should likely figure (in some form) in a fully articulated and developed philosophical framework—one that can also be used to support ethical reasoning and decision-making about the societal impacts of digital twins. However, we do not wish to take this argument as our starting point without first considering the fundamental *purpose* of a digital twin ourselves. In Section 2.3 we argued that *utility, trust, and insight* are the three key *generic* properties we want for digital twins. These three characteristics form the basis of our characterization of *purpose*.

Specifically, we take *utility* to mean context-specific usefulness that relates to the application at hand, and is expressed as a set of *functional requirements* within the *contextual setting* that the digital twin operates—here, the contextual setting relates to the specific properties of the physical twin, such as its geometry, materials, the environment in which it is located or deployed and so forth. The functional

---

[31] It is also due to the *universality* and *versatility* of the digital twin concept.

requirements could be, for example, to support decisions, to learn patterns of behavior, or to develop more efficient ways of operation.

The attribute of (unbiased) *trust* is related to the *uncertainty* within the digital twin and is also connected to security, openness, and quality (Bolton et al., 2018). Trust is therefore essential for supporting the functional requirements of the digital twin. Lastly, the role of *insight* is related to *knowledge*, but not just lists of facts, insight relates to *enhanced understanding* of the physical twin within the contextual setting. The insight(s) gained from the use of a digital twin could be some *measurable* improvement in understanding the behavior of the physical twin, or the learning acquired via the successful completion of a sequential decision-making problem(s) over time (such as mentioned in Section 3.4.)

Since the concept of digital twins was first suggested there has been lots of discussion and debate over what exactly the *definition* of a digital twins actually is—see for example Negri et al. (2017); Miller et al. (2018); Wright and Davidson (2020); Wagg et al. (2020); Arthur et al. (2020). This is natural when the idea is new, but can be unhelpful to the overall debate at times. Therefore, in an attempt to give some additional clarity about digital twins, but without getting overly restricted by a technical definition (at least for now), a set of principles for digital twins is proposed here, based in part on the discussion above. While these principles fall short of establishing a full philosophical framework, they are anchored in the philosophical concepts discussed in the previous two sections and are set out in three categories: (a) *what digital twins are*, (b) *what properties they should have*, and (c) *what they should enable*.

We begin with what digital twins are. Digital twins are:

1. **Holistic** in nature, but may use reductionist ideas when appropriate. e.g. both the whole *and* the parts are considered important, in order to capture any *heterogeneity*;
2. **Purpose driven** where the clearly articulated *useful purpose* (or set of purposes) is underpinned by a set of *functional requirements*;
3. **Time evolving** dynamic systems that can reflect changes in the physical twin that occur over time via *updating* and *evolution* of the digital twin;
4. **Context-specific** representations that are bespoke to an individual physical twin, and which can be both artefacts (objects) and/or processes within the *contextual setting*;
5. **Counter-parsimonious**, meaning not seeking simplicity for its own sake, but instead aiming to reflect the *required level of complexity*—although they may make use of parsimonious concepts, when appropriate;
6. **Reconstructivist**, meaning they aim to reconstruct (some or all of) the behavior of a physical twin by assembling the components of the digital twin, including *emergent behaviors*; and
7. **Biased**, due to the philosophical worldviews of the communities that constructed them, but able to acknowledge the limitations that this brings.

Digital twins should have:

8. A set of **components**, which can include agents, models, networks, data sets, and other digital objects;
9. Access to **real-world data**, recorded/streamed from the physical twin, or its surrounding environment.
10. A means of **dynamic assembly**, so that the components can be connected, or otherwise integrated together in a time dependant way;
11. An **operational platform**, consisting of software, hardware and network infrastructure, including a user interface, data storage and other computational resources;
12. A method for representing and updating **knowledge** that is shared between the users and the digital twin;
13. A time-dependent **connectivity** to the physical twin, usually via an internet-of-things (IoT) network or similar, so that data, control, and other signals can be exchanged between the twins; and

14. An **integration architecture** that enables components and/or other parts of the digital twin to interoperate and/or federate with each other, and in some cases entire other digital twins.

Digital twins should enable:

15. **Outputs** to be produced that relate to observed quantities of interest (QoIs) in the physical twin and to the functional requirements;
16. **Trust** in the outputs to be expressed through processes of assurance, including validation, verification, and/or error detection and correction[32] in order to account for relevant forms of uncertainty[33];
17. **Inheritance** of (at least) some of the properties of the components within the digital twin (e.g. object-property inheritance, described below);
18. **Interaction**, such that the components are able to interact with the aim of reconstructing emergent behavior (s);
19. **Learning** both from data (e.g. QoIs and outputs), and more broadly from the deployment of advanced techniques such as those from AI, statistics, dynamical systems, etc;
20. **Insights** to be obtained that serve the purpose of the user, and maximize explainability and interpretability of the outputs; and
21. **Exploitation** of the insights to provide **value** (e.g. improved decisions, efficiency gains, etc.) and/or enable real-world **actions** to be taken such as control/scheduling actions for the physical twin.

These 21 principles incorporate the key attributes of a digital twin and capture the holism of systems engineering, emergent behaviors from complexity science, uncertainty analysis from statistics, time-evolution from dynamical systems theory, techniques from AI, control actions, and decision theories—amongst other things. We believe that such a set of principles is sufficiently versatile and universal to fit a wide range of digital twin applications, across multiple domains, whilst still capturing some of the most important specific aspects of digital twins.

However, to help justify this belief, we now consider how these philosophical principles can be applied to explain some common questions relating to digital twins.

### 4.1. Why is a digital twin, not a model?

We will offer more than a single answer to this particular question, all of which can coexist with each other. The first is shown in Figure 2 and relates to the connectivity of the physical twin and the digital object. Kritzinger et al. (2018) make the following distinctions between three concepts which are shown schematically in Figure 2:

1. Digital model—no connection between virtual and physical (Figure 2a). This is the "traditional" approach to modeling in science and engineering.
2. Digital shadow—data received from a connection (e.g. over an IoT network) with the physical twin is used to update and "shadow" the state of the physical twin (Figure 2b). In this way, the digital shadow will evolve over time to reflect changes that occur in the physical twin.
3. Digital twin—as for the digital shadow, but with the addition of control actions, or interventions (in the case of a system that cannot be directly controlled) being given over the network to the physical twin domain (Figure 2c)

---

[32] See, for example, Chapter II of MacKay (2003).

[33] Note that this form of trust depends on the inherent *trustworthiness* of the outputs. Trust without trustworthiness is misplaced. Another way of stating this, therefore, is to say that digital twins should enable *justified trust*.
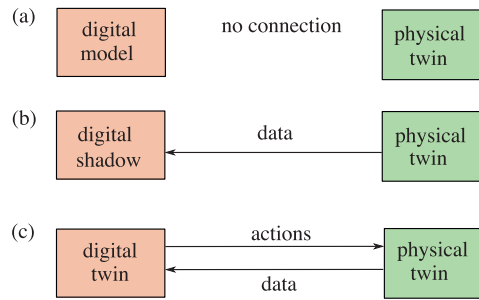
Figure 2. *Schematic diagram of a common way to interpret the distinction between a digital model and a digital twin, showing (a) a digital model, (b) a digital shadow, and (c) a digital twin. This is a common interpretation found within the literature and helps explain why a digital twin that is "bidirectionally connected" with a physical is best seen as a broader "cyber-physical system" rather than two separate components (see for example Kritzinger et al., 2018).*

The 21 principles set out above relate to *digital twins*, but digital models and shadows could also be represented by selecting fewer principles to apply.

However, the model/shadow/twin explanation does not capture some aspects that we have discussed above relating to digital twins. Critics can point out that using existing terminology, Figure 2a shows a model. Figure 2b shows an updated model, and Figure 2c shows a control system. For example, the explanation given in relation to Figure 2 has little or no sense of timing or mechanisms. For example, when does the digital become connected to the physical? Is the data transfer to the shadow continuous or intermittent? Are the actions taken part of the digital twin or something separate. Another criticism is that Figure 2 does not show (or even anticipate) connections between digital twins, via federation.

