The philosophical foundations of digital twinning

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

1. The Alan Turing Institute, London, NW1 2DB, UK.
2. Department of Mechanical Engineering, University of Sheffield, Sheffield, S1 3JD, UK.
3. Fujitsu Services Limited, Lovelace Road, Bracknell, RG12 8SN, UK.
4. Mott MacDonald, 8-10 Sydenham Road, Croydon, CR0 2EE, UK.
5. Imperial College London, South Kensington Campus, London, SW7 2AZ, UK.

[*] dwagg@turing.ac.uk, david.wagg@sheffield.ac.uk

25 January 2024

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 practitice, 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, because they have different starting points, assumptions and cultural practices. These differences are often due to the established philosophical world-view adopted within specific communities of practice. Therefore, in this paper we explore the philosophical context which underpins the concept of digital twins. As part of this effort we offer a set of philosophical principles for digital twins, which are intended to help facilitate their further development. Specifically, we argue that the philosophy of digital twins is fundamentally *holistic*. We further argue that digital twins are *reconstructivist*, meaning they aim to reconstruct the behaviour of a physical twin by assembling multiple “components”, e.g. models, agents and data sets. Importantly, these digital twin components have the potential to capture emergent behaviours when they are *dynamically assembled*. Lastly, 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?

Keywords: Digital twin, philosophy, modelling, complexity, systems, artificial intelligence

Impact Statement Creating digital twins (or the process of digital twinning) is an 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 philosophical framework for digital twins that considers how such a holistic integration can be achieved, including current questions of interest, and challenges for future research.

1 Introduction

A digital twin is a virtual representation of a physical system (called the *physical twin*) that enables a two-way coupling between the digital and physical domains, using some form of network-based connectivity. The digital twin evolves over time and is constructed from digitised information such as recorded data and the output of computational models.

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 was 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 realised, the improvements in structural life prediction will be substantial” — Tuegel et al. (2011).

This is certainly a very exciting prospect, however, as engineers we always need to be cautiously pragmatic and it is worth keeping in mind the observations of many experienced practitioners. For example, to pick just one related insight, even before the idea of digital twins was proposed:

> “Mete A. Sozen, Kettelhut Distinguished Professor of Structural Engineering at Purdue University, presented the 2002 Distinguished Lecture in February at the EERI Annual Meeting in Long Beach, California. His lecture was entitled *A Way of Thinking*. Sozen was motivated in selecting his topic by the fact that at the present time, ready access to versatile and powerful software enables the engineer to do more and think less, which in his opinion makes it especially important to reflect thoughtfully on the role of analysis in design.” — Sozen (2002).

Although it couldn’t be known at the time, Sozen’s observation about software enabling engineers to “do more and think less” is relevant not just from the point of view of over-reliance on software, but also because the recent advent of large language models like ChatGPT (see for example Teubner et al. 2023), and other developments in artificial intelligence (AI), offer the prospect of a non-human AI doing at least some of the thinking for us as so-called "cognitive surrogates" Leslie (2021).

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”, meaning models, data, AI methods, processes, controls, decision, etc. in some optimal way. In addition, it is also often implied that digital twins will somehow overcome the fundamental challenges and limitations related to modelling 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, scepticism can become cynicism, and genuine scientific and technological progress can become stalled.

We believe that part of the underlying issue is that because the concept of a digital twin is so versatile and universally applicable 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 built upon their specific world view. Typically these communities are domain-specific, and have excellent reasons for their adopted philosophical culture, but are often operating in a silo, or at least only interacting with those who share a similar approach to themselves.

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

— see for example Okita et al. (2019); Wang et al. (2020); Zhang et al. (2021a); Savage et al. (2022); Yossef Ravid and 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 which underpins the concept of a digital twin. Firstly, in Section 2 a brief review of the historical and philosophical context of digital twins is presented. The role of modelling will be key to this discussion, and the distinction between models and digital twins, and the discussion related to this theme will be started in this section. Then in Section 3 we consider the types of complexity that occur in engineering systems, and how this might be represented in a digital twin. In Section 3.5 we briefly consider the role of human interpretations and bias. Then in Section 4 we introduce a philosophical framework for digital twins. This includes a series of philosophical principles that underpin the concept of digital twins, or the process of *digital twinning*. These principles are then used to suggest answers to four key questions relating to digital twins. Finally Conclusions are drawn in Section 5.

2 Philosophical context of digital twins

The origins of the twinning concept is usually attributed to the work of NASA during the Apollo programme, where physical duplicates were used (Rosen et al., 2015). 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 — to name just a few application examples. 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); Niederer 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).

Although there has been much discussion on the potential definitions relating to digital twins (see for example the review in Semeraro et al. (2021)) one area that has not received much attention is the philosophical underpinnings of digital twins[1]. We introduce this topic by first reviewing the philosophy of modelling as applied to a wide range of scientific and engineering domains. Models are very important for digital twins because they are one of the key *components* that make up a digital twin. In addition, many of the techniques previously applied to models, such as verification, validation, calibration and prediction, are also functions that are often required of a digital twin.

The relationship between models and digital twins is a common source of confusion, and will be discussed in more detail in Section 4.2. For now we note that models and the philosophical concepts relating to models have a very important role for digital twins. Therefore we start with a selected introduction to the philosophy of modelling.

[1] Though see Korenhof et al. (2021) for one notable exception.

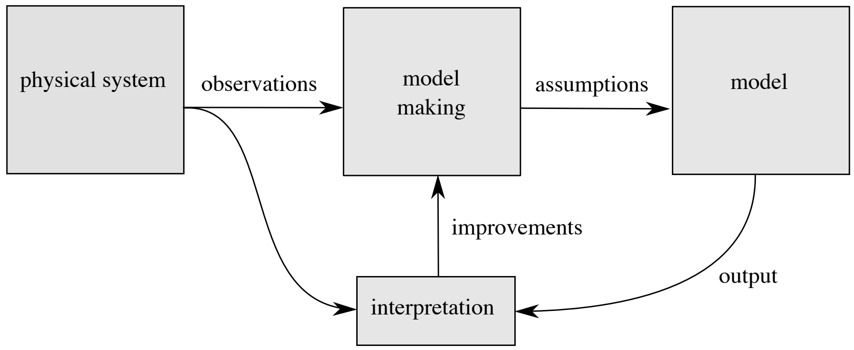

Figure 1: Schematic diagram showing the typical method of making a model of a physical system. The physical system can be a process or a material object

2.1 A selected introduction to the philosophy of modelling

In order to try and understand the philosophical context, we first consider the method via which a *model*[2] of a physical system is typically made in an engineering context. An example of a model making process is shown schematically in Figure 1. The first stage of the method in Figure 1 is to make observations from a physical system[3]. These observations are then used to make a model based on a set of assumptions. The subsequent output(s) from the model are then interpreted, and in many cases, this leads to improvements being made to the model, and the process is repeated as often as deemed necessary. There are several important philosophical viewpoints that can be understood in the context of Figure 1 that will be important for our later discussion on digital twins.