Furthermore, it is difficult to understand the ideas of holism, or emergent behavior with the model/shadow/twin explanation. So, we believe it is useful to also suggest an additional explanation that can complement the rationale of Figure 2. This additional explanation relates to the use of models in digital twins, as we have described in this paper (e.g. as a combination of multiple digital objects). As a result, digital twins will have the property of *object-property inheritance*. Therefore, digital twins include models among their components, such that digital twins are more than just models (and models are not digital twins). In other words, a digital twin is something more than a model but can be used to perform functions that have been previously carried out using models, because it inherits the properties of the model. In general, object-property inheritance relates to all the components within a digital twin and will be explained in further detail in the next section.

4.2. What previously unseen results can we expect from a digital twin?

"It is the mark of an educated mind to rest satisfied with the degree of precision which the nature of the subject admits and not to seek exactness where only an approximation is possible"—Aristotle (384 BC–322 BC).

It will be fundamental to the purpose of a digital twin to establish whether the digital twin can *produce an output that suits our particular purpose(s).* As the quote from Aristotle reminds us, every output from a digital twin will most likely include (multiple) approximations, and we should *be wary of seeking exactness beyond that which is possible.* The "degree of precision," as Aristotle puts it, relates to the fidelity of an output. However, before any attempt to assess fidelity can be made, we need to consider if a viable output for our particular purpose is possible.

Grieves and Vickers (2017) have considered how the outputs from digital twins could be used to anticipate the types of emergent behaviors that may arise. They proposed a categorization of outcomes for
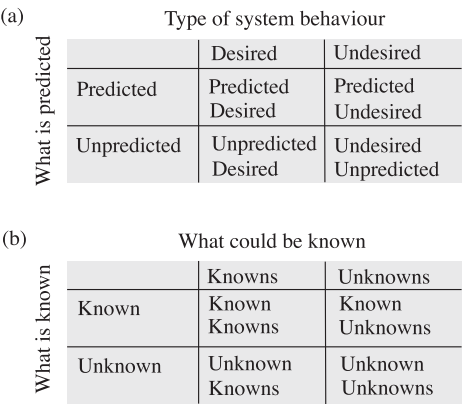
Figure 3. *Schematic diagram showing (a) how the outputs from a digital twin might be able to predict emergent behaviors proposed by Grieves and Vickers (2017), and (b) the "Rumsfeld" matrix.*

the digital twin that is shown in Fig. 3(a). Here there are four categories of outcome that depend on what the digital twin predicts and whether the predicted behavior was desirable in a design context (meaning the intended design) or undesirable (problematic and/or unwanted designs). This framework is then used iteratively to try and minimize the undesirable and unpredicted aspects as much as possible[34].

However, this approach also suffers from the problem of the need to know in advance what to include in the digital twin to get the desired outcome. As pointed out by Kauffman (2000) for example, this is a particular problem in the field of emergent behavior. In fact, problems relating to prior knowledge are well known in other fields, such as the domain of uncertainty and risk management (Okashah and Goldwater, 1994; Lanza, 2000). The "Rumsfeld" matrix captures the key issue as shown in Fig. 3(b).[35]

In the Rumsfeld matrix, we create four categories based on what *is* known (e.g. meaning what we know at this present moment) and what *could be* known (e.g. all possible knowledge, if we had a way to access it). It should be clear that if we do not know something at the present moment, then we cannot include it in our digital twin, and therefore we can never access the "unknown unknowns" category[36]. Knowing in advance, for example by prescribing a specific solution space, is a practical necessity for modeling, but will exclude the more advanced behaviors, particularly evolutionary forms of emergence—see for example Tononi et al. (2016); Kauffman (2000) and references therein.

To take one example, emergent behaviors are often modeled using multiple *agents* that interact with each other according to a predefined set of "rules", typically relating to the environment and their nearest neighboring agents (Jensen, 2022). The idea has already been explored in a digital twin context by several authors—see for example Croatti et al. (2020); Zheng et al. (2020); Vrabič et al. (2021); Clemen et al. (2021); dos Santos et al. (2022). So, although the emergent behaviors are not necessarily known in advance, the rules for the agents have to be prescribed in advance, and so the rules are therefore *known knowns.* The emergence will be a product of the prescribed rules (as was the case for Deepmind AlphaGo algorithm (Silver et al., 2016; Chouard, 2016)), and so if we have never observed a particular type of interaction before, it cannot be included in the digital twin. It also will not be in any of our previously recorded data sets, or associated data-based models.

With this in mind, let us consider what can be reasonably expected from digital twins in terms of emergent behaviors. Object-property inheritance can be interpreted as both related to *individual*

---

[34] Another variation would be to replace the desirable and undesirable with "authentic" and "spurious" to try to capture when the digital twin succeeds or fails to give a valid output.

[35] Made famous by Donald Rumsfeld in 2002, this is an adaptation of the Johari window.

[36] This is the category which is associated with *black swans* Taleb (2007); Aven (2013).
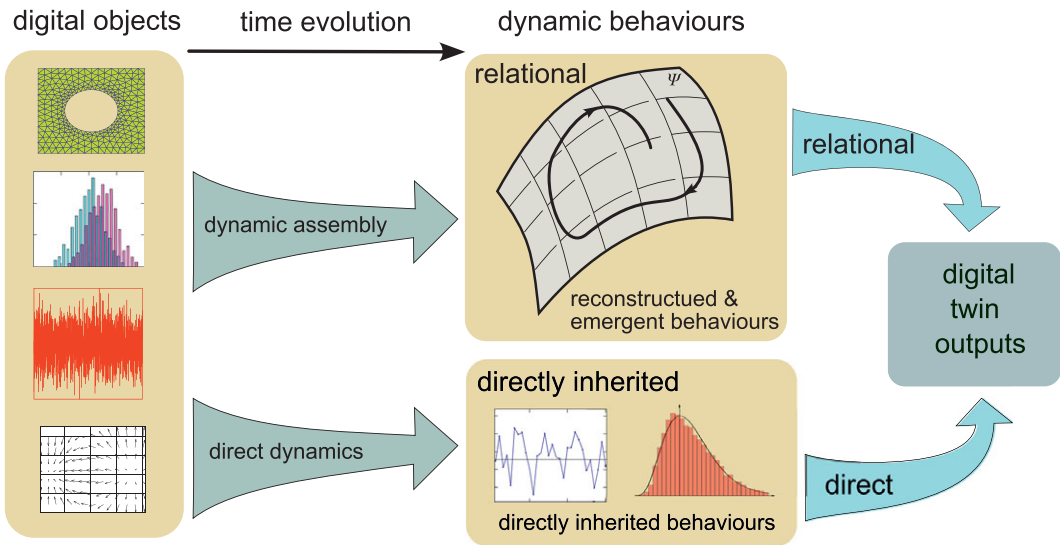
Figure 4. *Schematic diagram showing how the outputs from a digital twin might be created using a series of digital objects (e.g. the components of the digital twin). The directly inherited properties come from each of the components are grouped together. The relational properties, such as any reconstructed or emergent behaviors, come from the process of dynamic assembly. Both the directly inherited and relational properties can be used to form digital twin outputs.*

components (objects) in the digital twin, and *relational combinations* of the components.[37] The relational combinations of the components are achieved using what we can call "dynamic assembly"—an example of which is described in the next section. Therefore, if a digital twin consists of n objects it would be expected to have a number (say d) of *directly inherited properties* that come from the n objects without any interaction between them. In addition, the digital twin would have a combinatoric number (say r) of *relational properties*, including any emergent behaviors, which are generated from the dynamic assembly. Note also that the combinatoric metric will depend on the specific context of the digital twin.

A simplified schematic example for a series of digital objects (e.g. components) is shown in Figure 4, where dynamic assembly methods are used to obtain interactions between the components. In Figure 4, the directly inherited properties are shown to come from the components, and relational properties come from the dynamic assembly of the components. Both direct and relational properties can be then used as digital twin outputs.

It is important to emphasize that all the emergent (and non-emergent) behaviors observed in digital twin outputs are contained in the categories of *known knowns*, *known unknowns*, and *unknown knowns*, shown in Figure 3b. The *unknown unknowns*, shown in Figure 3b *are not accessible to the digital twin by definition*, and could only be known by the addition of new information not known at the current time.

As a result, assuming that the known knowns category is already well understood, it is the known unknowns, and particularly the unknown knowns categories where value can be obtained from using digital twins. Note that we would expect to see more previously unseen results from an ecosystem of connected digital twins (Nativi et al., 2021). This is simply because of the nature of systems—the more connections there are, the more potential there is for emergence. We now consider an example of dynamic assembly.