The first is *objectivism* which can be defined as the belief that there is an objective truth represented by the behaviour of the physical system that our model is trying to represent. In this worldview, the modeller and modelling process are separate from the physical system of interest, and do not have any direct influence on it. In addition, the “objective truth” is the same for all observers. Objectivism relates closely to the Newtonian scientific worldview, which will be described shortly.

In contrast, the idea of *subjectivism* is the belief that the process of making observations is not objective and actually any time an observation is made, then an interaction with the physical system takes place. This relates to the idea of *relativism*[4] 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).

Regardless of whether the objective or subjective view is taken, real-world physical systems are typically complicated[5], and can often be (conceptually) decomposed into “simpler" parts in order to be more effectively studied. The philosophy that supports this approach is the idea of *reductionism* which assumes that the physical system can be reduced to something simpler, and that by studying the reduced version, useful information about the complete system can be obtained (Heylighen et al., 2007).

For the purpose of this discussion we will consider two primary forms of reductionism. The first is

[2]The review is not intended to be comprehensive in any way, and readers are referred to Oberkampf and Roy (2010) for a comprehensive overview of modelling in the scientific and engineering domains. Note also that the word ‘model’ in this context is used to mean a representation 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*.

[3]This obviously assumes that there is already a physical systems in existence, which is often not 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 the physical system is available for observation.

[4]And also *perspectivism* (Giere, 2019).

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

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 to sub-components of the whole system after component-based reductionism has been carried out.

In Figure 1 the reductions are encoded as a set of assumptions. The reductionist philosophical approach has come to dominate scientific and engineering practices over time, and has been used to great effect. Reductionism has become associated with a Newtonian world-view, and 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).

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. There is also the interpretation that nature intends or prefers simplicity, which is related to the idea of *parsimony* which is discussed later.

However, it was not just Newton who contributed to what has become known as Newtonian (or classical) mechanics. There are (at least) two other key philosophical components that are important for our current discussion. First, the idea of a rules based *mechanistic* world view, leading to a set of "laws" that could be relied upon to apply "universally". A major contribution came from 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).

In addition to reduction (i.e. simplification) and laws, there is the idea of *determinism* which (has come to mean) that the state of something in the future can be determined entirely from it's current state. This idea is generally attributed to Laplace (see reprint: Laplace, 2012, first published 1795) who also made key contributions to the ideas of probability and "ignorance" (i.e. lack of knowledge), both of which we return to later. Classical mechanics has been built on these principles of a reductionist, mechanistic and deterministic approach, with huge success, and the plethora of models of this type are typically defined with a high degree of mathematical rigour.

However, it has also been long recognised that reductionist models, by definition, cannot capture the entire physical behaviour of the physical system, and the difference between a model output and an observation is described as the *error* or *uncertainty* related to the model — see for example Smith (2013). In particular, in the case when a deterministic model cannot capture the observed behaviour of the physical system, the model is often considered to be "missing" some important part of the physics. The missing knowledge is the *model inadequacy* of the reduced model, and is often defined as the *epistemic* uncertainty. In other words, this type of uncertainty represents the lack of knowledge (e.g. our ignorance of) of the real-world physical system.

In addition to this, observations of physical systems will always exhibit time varying fluctuations, and the more precisely one tries to make an observation, the greater (typically) these fluctuations grow. These fluctuations are often referred to as "noise" or "disturbances" and are typically considered to be an inherent part of the physics and observation processes[6]. Collectively they are known as *aleatory* uncertainty — see for example Hughes and Hase (2010).

In contrast to reductionism is *holism* where the physical system is not reduced, but treated as a whole. In the engineering domain, this philosophy has been developed primarily through the field of *systems engineering*[7] (and the related, overlapping subjects of cybernetics (see e.g. Ashby 1956), systems science (see e.g. Edson et al. 2016), operations research (see e.g. Hillier and Lieberman 2001), complexity science (see e.g. Waldrop 1993) and artificial intelligence (see e.g. Russell and Norvig 2010). Systems engineering uses a hierarchical landscape of systems including the "closed" (relatively simple) systems that can be modelled using deterministic (Newtonian) mathematical models such as those for the motions of point masses (e.g. like billiards), which we call classical (linear or nonlinear) dynamics. There is also the possibility of closed systems-of-systems, when many deterministic systems can interact with each other. Beyond this are "open" complex systems (such as living organisms or social systems) where closed, mechanistic models can fail to sufficiently represent the complexity of the underlying processes. Open systems, are critically dependent of their environment and the interactions they have with outside effects, such as other systems.

In the case of complex interacting systems, the occurrence of *emergent* behaviours can be induced by interactions between different parts of the overall system. Crucially, the emergent behaviours cannot typically be anticipated just from a knowledge of the parts of the system. Often such interacting systems contain intricate hierarchies or interdependencies, and emergence (e.g. self-organisation) can happen within a part, or across the entire system (see e.g. Bedau and Humphreys 2008). Systems engineering recognises that systems exist within an *environment*, and that systems can interact with each other to create systems-of-systems. An ongoing challenge with the systems engineering approach is how to represent the boundary between the system and its environment or other systems.[8] For instance, the choice of where to draw boundaries around a system is inherently a framing decision which influences what is included and excluded from analysis. This framing can be influenced by various values and cognitive biases, such as the researcher's theoretical commitments, interests, and goals (e.g. the perceived value of knowledge or insights to be derived). We will say more about this later in the paper. Emergent behaviour has also been studied using *complexity theory*[9] which (typically) uses coupled systems of dynamic models acting as "agents" to create models of emergent behaviours (typically in a deterministic sense — see for example Jensen 2022), although a non-deterministic framework can also be adopted[10].

More broadly, the role of humans is important in this discussion (more details are given in Section 3.5), because humans are imperfect, and they make decisions and assumptions based on their own worldview (or lens). Furthermore, communities and organisations can adopt and develop their own biases based on a range of factors, and perpetuate these over long periods of time[11]. Often this results in current views and approaches that have been "inherited" from previous generations without being appropriately examined (a phenomena known as *social learning* (Hoppitt and Laland, 2013)). For example,

[6]In control engineering, the a disturbance is considered to be an input that acts on the underlying dynamic system defining the plant, and noise is the unwanted corruption of the measured output signal.

[7]For a discussion of how systems engineering relates to the digital twin concept see Madni et al. (2019). Recent discussions regarding the principles and hypotheses that underpin systems engineering can be found in Watson (2019); Watson et al. (2019).

[8]This challenge is one area in which societal values enter into scientific theorising and engineering design choices.Longino (1990)

[9]An excellent overview of the philosophical relationships between classical reductionist science, systems thinking and complexity theory is given in Heylighen et al. (2007). The topic is discussed further in Section 3.

[10]Note the terminology of "agents" and "environment" used here is not to be confused with deep reinforcement learning, which also uses this terminology — see for example Graesser and Keng (2019).

[11]The interplay between science and society, and the roles of both scientists and laypersons is also relevant here — see for example Kitcher (2011).

in science and engineering, the type of model iteration process shown in Figure 1 happens repeatedly, often over many years, or even decades and involves multiple humans during that time. It is perfectly possible for people to be working on a model for which they did not do any model making, and therefore be unaware of the philosophical approach 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 "standardised" measurement scale over another).

In many domains (engineering being one) the separation of practitioners from the model making process (and the associated assumptions) is increasingly the case as modelling becomes more frequently integrated into sophisticated software tools. This leads to the obvious questions of (i) is the model being computed correctly in the software? (verification) and (ii) does the model output correctly represent the behaviour of the observed physical system? (validation)[12]. These are important questions that naturally can be extended to digital twins, and to which we will return later in Section 3.5.

2.2 The role of knowledge in model making

It will be argued here that a primary useful purpose of a model is to gain (or enhance, extend, and/or clarify) knowledge. Once additional knowledge is gained, it can be used (exploited) to create *value*, for example by supporting decisions. For the purpose of this discussion we focus on the process of gaining knowledge rather than exploiting it. We return to the topic of decision making in Section 5.3.

There are a multiple theories of knowledge relating to science that have been developed over many centuries[13]. One of the most important is the idea of *empiricism* which is the epistemological idea that knowledge can only be obtained by physical observations (e.g. the sensory experience of the observer). In the scientific context, during the 19th Century, empiricism led to what has become known as the *scientific method* where physical observations are made and used to test a specific *hypothesis*, primarily using statistical models to assess whether the hypothesis could be proven or not (see e.g. Lehmann et al. (1986)).

An example of a this type of hypothesis-based model making process is shown schematically in Figure 2. Here the model-making process of Figure 1 is expanded to show (very simplistically) the role of knowledge in this type of model making. The first step is to establish a research question. Next (or often as part of the first step) a review of existing knowledge that is relevant to the research question is carried out. From there, a testable hypothesis needs to be created, after which an experiment that can actually test the hypothesis needs to be designed and then performed. After the experimental outputs have been reviewed, the process can be improved. All these steps require expert knowledge of the specific application and the wider context. Multiple judgements are needed if the results are to be of use at the end of the process.

The scientific method is still used extensively, but two significant developments in the 20th Century have had a major influence[14]. Firstly, following the development of quantum mechanics, the philosophy of science underwent a major shift in perspective (see for example, the discussion and references in Chapter 2 of Oberkampf and Roy (2010)), resulting in far less certainty of what can be "proven"[15]. Secondly, the 20th Century saw the development of computational power that has given birth to high-

[12]These terms are sometimes defined the opposite way around, but generically the abbreviation "V& V" means carrying our the processes of model verification and validation.

[13]Epistemology is the philosophical study of knowledge theories. Although there is a longstanding tradition of relating scientific research to a relevant epistemology, there has been, until very recently, been almost no equivalent in engineering (Edson et al., 2016; Van de Poel and Goldberg, 2010).

[14]There is at least one other major factor, which is the development of artificial intelligence, but we will come to that later.

[15]There are also significant questions relating to "truth", induction and inference (Kuhn, 2012; Popper, 2014) that we will not discuss in detail here.

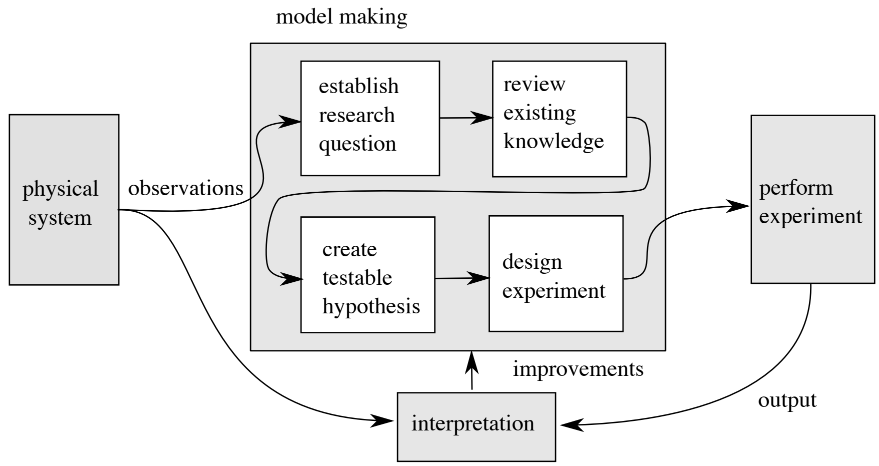

Figure 2: Schematic diagram showing the hypothesis-based method of making a model of a physical system. The physical system can be a process or a material object