---

[37] For the purposes of this paper we use the labels "objects" and "components" to mean the same thing.

4.3. How can emergent behaviors be predicted using a digital twin?

Emergent behavior can be reconstructed via interaction. This can be achieved using certain components in digital twins (e.g. models, agents, etc.) which can be dynamically assembled (e.g. joined together) as was shown schematically in Fig. 4.

We, therefore, consider a digital twin to be made up of a series of digital components that will be combined in such a way that they can *reconstruct* the time-evolving behavior of the physical twin[38]. Dynamic assembly is interpreted here based on the idea of creating "connectors" such that the resulting connections lead to interactions between the components with the aim of reconstructing emergent behavior. More specifically we define dynamic assembly as *a method for connecting the components together such that the subsequent interactions enable emergent behaviors, including time-dependent* (e.g. *dynamic*) behaviors.

We are using the terminology *dynamic assembly* to include a range of methods via which DT components might be able to interact. These include, but are not limited to:

- Coupling of system equations describing the behavior of the DT component. E.g. two separate models are "coupled" via the states and parameters and/or boundary/initial conditions.
- Synchronisation of all or part of the states of two DT components. E.g. in a similar way to methods such as hybrid testing, where synchronization is used to achieve a joining effect—see for example Gonzalez-Buelga et al. (2005).
- Organisation of DT components. E.g. into networks and/or hierarchies or more complex structures such as holarchies—see for example Calabrese et al. (2010); Cardin et al. (2018).
- Interoperation via a bespoke protocol to enable communication between DT components that are not using standardized formats. For example, this is often the case when legacy systems are present.

As an example of dynamic assembly, consider the case where it is desired to create a digital twin of the transport systems in a city. Ideally, we would include all modes of transport, but for argument's sake, let us assume we had separate models of the road, rail, and pedestrian traffic in a specific area of the city. Running the models separately misses the interactions between transport modes, so we might want to create a method for dynamically assembling the road, rail, and pedestrian traffic models together[39]. In order to do this, we create a "dynamically assembled" system, where information is passed between the three simulations to represent the interactions between the transport modes. As a result of the interactions between models, they are (i) more likely to represent the real world, and (ii) made lead to emergent behaviors—see for example discussions in Ambra et al. (2019); Busse et al. (2021); Wolf et al. (2022); Jafari et al. (2023).

4.4. How can we assess the existence and uniqueness of digital twin outputs?

As we said above it will be fundamental to the purpose of a digital twin that some type of output exists that is relevant to the context of the physical twin.[40] One example of an output is to choose a quantity of interest. In the study of differential equations, an important underlying concept is the idea of the *existence and uniqueness* of a solution to the problem (Hirsch and Smale, 1974; Guckenheimer and Holmes, 1983; King et al., 2003). The concept asks the questions (1) does a solution exist?, and (2) if it does, is it a unique solution? If the solution is *nonunique*, then other solutions will exist that also satisfy the same defined problem[41].

---

[38] Although the reconstructed outputs still need to be validated.

[39] To keep the argument simple, let us ignore the other aspects of digital twinning, such as data flow and connectivity.

[40] Note the similarity between the hypothesis H1 listed above in Section 3.2 and the questions of existence and uniqueness. Firstly H1 is prefaced with "if a solution exists", which is an assumption of existence. Furthermore, H1 contains the phrase "at least one", which we interpret as the possibility that there may be many solutions, and therefore allowing for *nonuniqueness.*

[41] Note that in the study of differential equations, even though there are many examples where solutions are nonunique, nonuniqueness is generally seen as a situation to be avoided.

Although the idea of existence and uniqueness is typically applied in a deterministic worldview, in the absence of a developed theory for digital twins, we consider how questions (1) & (2) could be applied to the case of digital twins in general. To widen the application beyond the deterministic realm instead of "solution" (that typically implies a precise answer to a specific set of equation(s)), we will instead take the idea of an *output* from the digital twin.

In practical terms, there would appear to be two potential approaches (and at least one caveat) to determining the existence and uniqueness of digital twin outputs. The first approach is to rely on the object-property inheritance of the digital twin, so that if the underlying objects (components) in the digital twin have the property of existence and uniqueness, then the digital twin can also inherit those properties (under some defined conditions). For example, if the digital twin has an ordinary differential equation (ODE) as one component, and that ODE has solutions that exist and are unique, then the digital twin can also inherit those properties—see for example Han et al. (2022); Area et al. (2022). The caveat is that the philosophical framework for differential equations is (almost always) deterministic, and so this will act as a limiting factor with this approach.

The second possible approach (either separately or in combination with the above) is to consider the behavior of the interconnections between components in the digital twin. It might be possible that the existence and uniqueness of digital twin outputs (e.g. the *reconstructed* behaviors) could be assessed using the information at the interface of components. Further work is needed to develop a more formal analysis relating to the existence of digital twin outputs.

Now turning to the question of uniqueness, it is perhaps obvious to state that digital twin outputs may or may not be unique. Nonuniqueness could be a major problem for digital twin users if they are *expecting* (or assuming) a unique output but do not obtain one. However, the precise nature of what is meant by the uniqueness of an output will depend on the context and components that make up the digital twin.

Finally, we note that nonuniqueness relates to a broader issue of *spurious solutions* and related problems such as missing solutions, and false emergent behaviors—this could be considered to be a *failure mode* of the digital twin. We will not consider these problems explicitly here, but we would need to consider the possibility of these outcomes when building a digital twin—see discussion in Grieves and Vickers (2017).

5. Conclusions and future directions for research

In this paper, we have explored and discussed key philosophical concepts that apply to the concept of digital twins—particularly, those that underpin the series of principles presented in Section 4, which could be used as the building blocks for a more complete philosophical framework. This discussion also enabled us to consider how the philosophical context could help define a purpose for a model, and it was concluded that *utility, trust, and insight* are the three key *generic* requirements of models that we wanted to extend to digital twins.

A key part of the digital twinning philosophy is representing complicated/complex systems. This was discussed in detail in Section 3, where we considered the limitations of traditional reductionist methodologies. We then discussed how systems engineering and complexity science had been used to attempt to overcome these limitations by adopting a more holistic worldview. In particular, we discussed the importance of modeling emergent behaviors, that cannot be captured in a reductionist paradigm. Importantly, it is *interactions* that lead to emergent behaviors, and these have to occur dynamically—depending on the exact context, we note that the environment might also influence the emergent behavior.

In Section 4 of the paper, we presented 21 principles set out in three categories; what digital twins are, what properties they should have, and what they should enable. We then used the 21 principles to consider some common questions that arise regarding digital twins. Namely, the questions were: Why is a digital twin not a model? What previously unseen results can we expect from a digital twin? How can emergent behaviors be simulated using a digital twin? How can we assess the existence and uniqueness of digital twin outputs? We do not claim to have provided definitive answers to these questions, rather we have used

the philosophical principles to frame the questions in a way that might help provide more insight and understanding of the questions and the associated topics they relate to.

In concluding this paper, we draw together some further comments and open research questions.

5.1. Further comments and open questions

As a reflection of some of the key points raised in this paper, we offer the following further comments that lead to some open research questions.

1. *Potential limitations of model-dependent realism.* In Section 2, we mentioned the concept of model-dependent realism, which commits us to the following three beliefs/attitudes:

   (a) Pragmatism: a digital twin (or model) is deemed successful if it is able to explain and predict phenomena according to some validation criteria (e.g. making observations). The issue of realism versus non-realism is effectively side-stepped.

   (b) Utility as an over-arching value for digital twin (or model): the new emphasis is on the utility of a digital twin output(s) rather than on finding a digital twin (or model) that is ontologically "true" in terms of representing the behavior (s) of the physical twin.

   (c) Pluralism: as there may be multiple digital twin output(s) that adequately describe the same phenomena, or have similar levels of utility, the choice between different twins may depend on additional (so-called, extra-theoretic virtues)—which also links to the issue of uniqueness of outputs.