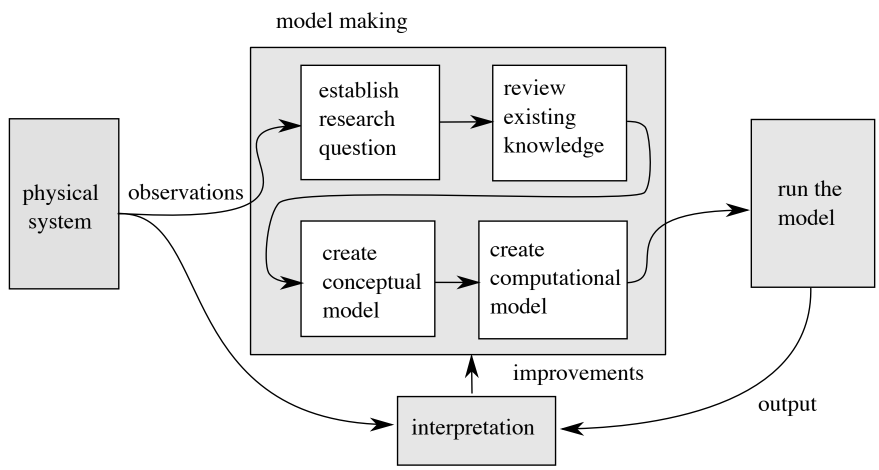

Figure 3: Schematic diagram showing the typical method of making a model of a physical system for a computational model. The physical system can be a process or a material object

powered software models that have surpassed all previous human capacities to simulate the physical world.

Many computational models are not created with an explicit hypothesis to test, and this variant of the model making process is shown in Figure 3. Here, the research question and knowledge review leads to a conceptual model, from which a computational model is derived. Throughout this process, assumptions will need to be made in order to define the precise form of the computational model.

There are several important points that can be understood from the processes shown in Figures 2 and 3 that will be important for our later discussion on digital twins. In particular, the role of knowledge and expertise in the process of creating the model and making the associated assumptions, a topic we return to later.

2.3 Defining a philosophical purpose for a model

In 1982, British statistician George Box published the now famous adage, *"all models are wrong, some are useful"*. The context for this comment comes from a discussion regarding the level of validation a model can have when compared to the real world system (Vining, 2013). Box's main point is that no (statistical) model can ever be "correct" in the sense that there is a "perfect" match with the physical system. The principle is key to understanding the limitation of all theoretical and computational models.

Box's statement also introduces the idea of model usefulness (or utility[16]). In addition, Box's statement implies that models can have a *useful purpose* even though they can never be perfect. In Section 2.2 above we asserted that the key useful purpose is to gain knowledge[17]. We will expand on this idea shortly, after first considering another important concept, captured in a quotation by Harlow Shapley, the American astronomer, who said that *"No one trusts a model except the man who wrote it; everyone trusts an observation except the man who made it."* This quote from Shapley introduces the idea of *trust* which is intrinsically linked to uncertainty. Shapley's quote also captures two *human biases*. First, trust in observations over a model (e.g. most people always assume an observation is more likely to be 'true' than a simulated model, even if they have no knowledge of how the observation was recorded or its closeness to the 'true values'). And, second, the difference between model makers (and data collectors) and users (e.g. the idea that a proponent of a new model (or theory) are likely to be biased in over-stating the value and fidelity of that new model, compared to those collecting data via observations, who in general terms, are assumed to be unbiased seekers of real-world 'truth'). It will be key to our later discussion to understand why trust in models might be as important as usefulness.

Returning to Box's main idea, what is the "useful purpose" of a model? It will be argued here that a primary useful purpose of a model is to gain (or enhance, extend, and/or clarify) knowledge. Furthermore, the statements from Box and Shapley are key to understanding the limitations of *all* theoretical and computational models. In the authors' opinion, these limitations are broadly aligned to the idea of "model dependent realism" expressed by Hawking & Mlodinow's Grand Design (Hawking and Mlodinow, 2010).

The model dependent realism philosophy essential says that absolutely certainty is an impossible goal, and therefore the most important thing is the usefulness of the model. Hawking & Mlodinow also say:

> "Model dependent realism short-circuits all this argument and discussion between the realist and anti-realist schools of thought. 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).

In the context of digital twins, described later, the usefulness will be particularly important in terms of *explanatory capability.* Therefore, for both models and digital twins, we will contend that the primary useful purpose is to gain knowledge/insights that will ultimately lead to explanatory capability[18]. We also acknowledge (following Shapley) that if utility is the primary criteria, then unbiased and trustworthy models (and digital twins) are crucial secondary requirements, in order to gain this new knowledge and insight. Therefore the claim made here is that *utility, trust and insight* are the three key *generic* requirements (or properties) of models that we would like to extend to digital twins[19].

But what about other important characteristics like fidelity, parsimony, cost or optimality? We argue here, that these characteristics will depend on the specific *context* of the model (or digital twin). Here, context means the specific application, objectives and other details relating the the physical system under consideration. It is important to bear in mind that our discussion here is ultimately aimed at creating digital twins that are *not* models (because of object-property inheritance — further explanation given later). Despite that, the characteristics like fidelity, parsimony, cost, tractability, or optimality will be considered to be context dependent, whereas utility, trust and insight are *generic*.

[16]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).

[17]As mentioned in Section 2.2, typically the *value* may come from how the additional knowledge is exploited, e.g. by supporting decisions. But the gain in knowledge is required first.

[18]Remembering that we are not including the exploitation of the gained knowledge in the purpose. That is considered to be an additional step.

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

Parsimony requires a further comment as it is often interpreted (following Newton's quote above) as "the natural order of things". Essentially the parsimony principle for models means that a simpler model with fewer parameters is regarded as better than a 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 favour 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[20]. The relationship between parsimony and purpose will have important consequences for digital twins that will be discussed in Section 4.2.

We finish this section, by noting that until relatively recently engineering has had no equivalent philosophical epistemological foundations, such as those that have developed for science — see for example, discussions in Vincenti et al. (1990); Bucciarelli (2003); Van de Poel and Goldberg (2010). Engineers use models extensively, but their use has developed as a series of overlapping *practices* associated with other functions such as the design, creation, testing, management, operation and decommissioning of engineering applications. Any associated philosophical implications have primarily been considered in terms of ethics, risk and safety — see for example Blockley (1980); Vincenti et al. (1990); Mitcham (1998); Martin and Schinzinger (2008) — and not the philosophy of modelling itself. As a result, engineers tend not to be trained and educated to consider the philosophy of modelling or the philosophy of decision-making — something that will be explored in more detail in Section 3.5.

We will now take a more detailed look at topics relevant to digital twins for engineering applications.

3 Complexity in engineering systems

> "Having been deeply enamoured of physics and reductionist goals, I was going through my own antireductionist epiphany, realising that not only did current-day physics have little, if anything, to say on the subject of intelligence but that even neuroscience, which actually focused on those brain cells, had very little understanding of how thinking actually arises from brain activity." — Melanie Mitchell, from Complexity: A guided tour (Mitchell, 2009).

The reductionist philosophy described in Section 2.1 is strongly associated with *classical mechanics* as initiated by Descartes, Newton, Laplace and many others since. In this (Newtonian) world view, any physical process is reduced by division (and/or other simplification) until a deterministic, mechanistic model can be used to explain it's behaviour. If a model cannot be found, then the division is applied again, and the logic is that eventually one reaches particles (the concept of indivisible atoms as defined by Greek philosophers) that it was once thought, (before quantum mechanics) could no longer be divided.

The concept of division in classical mechanics is based on the division of *material*, and so we say that the associated *ontology*[21] is materialistic (e.g. related to physical matter)[22] (Heylighen et al., 2007).

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

[21]Ontology (a central part of metaphysics) is the branch of philosophy which examines the fundamental categories of things. The other relevant branches related to engineering are epistemology and value theory (e.g. aesthetics or ethics).

[22]The idea that some things, like the human mind, are *non-physical* extended back to the ideas of Greek philosophers, and *meta-physics* has become the established as the study of non-material phenomena. More specifically related to the human mind, Descartes developed the idea of *mind-body dualism*.

Classical mechanics broadly developed into the study of solids, liquids and gases, with more advanced fields evolving to cover phenomena related to topics such as thermodynamics and electromagnetism etc. The overall ethos is based on finding simplicity when modelling apparently complicated physical processes e.g. the basic premise of creating parsimonious models. This approach works particularly well for *ordered systems*, such as materials with lattice-like structures, or the dynamics of point-mass systems. In these cases the behaviours can be encoded into a set of deterministic "laws", as envisaged by Descartes.

The classical mechanics approach couldn't be applied in the same way to *disordered systems*, such as a gas, consisting of molecules that act without any apparent constraints. To cope with this apparent disorder, *statistical mechanics* was developed whereby small particles (such as molecules in a gas) can be treated statistically with probability theory and related techniques. This allowed for average behaviours to be modelled, based on some basic assumptions about the independence of each particle and the nominally identical nature of the associated probability (see the early contribution of Laplace 2012). Using these simplifying assumptions allowed disordered systems to be treated within an essentially mechanistic modelling framework as well[23].

In a reductionist world-view, processes that could not be readily reduced were thought to be made up from many coupled-together simpler systems. In that sense they were still thought to be mechanistic and deterministic from the objectivist view point. Any observable complicated behaviour was believed to be explainable in terms of the underlying coupled systems[24].

Large numbers of coupled oscillator systems would fit into this category of systems, and they would be considered to be a *closed system* from a systems engineering perspective[25]. It is interesting to note, that as computational tools became available in the 20th Century, the Fermi-Pasta-Ulam-Tsingou paradox readily demonstrated how assumptions about a mechanistic model of many coupled oscillators did not necessarily exhibit the behaviour the model makers expected — and the subtleties of such systems are still being discussed and expanded on to the present day — see for example Weissert (1999); Berman and Izrailev (2005); Dauxois (2008).

Computational tools also led to the exploration of deterministic systems that contained *nonlinear* behaviours, using numerical approximations, because the majority of nonlinear systems could not be solved exactly. As Fermi and collaborators had discovered, the ability to compute numerical approximations of these systems led to the exploration of multiple new phenomena, and was a major contributory factor to the subsequent expansion of the field of *dynamical systems theory* (that during the 1970s and 80s became known as *chaos theory*) (e.g. see Hirsch and Smale 1974; Guckenheimer and Holmes 1983; Moon 1987; Glendinning 1994; Thompson and Stewart 2002; Strogatz 2019 and references therein).

The nonlinear systems that exhibited chaotic oscillations, and other related phenomena were entirely deterministic and often quite straightforward to write down mathematically. They exhibited interesting new behaviours, such as sensitivity to the initial starting conditions of the system (*the butterfly effect* (Hall, 1992)), and *bifurcations* (Kuznetsov, 2004; Haragus and Iooss, 2010), all of which were eventually explained in a rigorous, deterministic mathematical framework (which continues to be expended even today), but raised awkward philosophical questions about our ability to make predictions.

Many others in the early and mid 20th Century were identifying that complex behaviours occurred in a diverse range of applications including those that were long established, such as life sciences (see e.g. Weaver (1948)), and those that were just forming 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)). All of these topics would become large fields of research in their own

[23]The pioneering work of L. Boltzmann, J. Clerk Maxwell & J. W. Gibbs was key to the development of this field, and a modern introduction to the topic can be found, for example, in Pathria and Beale (2011).

[24]Or for continua, by reducing the *order* of the model needed to represent the behaviour

[25]As opposed to *open systems* such as those in biology that interacted with their surrounding environment.

right, but they also all had a relationship with the current fields of *complexity science*[26], which primarily focuses on emergent and adaptive behaviours (see e.g. Waldrop 1993; Mitchell 2009; Jensen 2022), and *systems research* which is focused on managing large-scale socio-technical systems — see for example Meadows (2008)[27]. The interesting distinction is that complexity science has typically focused on, and made extensive use of, *deterministic* agents or network models[28], whereas large-scale socio-technical systems considered in systems research, cannot (in general) be treated in such a deterministic framework (and are therefore often not even discussed)[29].

Within this general context, it is possible to distinguish between different categories of system based on linear vs nonlinear, ordered vs disordered, deterministic vs non-deterministic[30], reduced vs holistic, etc, and combinations of these categories. Here we will adopt the broad distinction that *complex* relates to a system which can have emergent behaviour whereas *complicated* relates to a system that is not "simple" but does not have interacting components that could lead to emergent behaviours — see discussion in Grieves and Vickers (2017). These distinctions will become important when we set up the framework for creating a digital twin, and complexity techniques are already being promoted for digital twins of cities (see Rozenblat and Fernández-Villacanas 2023; Caldarelli et al. 2023).

3.1 Types of complexity in engineering systems

> "Engineering is the art of modelling materials we do not wholly understand, into shapes we cannot precisely analyse, so as 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.