   Furthermore, model-dependent realism is developed from a scientific worldview that is focused entirely on explaining the physical behavior of the Universe we live in. It could be considered that the *"direction-of-fit"* is one-way. In other words, the definition of utility is focused primarily on "representation" or "description". For engineering problems, we also need to consider other factors, such as (i) the consequences of utility on subsequent actions taken, such as decisions and interventions in the real world, and (ii) it could be the case that there is no physical system to represent if we are trying to engineer something completely new. In both these cases, the argument for a philosophy built on model-dependent realism is more difficult to make, and leaves open the question of whether there is a more appropriate philosophical approach in these cases? We note also, that more formally the utility, trust, and insight requirements could be contextualized using a more detailed philosophical analysis such as that proposed by Douglas (2013) which distinguishes between internal consistency (minimal criteria) and external consistency (an ideal desiderata, presuming general confidence in other scientific theories). While internal consistency is a minimum requirement for acceptance of any scientific theory, external consistency is not, as it depends on confidence in other theories and external bodies of knowledge.

2. *Emergence is counter-parsimonious.* As was described in Section 4.2, a digital twin will only be able to exhibit behaviors within the constraints of the choices and assumptions that have been made during its construction. Therefore the less simplification in the process of constructing the digital twin, the more likelihood there is for a wider range of emergent behaviours to be exhibited in the subsequent digital twin outputs. The aim stated in Principle 5 (and system engineering Hypothesis H2 from Section 3.2) is to *represent the observed complexity* rather than seek parsimony. Furthermore, it is also possible that if the digital twin maker has been too parsimonious (and/or biased in worldview), there is a possibility of creating a digital twin that is only capable of reinforcing your own (or an inherited) prejudicial expectation. The exact relationship between emergent behaviors and parsimony is an open question.

3. *Purpose dictates your parsimony.* Following on from the comments above, digital twins developed for different purposes will enable different levels of parsimony to be used. Therefore, care should be exercised if transferring a digital twin developed for one purpose into a new domain or purpose. One way to help mitigate these effects is to make use of error detection and correction (EDAC)

(MacKay, 2003). Similar comments relate to the inter-operation or federation of digital twins that might have been constructed using different levels of assumed parsimony. It is an open question of how such systems might be integrated systematically.

4. *Validation of digital twins.* Three comments can be made regarding the validation of digital twin outputs:

    (a) In general the validation of a digital twin is context-specific and will be relevant to a specific application[42]. In some cases, validation can be defined as a function of utility, where the metric of validity relates to the output of some utility function. This situation enables a strong connection with the model-dependent realism philosophy.

    (b) In some applications, the accuracy of a digital twin output does not serve well as a universal metric for validation. For example, from a control perspective, the stability and robustness of a predictive model might be more important than the tracking accuracy of any particular output.

    (c) In Section 4, we presented a framework for defining what potential outputs can be expected from a digital twin. When a system is relatively "simple," it is often possible to *know in advance* what behavior to expect, and therefore validate the output quite easily[43]. Cases where we cannot know what to expect in advance will obviously be more challenging to validate, and there is ongoing research as to how this might be most effectively achieved.

Next, we propose several areas for future research development.

5.2. *Future directions for research*

1. *Human factors.* Broadly this area of research, as it applies to digital twins, includes topics, such as (i) the role of humans in designing and building digital twins (partially discussed in Section 2.2), (ii) how human users interface with digital twins and act on the outputs they receive, and (iii) digital twins as instances of sociotechnical systems that include humans in some way (e.g. medicine or social systems. Early work in this area includes Nguyen (2022); Lin et al. (2022); Sun et al. (2021); Fan et al. (2021).

2. *Ethical, legal, and societal issues.* In their original context as tools for product engineering, digital twins raised a (relatively) narrowly circumscribed set of ethical, legal, and societal issues (e.g. safety compliance). However, as digital twins are now used increasingly to represent not just products or objects, but living entities and systems (from the cellular level to whole ecosystems) they enable new forms of knowledge generation (that is principles 19 and 20: *learning* and *insights* obtained from the digital twin and means for interacting with and influencing the coupled physical systems (principle 21: *exploiting* the derived value of the relevant insights). A number of papers have already explored a variety of normative issues that arise in the context of digital twins, especially in high-stakes and fault-intolerant environments such as health and healthcare (Kuersten, 2023; Huang et al., 2022; Tigard, 2021; Popa et al., 2021; Korenhof et al., 2021; Braun, 2021). In combination with current and emerging frameworks for regulation, governance, and assurance, these analyses provide significant value for identifying and mitigating possible risks that could arise when deploying and using digital twins within society (e.g. unintended behaviors caused by model drift, data privacy, and security violations). There is a lot to explore here, and the presence of a robust and comprehensive philosophical framework could provide a systematic means for grounding and evaluating the myriad normative issues associated with digital twins.

---

[42] This is also true in the assurance of systems and technologies more generally, as claims made about some goal of a system, such as the safety or efficacy of the system, are always contextual to some environment (e.g. an airplane may only be assured as *safe* when in operation by a qualified operator and in a highly-regulated airspace).

[43] Also a limitation in supervised learning techniques, also sometime referred to as an "oracle" e.g. as a source of the "correct" solution.

3. *Methods for dynamic assembly.* In practical terms, one of the most interesting areas for future research is methods that enable dynamic assembly of digital objects within a digital twin. As we have already noted, dynamic assembly is the method by which we can recreate interaction within the digital twin, and thereby reconstruct emergent behaviors. There are already techniques, such as agent-based modeling including intelligent agents, and heterogeneous multi-agents, as discussed and reviewed in Sections 3.3 and 3.4. Such models have the potential to recreate the type of multi-level interactions that occur in a complex system, including socio-economic systems (see for example Yossef Ravid and Aharon-Gutman (2022); Wang et al. (2020); Okita et al. (2019); Tadić (2019)). However, creating appropriate "connectors" for heterogenous sets of digital objects is an open area of research, several of the most obvious potential methods were listed in Section 4.3. The scope for new developments in this area is significant.

4. *The role of knowledge.* While this relates to the topic of human factors listed above, it is significant enough to warrant a separate discussion point—in particular, the role of knowledge and insight, in supporting subsequent actions taken, such as decisions. One way we can distinguish this topic from human factors is the idea of removing the process from the human, and automating the action/decision process, possibly using some form of artificial intelligence. From a practical perspective, a starting point would be to establish with more rigor what knowledge means in the context of a digital twin, particularly linking to topics such as knowledge representation, inference, and model interpretability.[44] This would extend and complement much of the existing work, which primarily focuses on ontologies (e.g. Nguyen, 2022; Akroyd et al., 2021; Singh et al., 2020; West, 2011).

5. *Logic* versus *learning.* In Section 3.4 we touched upon symbolic and neurosymbolic AI but did not explicitly discuss the types of logical approaches that could be applied to digital twins. There has long been a philosophical discussion about how logic, learning, and probability interrelate (see for example the discussion in Belle (2022)). This is an interesting topic that has several relevant questions for digital twin research. For example, is there an underlying logical methodology relating to digital twins, or is the logic dependent on the context? How is the logic and learning combined? How does the logic relate to a "top-down" versus "bottom-up" approach to creating a digital twin?[45] It should also be noted that statistical relational learning and hyperdimensional computing are novel approaches that enable knowledge representation, logic, and learning to be brought together (Getoor and Taskar, 2007; Thomas et al., 2021). These processes offer the possibility of bringing logic more formally into the digital twin operation. The exact details of how this might work are an open question.

---

[44] See for example the discussions and examples in Pan et al. (2023); Olsson and Axelsson (2023); Akroyd et al. (2021); Li et al. (2021).

[45] For example, there may be things to learn from the idea of polycentricity Ostrom (2010).

References

Åström KJ and Wittenmark B (1995). *Adaptive Control*. Addison Wesley: Menlo Park, CA, USA.

Abar S, **Theodoropoulos GK**, **Lemarinier P and O"Hare GM** (2017). Agent based modelling and simulation tools: A review of the state-of-art software. *Computer Science Review*, *24*:13–33.

Adams KM and Meyers TJ (2011). Perspective 1 of the sose methodology: Framing the system under study. *International Journal of System of Systems Engineering*, *2*(2–3):163–192.

Akroyd J, **Mosbach S**, **Bhave A and Kraft M** (2021). Universal digital twin—a dynamic knowledge graph. *Data-Centric Engineering*, *2*:100079.

Ambra T, **Caris A and Macharis C** (2019). *The digital twin concept and its role in reducing uncertainty in synchromodal transport. Proceedings of the IPIC*. 1–14.

Area I, **Fernández FJ**, **Nieto JJ and Tojo FAF** (2022). Concept and solution of digital twin based on a stieltjes differential equation. *Mathematical Methods in the Applied Sciences*, *45*(12):7451–7465.

Arthur R, **French M**, **Ganguli JS**, **Kinard DA**, **Kraft EM**, **Marks I**, **Matlik JF**, **Fischer OJP**, **Sangid MD**, **Seal D**, **Tucker E and Vickers J** (2020). *Digital Twin: Definition & Value—An AIAA and AIA Position Paper*. AIAA: Reston, VA, USA. 1–16.

Ashby WR (1956). *An Introduction to Cybernetics*. Chapman & Hall, New York, USA.

Aven T (2013). On the meaning of a black swan in a risk context. *Safety Science*, *57*:44–51.

Axelrod R (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration: Agent-Based Models of Competition and Collaboration*. Princeton University Press, Princeton, NJ, USA.

Bainbridge WS and Roco MC (2016). *Handbook of Science and Technology Convergence*. Springer Verlag, New York.

Baker A (2013). Complexity, networks, and non-uniqueness. *Foundations of Science*, *18*:687–705.

Banerjee S, **Carlin BP and Gelfand AE** (2014). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press, Boca Raton, FL, USA.

Barber D (2012). *Bayesian Reasoning and Machine Learning*. Cambridge University Press, Cambridge, UK.

Barlow JS (2002). *The Cerebellum and Adaptive Control*. Cambridge University Press, Cambridge, UK.

Barthelmess P and Anderson KM (2002). A view of software development environments based on activity theory. *Computer Supported Cooperative Work (CSCW)*, *11*(1–2):13–37.

Bedau MA and Humphreys P (2008). *Emergence: Contemporary readings in philosophy and science*. MIT press, Cambridge, MA, USA.

Beer S (1985). *Diagnosing the System for Organizations*. John Wiley & Sons, New York, USA.

Belle V (2022). Logic meets learning: From aristotle to neural networks. *Neuro-Symbolic Artificial Intelligence: The State of the Art*, *342*:78.

Boje C, **Guerriero A**, **Kubicki S and Rezgui Y** (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction*, *114*:103179.

Bolton A, **Butler L**, **Dabson I**, **Enzer M**, **Evans M**, **Fenemore T**, **Harradence F**, **Keaney E**, **Kemp A and Luck A** (2018). Gemini Principles. https://www.cdbb.cam.ac.uk/system/files/documents/TheGeminiPrinciples.pdf

Botín-Sanabria DM, **Mihaita A-S**, **Peimbert-García RE**, **Ramírez-Moreno MA**, **Ramírez-Mendoza RA and Lozoya-Santos JDJ** (2022). Digital twin technology challenges and applications: A comprehensive review. *Remote Sensing*, *14*(6):1335.

Boyd SP and Vandenberghe L (2004). *Convex Optimization*. Cambridge University Press, Cambridge, UK.

Braun M (2021). Represent me: Please! Towards an ethics of digital twins in medicine. *Journal of Medical Ethics*, *47*(6):394–400.

Busse A, **Gerlach B**, **Lengeling JC**, **Poschmann P**, **Werner J and Zarnitz S** (2021). Towards digital twins of multimodal supply chains. *Logistics*, *5*(2):25.

Calabrese M, **Amato A**, **Lecce VD and Piuri V** (2010). Hierarchical-granularity holonic modelling. *Journal of Ambient Intelligence and Humanized Computing*, *1*:199–209.

Caldarelli G, **Arcaute E**, **Barthelemy M**, **Batty M**, **Gershenson C**, **Helbing D**, **Mancuso S**, **Moreno Y**, **Ramasco J**, **Rozenblat C**, **Sánchez A and Fernández-Villacañas JL** (2023). The role of complexity for digital twins of cities. *Nature Computational Science*, pages 1–8. https://doi.org/10.1038/s43588-023-00431-4

Cardin O, **Derigent W and Trentesaux D** (2018). Evolution of holonic control architectures towards industry 4.0: A short overview. *IFAC-PapersOnLine*, *51*(11):1243–1248.

Cash P, **Hicks B and Culley S** (2015). Activity theory as a means for multi-scale analysis of the engineering design process: A protocol study of design in practice. *Design Studies*, *38*:1–32.

Checkland P (1999). *Systems Thinking, systems practice*. Chichester: John Wiley, New York, USA.

Chouard T (2016). The Go files: AI computer clinches victory against Go champion. *Nature*. https://doi.org/10.1038/nature.2016.19553

Christakis A (2006). Dialogue for conscious evolution. *Wisdom, Knowledge, and Management: A Critique and Analysis of Churchman's Systems Approach*. Springer Verlag Berlin Heidelberg, pages 279–301.

Churchman CW, **Ackoff RL and Arnoff EL** (1957). *Introduction to Operations Research*. John Wiley, Hoboken, NJ, USA.

Cimino C, **Negri E and Fumagalli L** (2019). Review of digital twin applications in manufacturing. *Computers in Industry*, *113*:103130.

Clemen T, **Ahmady-Moghaddam N**, **Lenfers UA**, **Ocker F**, **Osterholz D**, **Ströbele J and Glake D** (2021). Multi-agent systems and digital twins for smarter cities. In *Proceedings of the 2021 ACM SIGSIM conference on Principles of Advanced Discrete Simulation*, pages 45–55.

Committee on Foundational Research Gaps and Future Directions for Digital Twins, Board on Mathematical Sciences and Analytics, Committee on Applied and Theoretical Statistics, Computer Science and Telecommunications Board, Board on Life Sciences, Board on Atmospheric Sciences and Climate, Division on Engineering and Physical Sciences, Division on Earth and Life Studies, National Academy of Engineering, and National Academies of Sciences, Engineering, and Medicine (2024). *Foundational Research Gaps and Future Directions for Digital Twins*. National Academies Press, Washington, DC, USA.

Crisfield MA (1997). *Non-linear finite element analysis of solids and structures. Volume 2: Advanced Topics*. John Wiley, New York, USA.

Croatti A, **Gabellini M**, **Montagna S and Ricci A** (2020). On the integration of agents and digital twins in healthcare. *Journal of Medical Systems*, *44*:1–8.

Curd M, **Psillos S**, *et al.* (2014). *The Routledge Companion to Philosophy of Science*. Routledge, London, UK.

Dale KI, **Pope EC**, **Hopkinson AR**, **McCaie T and Lowe JA** (2023). Environment-aware digital twins: Incorporating weather and climate information to support risk-based decision-making. *Artificial Intelligence for the Earth Systems*, *2*(4):e230023.

Dandashi F, **Siegers R**, **Jones J and Blevins T** (2006). The open group architecture framework (togafTOGAF) and the us department of defense architecture framework (DoDAF). *The Open Group*, *4*. 1–59.

DebRoy T, **Zhang W**, **Turner J and Babu S** (2017). Building digital twins of 3d printing machines. *Scripta Materialia*, *135*: 119–124.

Dennett DC (1991). Real patterns. *The Journal of Philosophy*, *88*(1):27–51.

Descartes R (1985). *The Philosophical Writings of Descartes*. Cambridge University Press, Cambridge, UK.

dos Santos CH, **Montevechi JAB**, **de Queiroz JA**, **de Carvalho Miranda R and Leal F** (2022). Decision support in productive processes through des and abs in the digital twin era: a systematic literature review. *International Journal of Production Research*, *60*(8):2662–2681.

Douglas H (2013). The value of cognitive values. *Philosophy of Science*, *80*(5):796–806.

Enders MR and Hoßbach N (2019). *Dimensions of digital twin applications – a literature review. In Proceedings of Twenty-fifth Americas Conference on Information Systems*. 1-10.

Errandonea I, **Beltrán S and Arrizabalaga S** (2020). Digital twin for maintenance: A literature review. *Computers in Industry*, *123*:103316.

Estefan JA *et al.* (2007). Survey of model-based systems engineering (MBSE) methodologies. *Incose MBSE Focus Group*, *25*(8): 1–12.

Evensen G *et al.* (2009). *Data Assimilation: The Ensemble Kalman Filter*, volume *2*. Springer, Berlin Heidelberg, Germany.

Fan C, **Zhang C**, **Yahja A and Mostafavi A** (2021). Disaster city digital twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management*, *56*:102049.