Having (briefly) described the development of multiple related fields, including complexity and systems science, and distinguished the difference between complex and complicated, we now consider what types of complexity occur 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 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's typical for engineering applications to have multiple types of complexity contained within it 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, modelled 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 out and consider each type independently. Usually this is mapped onto our siloed (e.g. reduced) set of

[26]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).

[27]We will discuss the related field of *systems engineering* further in Section 3.2.

[28]Another related subfield is that of *complex networks* or *network science*, which we do not discuss explicitly here, but a historical and philosophical introduction can be found in Baker (2013).

[29]Although there are overlapping concepts relating to phenomena such as hierarchies, feedbacks, interconnections and delays.

[30]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.

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; geographical effects
Geometric	Multiple compliments of varied shape & geometries; joints and jointing between components; mechanisms & interactions
Material	The physical & chemical properties of matter; combined & composite materials; wear, ageing & damage
Behavioural	Mechanistic behaviour of solids & fluids; vibrations & time-dependent behaviours; emergent behaviour; multi-physics; length-scales
Operational	Control & feedback; updates & changes; faults & failures; networks & connectivity; computational hardware & software
Computational	deterministic vs non-deterministic; time & memory requirements; processing resources; data size & formats, Kolmogorov complexity
Processes	Design; decisions & interventions; sequencing & workflow; human behaviour; communications; heuristics
Organisational	Structure & hierarchies; practices & organisation culture; rewards & incentives
Social	Attitudes; motivations; culture; education level; religion; beliefs; gender etc.

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 enquiry, where the focus is on understanding and explaining the behaviour 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[31]. In order to try to address some of the related challenges, the field of *systems engineering* has developed some useful methodologies, which we discuss next.

3.2 Systems engineering

> "Engineering is not merely knowing and being knowledgeable, like a walking encyclopaedia; engineering is not merely analysis; engineering is not merely the possession of the capacity to get elegant solutions to non-existent engineering problems; engineering is practicing the art of the organised forcing of technological change... Engineers operate at the interface between science and society" — Gordon S. Brown, Dean of the School of Engineering, MIT. 1959 – 1968.

Systems engineering was developed during the 20th Century alongside the related other fields of systems research and complexity already described above[32] (Schlager, 1956). 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). NASA and the space programme was undoubtedly a major influence in the development of systems engineering, and continues to be a driving force for the further development of the topic (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.

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

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

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 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 which 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/optimised whereas complex systems typically can’t.

The systems engineering community has given a considerable amount of time and thought into the philosophical and pragmatic frameworks needed to deal with complex/complicated engineering applications. For example, in recent papers (Watson, 2019; Watson et al., 2019) 15 principles and 3 hypothesis for systems engineering were articulated. The three hypotheses given in Watson et al. (2019) are:

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 Sect. 4.2, in the context of digital twins, but to mention just briefly, H1 relates to the concept 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[33]. Lastly, H3 is anticipating the stakeholders preference for quantitative solutions.

Other important concepts that are emphasised 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)).

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 (Grieves, 2005), expresses reservations about treating everything as a process;

> “We like to think that what we do in our organisations is process. Under systems theory, 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 — Micheal Grieves (Grieves, 2005), from Product Lifecycle Management”. 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*[34] instead of optimisation (Grieves, 2005).

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

[34] 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.

The broader point is that engineering contains some form of "art" (as alluded to by both A. R. Dykes and Gordon S. Brown in the quotes above) typically encoded in the form of attributes like *engineering judgements* and *design choices*[35]. 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)[36].

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's certainly true that the systems engineering approach favours defining multiple processes with associated inputs and outputs, and that in itself could be an over-constraining structural format for some applications. It's also true that the role of reductionism and deterministic modelling 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 recognise that[37].

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).

3.3 Emergent behaviours

> "What does this mean? That the essential reality of a system is indescribable?...Or does it mean, as it seems to me, that we must accept the idea that reality is only interaction?" — Carlo Rovelli (Rovelli, 2016)

In the sections above, we have already mentioned the importance of emergent behaviours. The quotation from Rovelli emphasises the importance of interactions in the context of quantum physics. In the context of digital twins, the basic idea is to join components together to *reconstruct* the dynamical behaviour of the combined system. The simplest case is joining two components — and a detailed example will be shown in Section 4.3 (the example results are shown in Fig. 7)

In engineering we make extensive use of numerical simulation tools that essentially break up complex geometries and behaviours into an assemblage of simpler elements for which the behaviour can be defined. These techniques, such as the finite element method, have evolved into sophisticated tools that are widely used to simulate the behaviour of complex/complicated systems that cannot be captured using simpler modelling techniques (see e.g. Crisfield 1997). The outputs from element-based methods are, in fact, emergent behaviours. This usually relates to field quantities such as stress, displacement, flow rate or temperature which are approximated as a form of "self-organisation" between the elements, acting within the overall element-based model. Essentially, the overall behaviour arises from local interactions between the multiple elements. This is a type of time dependent emergent behaviour, generally considered to be a subset of *evolutionary dynamics* (Jensen, 2022).

[35] 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."

[36] 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).

[37] 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).

In addition to self-organisation, there are other types of emergent behaviour, and multiple authors have described how the various types might be categorised — 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 behaviours range from relatively simple types, such as self-organisation and synchronisation, (Jensen, 2022), through to *evolutionary* forms of emergence (Kauffman, 2000). The ability to make predictions for emergent behaviours is a significant capability that is seen as a very desirable functionality (Gershenson, 2013), including for digital twins. We will not spend time more time describing types of emergent behaviours. Instead we are more interested in how digital twins might be expected to produce such behaviours, especially for very complicated applications, something considered in Section 4.3.

3.4 Artificial Intelligence, and other methods for dealing with complexity and uncertainty

> "Early AI was mainly based on logic. You're trying to make computers that reason like people. The second route is from biology: You're trying to make computers that can perceive and act and adapt like animals." — Geoffrey Hinton

The roots of artificial intelligence (AI), as Geoffrey Hinton's quote says, can be found in the development of formal logical methods and the early attempts to create mechanical computation machines[38]. Developments by Alan Turing and others during the second world war (Turing, 1950) were the catalyst for the current incarnation of the field, and the name *artificial intelligence* came from a meeting at Dartmouth in 1956 organised by John McCarthy and colleagues.