Farhat MH, **Chiementin X**, **Chaari F**, **Bolaers F and Haddar M** (2020). *Digital Twin-Driven Machine Learning: Ball Bearings Fault Severity Classification*. Measurement Science and Technology. *32*(4): 04400.

Fernández N, **Maldonado C and Gershenson C** (2014). Information measures of complexity, emergence, self-organization, homeostasis, and autopoiesis. In *Guided Self-Organization: Inception*, *9*:19–51. Springer Verlag Berlin Heidelberg.

Fradkov AL, **Miroshnik IM and Nikiforov VO** (1999). *Nonlinear and Adaptive Control of Complex Systems*. Kluwer: Dordrecht, The Netherlands.

Frei R and Serugendo GDM (2012). The future of complexity engineering. *Central European Journal of Engineering*, *2*(2): 164–188.

Gelman A, **Carlin JB**, **Stern HS and Dunson DB** (2014). *Bayesian Data Analysis*, volume *3*. CRC press, Boca Raton, FL, USA.

Gershenson C (2007). *Design and Control of Self-Organizing Systems*. PhD thesis, Vrije Universiteit Brussel, Belgium.

Gershenson C (2013). The implications of interactions for science and philosophy. *Foundations of Science*, *18*:781–790.

Getoor L and Taskar B (2007). *Introduction to Statistical Relational Learning*. MIT Press, Cambridge, MA, USA.

Giere RN (2019). *Scientific Perspectivism*. University of Chicago Press, Chicago, USA.

Gonzalez-Buelga A, **Wagg DJ**, **Wallace MI**, **Neild SA and Macdonald JHG** (2005). Testing an autoparametric pendulum using a hybrid numerical experimental method. In *Proceedings of the 1st International Conference on Advances in Experimental Structural Engineering*. *1*:417–424.

Goodfellow I, **Bengio Y and Courville A** (2016). *Deep Learning*. MIT press, Cambridge, MA, USA.

Graesser L and Keng WL (2019). *Foundations of Deep Reinforcement Learning*. Addison-Wesley Professional, Boston, USA.

Greiner W (1994). *Quantum Mechanics, an Introduction*. Springer-Verlag, Berlin Heidelberg, Germany.

Grieves M (2005). *Product Lifecycle Management: Driving the Next Generation of Lean Thinking*. McGraw-Hill Professional, New York, USA.

Grieves M and Vickers J (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*, Springer Verlag Berlin Heidelberg, pages 85–113.

Grieves MW (2019). Virtually intelligent product systems: Digital and physical twins. In *Engineering: Theory & Practice*. American Institute of Aeronautics and Astronautics. 2–34.

Guckenheimer J and Holmes P (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag: New York.

Haenlein M and Kaplan A (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*(4):5–14.

Hamilton JD (2020). *Time Series Analysis*. Princeton University Press, Princeton, NJ, USA.

Han J and Kamber M (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Burlington, MA, USA.

Han Y, **Niyato D**, **Leung C**, **Kim DI**, **Zhu K**, **Feng S**, **Shen X and Miao C** (2022). A dynamic hierarchical framework for iot-assisted digital twin synchronization in the metaverse. *IEEE Internet of Things Journal*, *10*(1):268–284.

Hastie T, **Tibshirani R**, **Friedman JH and Friedman JH** (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Volume *2*. Springer Verlag, Berlin Heildelberg, Germany.

Hawking S and Mlodinow L (2010). *The Grand Design*. Random House Digital, Inc, London, UK.

He B and Bai K-J (2021). Digital twin-based sustainable intelligent manufacturing: A review. *Advances in Manufacturing*, *9*(1): 1–21.

Heber D and Groll M (2017). *Towards a digital twin: How the blockchain can foster e/e-traceability in consideration of model-based systems engineering. In DS 87–3 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 3: Product, Services and Systems Design, Vancouver, Canada, 21–25.08. 2017*.

Heylighen F, **Cilliers P and Garcia CG** (2007). Complexity and philosophy. In *Complexity, Science and Society*, pages 117–134. Radcliffe Publishing, Abingdon, UK.

Hirsch MW and Smale S (1974). *Differential Equations, Dynamical Systems , and Linear Algebra*. Academic Press, San Diego, CA, USA.

Hirshorn SR, **Voss LD and Bromley LK** (2017). *Nasa systems Engineering Handbook*. NASA, Washington, DC, USA.

Holland OT (2007). Taxonomy for the modeling and simulation of emergent behavior systems. In *Proceedings of the 2007 Spring Simulation Multiconference 2*:28–35.

Holland T (2018). *Foundations for the Modeling and Simulation of Emergent Behavior Systems*. CRS Press/Taylor and Francis, Boca Raton, FL, USA.

Huang P-h., **Kim K-h., and Schermer M** (2022). Ethical Issues of Digital Twins for Personalized Health Care Service: Preliminary Mapping Study. *Journal of Medical Internet Research*, *24*(1):e33081.

Huang Z, **Shen Y**, **Li J**, **Fey M and Brecher C** (2021). A survey on ai-driven digital twins in industry 4.0: Smart manufacturing and advanced robotics. *Sensors*, *21*(19):6340.

Huelsenbeck JP, **ANe C**, **Large B and Ronquist F** (2008). A bayesian perspective on a non-parsimonious parsimony model. *Systematic Biology*, *57*(3):406–419.

Hughes AJ, **Barthorpe RJ**, **Dervilis N**, **Farrar CR and Worden K** (2021). A probabilistic risk-based decision framework for structural health monitoring. *Mechanical Systems and Signal Processing*, *150*:107339.

Hughes I and Hase T (2010). *Measurements and Their Uncertainties: A practical Guide to Modern Error Analysis*. Oxford University Press, Oxford, UK.

Ioannidis JP (2005). Why most published research findings are false. *PLoS medicine*, *2*(8):e124.

Jafari M, **Kavousi-Fard A**, **Chen T and Karimi M** (2023). A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future. *IEEE Access*. *11*:17471–17484.

Jensen HJ (2022). *Complexity Science: The Study of Emergence*. Cambridge University Press, Cambridge, UK.

Jiang F, **Ma L**, **Broyd T and Chen K** (2021). Digital twin and its implementations in the civil engineering sector. *Automation in Construction*, *130*:103838.

Jin L, **Zhai X**, **Wang K**, **Zhang K**, **Wu D**, **Nazir A**, **Jiang J and Liao W-H** (2024). Big data, machine learning, and digital twin assisted additive manufacturing: A review. *Materials & Design*, page 113086.

Jinzhi L, **Zhaorui Y**, **Xiaochen Z**, **Jian W and Dimitris K** (2022). Exploring the concept of cognitive digital twin from model-based systems engineering perspective. *The International Journal of Advanced Manufacturing Technology*, *121*(9–10): 5835–5854.

Jones AJ (2000). *Game Theory: Mathematical Models of Conflict*. Elsevier, Amsterdam, The Netherlands.

Jones D, **Snider C**, **Nassehi A**, **Yon J and Hicks B** (2020). Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology. 29*: 36–52.

Kapteyn MG, **Knezevic DJ and Willcox K** (2020). Toward predictive digital twins via component-based reduced-order models and interpretable machine learning. In *AIAA Scitech 2020 Forum*, page 0418.

Katsoulakis E, **Wang Q**, **Wu H**, **Shahriyari L**, **Fletcher R**, **Liu J**, **Achenie L**, **Liu H**, **Jackson P**, **Xiao Y**, *et al.* (2024). Digital twins for health: A scoping review. *npj Digital Medicine*, *7*(1):77.

Kauffman SA (2000). *Investigations*. Oxford University Press, Oxford, UK.

King AC, **Billingham J and Otto SR** (2003). *Differential Equations*. Cambridge, UK.

Kiran BR, **Sobh I**, **Talpaert V**, **Mannion P**, **Al Sallab AA**, **Yogamani S and Pérez P** (2021). Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, *23*(6):4909–4926.

Korenhof P, **Blok V and Kloppenburg S** (2021). Steering Representations—Towards a Critical Understanding of Digital Twins. *Philosophy & Technology*, *34*(4):1751–1773.

Krishnamoorthy C and Rajeev S (2018). *Artificial Intelligence and Expert Systems for Engineers*. CRC press, Boca Raton, FL, UK.

Kritzinger W, **Karner M**, **Traar G**, **Henjes J and Sihn W** (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11):1016–1022.

Kruschke J (2014). *Doing Bayesian Data Analysis: A Tutorial With R, JAGS, and Stan*. Academic Press, San Diego, CA, USA.

Kuersten A (2023). Health digital twins, legal liability, and medical practice. *The American Journal of Bioethics*, 23(9):66–69.

Kutz JN (2013). *Data Assimilation Methods*, pages 539–555. Oxford University Press, Oxford, UK.

Kuznetsov YA (2004). *Elements of Applied Bifurcation Theory*. Springer Verlag Berlin Heidelberg, Germany.

Ladyman J, **Lambert J and Wiesner K** (2013). What is a complex system? *European Journal for Philosophy of Science*, 3:33–67.

Lanza RB (2000). Does your project risk management system do the job? *Information Strategy: The Executive's journal*, 17(1):6–12.

Laplace P-S (2012). *A Philosophical Essay on Probabilities*. Courier Corporation, North Chelmsford, MA, USA.

LeCun Y, **Bengio Y and Hinton G** (2015). Deep learning. *Nature*, 521(7553):436–444.

Li J, **Zhou G and Zhang C** (2021). A twin data and knowledge-driven intelligent process planning framework of aviation parts. *International Journal of Production Research*, 60(17), 5217–5234.

Lin Y, **Chen L**, **Ali A**, **Nugent C**, **Ian C**, **Li R**, **Gao D**, **Wang H**, **Wang Y and Ning H** (2022). Human digital twin: A survey. *Journal of Cloud Computing* 13(1), 131.

Liu M, **Fang S**, **Dong H and Xu C** (2020). Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems*. 58: 346–361.

Liu S, **Bao J and Zheng P** (2023). A review of digital twin-driven machining: From digitization to intellectualization. *Journal of Manufacturing Systems*, 67:361–378.

Lo C, **Chen C and Zhong RY** (2021). A review of digital twin in product design and development. *Advanced Engineering Informatics*, 48:101297.

Lu Q, **Chen L**, **Lee S and Zhao X** (2018). Activity theory-based analysis of bim implementation in building o&m and first response. *Automation in Construction*, 85:317–332.

MacKay DJ (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, Cambridge, UK.

Madni AM, **Madni CC and Lucero SD** (2019). Leveraging digital twin technology in model-based systems engineering. *Systems*, 7(1):7.

Marcus G (2018). *Deep learning: A critical appraisal. arXiv preprint* arXiv:1801.00631.

Marcus G (2020). *The next decade in AI: four steps towards robust artificial intelligence*. arXiv preprint arXiv:2002.06177.

Marques TA (2021). On how statistics is used and abused to find truth in science. *Kairos. Journal of Philosophy & Science*. 24:1–14.

Marsh HW and Hau K-T (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education*, 64(4):364–390.

McCorduck P (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. CRC Press, Boca Raton, FL, USA.

Meadows DH (2008). *Thinking in Systems: A Primer*. Chelsea Green publishing, White River Junction, VT, USA.

Melesse TY, **Di Pasquale V and Riemma S** (2020). Digital twin models in industrial operations: A systematic literature review. *Procedia Manufacturing*, 42:267–272.

Michael J, **Pfeiffer J**, **Rumpe B and Wortmann A** (2022). Integration challenges for digital twin systems-of-systems. In *Proceedings of the 10th IEEE/ACM International Workshop on Software Engineering for Systems-of-Systems and Software Ecosystems*, 1:9–12.

Miller AM, **Alvarez R and Hartman N** (2018). Towards an extended model-based definition for the digital twin. *Computer-Aided Design and Applications*, 15(6):880–891.

Minsky M (1988). *Society of Mind*. Simon and Schuster, New York, USA.

Mitchell M (2009). *Complexity: A Guided Tour*. Oxford university press, Oxford, UK.

Nativi, S Mazzetti, P and Craglia, M (2021) Digital ecosystems for developing digital twins of the earth: The destination earth case Remote Sensing, *MDPI. 13*:2019.

Negri E, **Fumagalli L and Macchi M** (2017). A review of the roles of digital twin in cps-based production systems. *Procedia Manufacturing*, 11:939–948.

Newton I (1686). *Philosophiae Naturalis Principia Mathematica*. Reg. Soc. Præss, London, UK.

Nguyen TN (2022). Toward human digital twins for cybersecurity simulations on the metaverse: Ontological and network science approach. *JMIRx Med*, 3(2):e33502.

Nilsson NJ (2009). *The Quest for Artificial Intelligence*. Cambridge University Press, Cambridge, UK.

Oberkampf WL and Roy CJ (2010). *Verification and Validation in Scientific Computing*. Cambridge University Press, Cambridge, UK.

Okashah LA and Goldwater PM (1994). Unknown unknowns: modeling unanticipated events. In *Proceedings of Winter Simulation Conference*, 1:689–694. IEEE.

Okita T, **Kawabata T**, **Murayama H**, **Nishino N and Aichi M** (2019). A new concept of digital twin of artifact systems: synthesizing monitoring/inspections, physical/numerical models, and social system models. *Procedia Cirp*, 79:667–672.

Olsson T and Axelsson J (2023). Systems-of-systems and digital twins: A survey and analysis of the current knowledge. In *2023 18th Annual System of Systems Engineering Conference (SoSe)*, 1:1–6. IEEE.

Ostrom E (2010). Beyond markets and states: polycentric governance of complex economic systems. *American economic review*, *100*(3):641–672.

Pan S, **Luo L**, **Wang Y**, **Chen C**, **Wang J and Wu X** (2023). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*, *36*:3580–3599.

Piras G, **Agostinelli S and Muzi F** (2024). Digital twin framework for built environment: A review of key enablers. *Energies*, *17*(2): 436.

Popa EO, **Van Hilten M**, **Oosterkamp E and Bogaardt M-J** (2021). The use of digital twins in healthcare: Socio-ethical benefits and socio-ethical risks. *Life Sciences, Society and Policy*, *17*(1):6.

Purcell W and Neubauer T (2022). Digital twins in agriculture: A state-of-the-art review. *Smart Agricultural Technology*, *3*:100094.

Ríos J, **Hernández JC**, **Oliva M and Mas F** (2015). Product avatar as digital counterpart of a physical individual product: Literature review and implications in an aircraft. In *ISPE CE*, *1*:657–666.

Ritto T and Rochinha F (2021). Digital twin, physics-based model, and machine learning applied to damage detection in structures. *Mechanical Systems and Signal Processing*, *155*:107614.

Rosen R, **von Wichert G**, **Lo G and Bettenhausen KD** (2015). About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, *48*(3):567–572.

Rovelli C (2016). *Seven Brief Lessons on Physics*. Riverhead Books, New York, USA.

Rozenblat A and Fernández-Villacanas J (2023). *Complexity science for digital twins. Preprint.* https://www.researchgate.net/publication/365852772_Complexity_Science_for_Digital_Twins

Russell S and Norvig P (2010). *Artificial Intelligence: A Modern Approach*. Pearson, London, UK.

Särkkä S (2013). *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge, UK.

Savage T, **Akroyd J**, **Mosbach S**, **Hillman M**, **Sielker F and Kraft M** (2022). Universal digital twin–the impact of heat pumps on social inequality. *Advances in Applied Energy*, *5*:100079.

Schlager KJ (1956). Systems engineering-key to modern development. *IRE Transactions on Engineering Management*, *1*(3): 64–66.

Schluse M, **Priggemeyer M**, **Atorf L and Rossmann J** (2018). Experimentable digital twins—streamlining simulation-based systems engineering for industry 4.0. *IEEE Transactions on Industrial Informatics*, *14*(4):1722–1731.

Scott JC (2020). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press, Yale, USA.

Semeraro C, **Lezoche M**, **Panetto H and Dassisti M** (2021). Digital twin paradigm: A systematic literature review. *Computers in Industry*, *130*:103469.

Sepasgozar SM, **Khan AA**, **Smith K**, **Romero JG**, **Shen X**, **Shirowzhan S**, **Li H and Tahmasebinia F** (2023). Bim and digital twin for developing convergence technologies as future of digital construction. *Buildings*, *13*(2):441.

Shahat E, **Hyun CT and Yeom C** (2021). City digital twin potentials: A review and research agenda. *Sustainability*, *13*(6):3386.

Shannon CE (1948). The mathematical theory of communication. *The Bell System Technical Journal*, .*27*:379–423.