The 20th Century saw multiple parallel developments of AI based on, for example, Turing machines, and computational complexity (Li et al., 2008), biologically inspired natural computing techniques (Worden et al., 2011), symbolic AI (Dingli and Farrugia, 2023), pattern recognition & machine learning (Bishop, 2006), and multiple other fields, including the recent development of large language models such as ChatGPT (Teubner et al., 2023).

The quest for 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 text book 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 (Kiran et al., 2021). Sequential decision-making problems are also highly relevant to digital twins, which by their nature are time evolving, and will be discussed in more detail later in this chapter. 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).

[38] We will give only the briefest sketch of the history here. For a more in-depth historical review, the interested reader should consult the wider literature including, for example Nilsson (2009); Russell and Norvig (2010); Haenlein and Kaplan (2019); Marcus (2020).

Another way to (very) broadly categorise different aspect of AI research and innovation is in terms of:

1. Symbolic AI, such as logical reasoning, knowledge models and expert systems (Krishnamoorthy and Rajeev, 2018)

2. Sub-symbolic AI (connectionism), which includes all types of machine learning (ML) (Bishop, 2006; Kelleher et al., 2020)

3. Neuro-symbolic AI, which is the fusion of the other two categories. (Marcus, 2003; Dingli and Farrugia, 2023; Garcez and Lamb, 2023)

In very general terms, it could be said that symbolic AI 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 neuro-symbolic AI[39] (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 twins 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's associated techniques[40]. 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 large part this is because of the *philosophical breadth* of the topic — something which hopefully is described by the preceding discussion[41].

3.5 Human interpretations and bias

> "Efficient learning requires an open mind. To be open-minded means you don't cast out new information before evaluating it and if it's useful, making an honest attempt to incorporate it into your present way of thinking. But beware! Few people actually admit to being or feeling close-minded. The ego doesn't allow that. We trick ourselves into thinking we are objective and open, when in fact we may be judgemental and closed." — Arno Linger (Linger, 2006).

It was already mentioned in Section 1, that digital twins have been open to a very wide range of interpretations and some hype, often causing confusion, frustration and scepticism regarding their value. We now return to this theme to consider the human interpretations related to digital twins, and the associated biases that often occur.

Firstly, human interpretations are problematic, and it is difficult for us to be objective when constructing models and interpreting the results. The discussion above has already described multiple different philosophical viewpoints. Humans tend to adopt worldviews that suits them, and we are all subject to

[39]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).

[40]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.

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

confirmation bias. In fact researchers are particularly susceptible as we are often searching for or selecting data and evidence that confirms and supports our specific ideas. In addition to that, 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).

In addition to this, all teams, groups and communities are subject to negative group dynamic effects. A lack of diversity and inclusivity combined with entrenchment and group think can exacerbate negative views of other groups, and their associated philosophies even further. 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).

It was already mentioned in Sect. 2.1 that there is a tendency for groups to adopt past assumptions without necessarily re-examining them. There are multiple other types of *philosophical tribalism* and dogmatic behaviour[42]. This can often impede the adoption of useful research methodologies. For example, individuals can become "locked-in" to a favoured methodology and fail to explore other potentially useful alternatives.

More broadly, 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) (Varga, 2018) would be more beneficial. This will be an important point for digital twins, where both quantitative and qualitative functions are often required.

3.6 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 modelling highly complex and uncertain applications. Some overlap with AI and other fields mentioned above, and others have developed their own areas of endeavour. For example (with just a few selected references) 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 modelling (Zhou, 2019), spatiotemporal modelling (Banerjee et al., 2014), agent-based modelling (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 optimisation methods (Boyd and Vandenberghe, 2004) to name just a few.

4 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 specific instances of digital twins. As noted by Machiavelli, introducing something new is fraught with potential difficulties, and we

[42]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).

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 analysed the dominant conceptualisations 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. 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 characterisation of purpose.

Specifically, we take *utility* to mean a 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 it's geometry, materials, the environment in which it is located or deployed etc. Digital twins should be able to capture the *heterogeneity* (e.g. the diverse nature of the representational states and processes), while also highlighting their differences in type or kind.

The functional requirements could be, for example, to support decisions, to learn patterns of behaviour, or to develop more efficient ways of operation. The attribute of (unbiased) *trust* is related to the *uncertainty* within the digital twin, and 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 behaviour 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. This is in part made more difficult because there are multiple communities involved, all of which have different philosophical cultures and disciplinary perspectives or assumptions (e.g. a tendency to favour one modelling technique over another due to familiarity with a contingent method or practice).

We know from philosophy that definitions can be challenging, and even today there are ongoing philosophical debates about the definitions of broad terminology — for example the distinctions and overlap between science, engineering and technology (Van de Poel and Goldberg, 2010). 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 philosophical principles for digital twins is proposed here, based in part on the discussion above.

These principles 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 as 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, which are bespoke to an individual physical twin, that can be both artefacts (objects) and/or processes within the *contextual setting*;
5. **Counter-parsimonious**, meaning not seeking simplicity for its own sake, instead aiming to reflect the *required level of complexity* — but may make use of parsimonious concepts, when appropriate;
6. **Reconstructivist**, meaning they aim to reconstruct (some or all of) the behaviour of a physical twin by assembling the components of the digital twin, including *emergent behaviours*; 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;
11. An **operational platform**, consisting of software, hardware & 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 such as validation and verification and/or error detection and correction[43] in order to account of relevant forms of uncertainty[44];
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 behaviour(s);

[43]see for example Chapter II of MacKay (2003).

[44]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*.

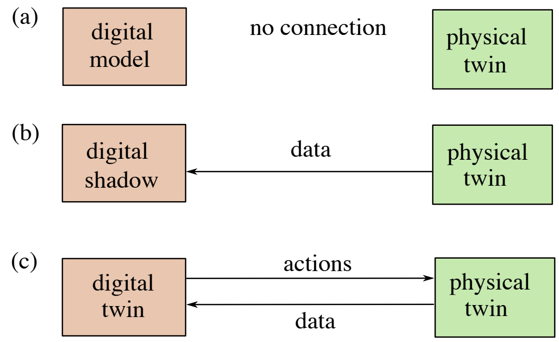

Figure 4: Schematic diagram showing (a) a digital model, (b) a digital shadow, and (c) a digital twin. Note that the digital twin together with the physical twin forms a ‘cyber-physical system’. It is in this digital/physical connection that the source of its value lies

19. **Learning** both from data (e.g. QoIs and outputs), and more broadly from the deployment of advanced techniques such as those from AI, statistic, dynamical systems etc;

20. **Insights** to be obtained that serve the purpose of the user, and maximise explainability and interpretability of the outputs; and

21. **Exploitation** of the insights to give **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. To summarise in a sentence, they capture the holism of systems engineering, emergent behaviours from complexity science, uncertainty analysis from statistics, time-evolution from dynamical systems theory, techniques from AI, control actions, and decision theories — amongst multiple other things! Our belief is that such a framework 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.