Silver D, **Huang A**, **Maddison CJ**, **Guez A**, **Sifre L**, **Van Den Driessche G**, **Schrittwieser J**, **Antonoglou I**, **Panneershelvam, V**, **Lanctot M**, **Dieleman S**, **Grewe, D**, **Nham J**, **Kalchbrenner N**, **Sutskever I**, **Lillicrap T**, **Leach M**, **Kavukcuoglu K**, **Graepel T and Hassabis, D** (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, *529*(7587):484–489.

Singh M, **Srivastava R**, **Fuenmayor E**, **Kuts V**, **Qiao Y**, **Murray N and Devine D** (2022). Applications of digital twin across industries: A review. *Applied Sciences*, *12*(11):5727.

Singh S, **Shehab E**, **Higgins N**, **Fowler K**, **Reynolds D**, **Erkoyuncu JA and Gadd P** (2020). Data management for developing digital twin ontology model. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, page 0954405420978117.

Siyaev A, **Valiev D and Jo G-S** (2023). Interaction with industrial digital twin using neuro-symbolic reasoning. *Sensors*, *23*(3): 1729.

Snow CP (2012). *The Two Cultures*. Cambridge University Press, Cambridge, UK.

Somers RJ, **Douthwaite JA**, **Wagg DJ**, **Walkinshaw N and Hierons RM** (2022). Digital-twin-based testing for cyber–physical systems: A systematic literature review. *Information and Software Technology*, page 107145.

Strogatz SH (2019). *Nonlinear Dynamics and Chaos*. Perseus Books Group, New York, USA.

Sun J, **Tian Z**, **Fu Y**, **Geng J and Liu C** (2021). Digital twins in human understanding: A deep learning-based method to recognize personality traits. *International Journal of Computer Integrated Manufacturing*, *34*(7–8):860–873.

Tadić B (2019). Self-organised criticality and emergent hyperbolic networks: blueprint for complexity in social dynamics. *European Journal of Physics*, *40*(2):024002.

Taleb N (2007). *The Black Swan:The Impact of the Highly Improbable*. Penguin, London, UK.

Tao F, **Xiao B**, **Qi Q**, **Cheng J and Ji P** (2022). Digital twin modeling. *Journal of Manufacturing Systems*, *64*:372–389.

Teubner T, **Flath CM**, **Weinhardt C**, **van der Aalst W and Hinz O** (2023). Welcome to the era of chatgpt et al the prospects of large language models. *Business & Information Systems Engineering*, *65*(2):95–101.

Thelen A, **Zhang X**, **Fink O**, **Lu Y**, **Ghosh S**, **Youn BD**, **Todd MD**, **Mahadevan S**, **Hu C and Hu Z** (2023). A comprehensive review of digital twin—part 2: Roles of uncertainty quantification and optimization, a battery digital twin, and perspectives. *Structural and Multidisciplinary Optimization*, *66*(1):1.

Thomas A, **Dasgupta S and Rosing T** (2021). A theoretical perspective on hyperdimensional computing. *Journal of Artificial Intelligence Research*, *72*:215–249.

Tigard DW (2021). Digital twins running amok? Open questions for the ethics of an emerging medical technology. *Journal of Medical Ethics*, *47*(6):407–408.

Tononi G, **Boly M**, **Massimini M and Koch C** (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience*, *17*(7):450–461.

Tripura T, **Desai AS**, **Adhikari S and Chakraborty S** (2023). Probabilistic machine learning based predictive and interpretable digital twin for dynamical systems. *Computers & Structures*, *281*:107008.

Tuegel EJ, **Ingraffea AR**, **Eason TG and Spottswood SM** (2011). Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering, 2011*, Article ID 154798.

Turing A (1950). Computing machinery and intelligence. *Mind*, *59*(236):433–60.

Van Beurden EK, **Kia AM**, **Zask A**, **Dietrich U and Rose L** (2013). Making sense in a complex landscape: How the cynefin framework from complex adaptive systems theory can inform health promotion practice. *Health Promotion International*, *28*(1):73–83.

Van der Vaart AW (2000). *Asymptotic statistics*, volume *3*. Cambridge university press, Cambridge, UK.

Varga L (2018). Mixed methods research: A method for complex systems. In *Handbook of Research Methods in Complexity Science*. Edward Elgar Publishing, Cheltenham, UK.

Vrabič R, **Erkoyuncu JA**, **Farsi M and Ariansyah D** (2021). An intelligent agent-based architecture for resilient digital twins in manufacturing. *CIRP Annals*, *70*(1):349–352.

Wagg DJ and Neild SA (2015). *Nonlinear Vibration With Control – For Flexible and Adaptive Structures*. Springer Verlag Berlin Heidelberg, 2nd edition.

Wagg DJ, **Worden K**, **Barthorpe RJ and Gardner P** (2020). Digital twins: State-of-the-art and future directions for modeling and simulation in engineering dynamics applications. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *6*(3). 030901.

Walden DD, **Roedler GJ**, **Forsberg K**, **Hamelin RD and Shortell TM** (2015). *INCOSE Systems Engineering Handbook*. John Wiley, New York, USA.

Waldrop MM (1993). *Complexity: The Emerging Science at the Edge of Order and Chaos*. Simon and Schuster, New York, USA.

Wanasinghe TR, **Wroblewski L**, **Petersen B**, **Gosine RG**, **James LA**, **De Silva O**, **Mann GK and Warrian PJ** (2020). Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. *IEEE Access*. *8*:104175–104197.

Wang B, **Zhou H**, **Li X**, **Yang G**, **Zheng P**, **Song C**, **Yuan Y**, **Wuest T**, **Yang H and Wang L** (2024). Human digital twin in the context of industry 5.0. *Robotics and Computer-Integrated Manufacturing*, *85*:102626.

Wang F-Y, **Qin R**, **Li J**, **Yuan Y and Wang X** (2020). Parallel societies: A computing perspective of social digital twins and virtual–real interactions. *IEEE Transactions on Computational Social Systems*, *7*(1):2–7.

Watson MD (2019). Systems engineering principles and hypotheses. *Insight*, *22*(1):18–28.

Watson MD, **Mesmer B and Farrington P** (2019). Engineering elegant systems: Postulates, principles, and hypotheses of systems engineering. In *Systems Engineering in Context: Proceedings of the 16th Annual Conference on Systems Engineering Research*, pages 495–513. Springer Verlag.

Weaver W (1948). Science and complexity. *American Scientist*, *36*(4):536–544.

Weil S (1968). Reflections on quantum theory. *Translated by Richard Rees. In On Science, Necessity, and the Love of God*. Royal Institute of Philosophy, Cambridge University Press, Cambidge, UK, pages 49–64.

Weinbaum DR (2015). Complexity and the philosophy of becoming. *Foundations of Science*, *20*(3):283–322.

West M (2011). *Developing High Quality Data Models*. Elsevier, Amsterdam, The Netherlands.

Wolf K, **Dawson RJ**, **Mills JP**, **Blythe P and Morley J** (2022). Towards a digital twin for supporting multi-agency incident management in a smart city. *Scientific Reports*, *12*(1):16221.

Wright L and Davidson S (2020). How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, *7*(1):1–13.

Xames MD and Topcu TG (2024). A systematic literature review of digital twin research for healthcare systems: Research trends, gaps, and realization challenges. *IEEE Access*. *12*:4099–4126.

Yossef Ravid B and Aharon-Gutman M (2022). The social digital twin: The social turn in the field of smart cities. *Environment and Planning B: Urban Analytics and City Science*, page 23998083221137079.

Zhang K, **Cao J**, **Maharjan S and Zhang Y** (2021a). Digital twin empowered content caching in social-aware vehicular edge networks. *IEEE Transactions on Computational Social Systems*, *9*(1):239–251.

Zhang W, **Valencia A and Chang N-B** (2021b). Synergistic integration between machine learning and agent-based modeling: A multidisciplinary review. *IEEE Transactions on Neural Networks and Learning Systems*. *34*(5):2170–2190.

Zheng X, **Psarommatis F**, **Petrali P**, **Turrin C**, **Lu J and Kiritsis D** (2020). A quality-oriented digital twin modelling method for manufacturing processes based on a multi-agent architecture. *Procedia Manufacturing*, *51*:309–315.

Zhou Z-H (2019). *Ensemble Methods: Foundations and Algorithms*. Chapman and Hall/CRC, Boca Raton, FL, USA.