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 4 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 4;

1. Digital model — no connection between virtual and physical (Figure 4 (a)). This is the ‘traditional’ approach to modelling in science and engineering.

2. Digital shadow — data received from a connection (e.g. over a IoT network) with the physical twin is used to update and “shadow" the state of the physical twin (Figure 4 (b)). 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 4 (c))

(a)

Type of system behaviour

What is predicted

	Desired	Undesired
Predicted	Predicted Desired	Predicted Undesired
Unpredicted	Unpredicted Desired	Undesired Unpredicted

(b)

What could be known

What is known

	Knowns	Unknowns
Known	Known Knowns	Known Unknowns
Unknown	Unknown Knowns	Unknown Unknowns

Figure 5: Schematic diagram showing (a) how the outputs from a digital twin might be able to predict emergent behaviours proposed by Grieves and Vickers (2017), and (b) the "Rumsfeld" matrix.

The 21 principles set out above relate to digital twins, but the categories 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 4 (a) shows a model. Figure 4 (b) an updated model, and Figure 4 (c) a control system. For example, the explanation given in relation to Figure 4 has little or no sense of timing or mechanisms. E.g. 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 4 does not show (or even anticipate) connections between digital twins, via federation.

Furthermore, it is difficult to understand the ideas of holism, or emergent behaviour 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 4. This additional explanation relates to the use of models in digital twins, as we have described it 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 & Vickers have considered how the outputs from digital twins might be used to anticipate the types of emergent behaviours that might arise (Grieves and Vickers, 2017). They proposed a categorisation of outcomes for the digital twin that is shown in Fig. 5 (a). Here there are four categories of outcome that depend on what the digital twin predicts and whether the predicted behaviour 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 minimise the undesirable and unpredicted aspects as much as possible[45].

However, this approach also suffers from the problem of the need to know in advance what to include in the digital twin to get a desired outcome. As pointed out by Kauffman (2000) for example, this is a particular problem in the field of emergent behaviour. In fact, problems relating to prior knowledge are well known in other fields. For example, in the domain of uncertainty and risk management (e.g. Okashah and Goldwater (1994); Lanza (2000)). The "Rumsfeld" matrix[46] captures the key issue as shown in Fig. 5 (b).

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 don't know something at the present moment, then we cannot include it in our digital twin, and therefore (using this type of philosophical framework) we can never access the "unknown unknowns" category[47]. Knowing in advance, for example by prescribing a specific solution space, is a practical necessity for modelling, but will exclude the more advanced behaviours, particularly evolutionary forms of emergence — see for example Tononi et al. (2016); Kauffman (2000) and references therein.

To take one example, emergent behaviours are often modelled using multiple *agents* that interact with each other according to a predefined set of 'rules', typically relating to the environment and their nearest neighbouring 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 behaviours 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 won't 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 behaviours. Object-property inheritance can be interpreted as both related to *individual components (objects) in the digital twin, and relational combinations of the components.*[48] The relational combinations of the components are achieved using *dynamic assembly* — an example of which is described in the next section — all of which we assume is prescribed in advance.

Therefore, if a digital twin consists of n objects it would be expected to have a number (say d) of *directly inherited properties* which 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 behaviours, which are generated from the dynamic assembly. Note also that the combinatoric metric will depend on the specific context of the digital twin.

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

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

[47] This is the category which is associated with *black swans* Taleb (2007); Aven (2013).

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

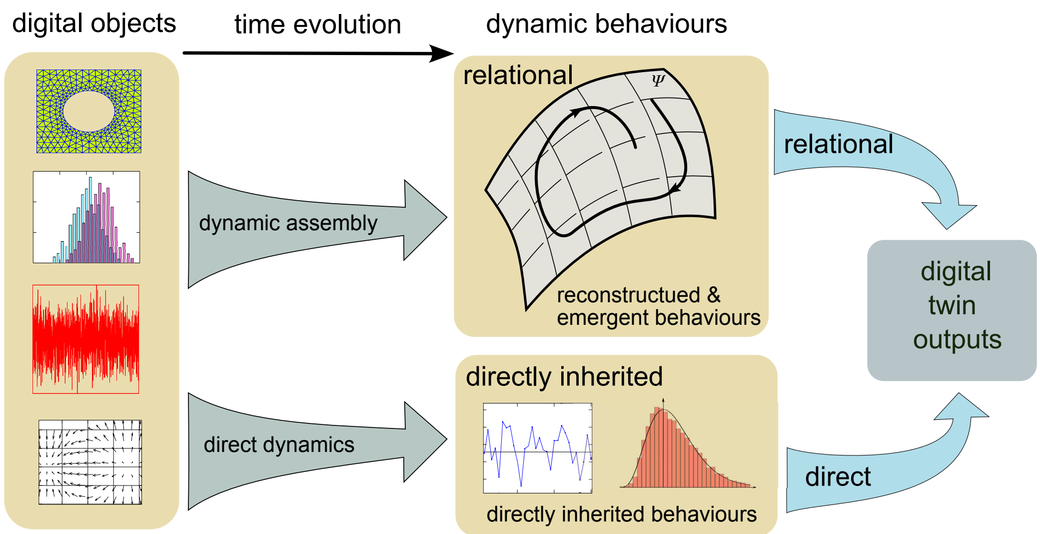

Figure 6: Schematic diagram showing how the outputs from a digital twin might be created using a series of digital objects (e.g. the components of the digital twin). The directly inherited properties come from each of the components are are grouped together. The relational properties, such as any reconstructed or emergent behaviours, come from the process of dynamic assembly. Both the directly inherited and relational properties can be used to form digital twin outputs

A simplified schematic example for a series of digital objects (e.g. components) is shown in Fig. 6, where dynamic assembly methods are used to obtain interactions between the components. In Fig. 6, 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 emphasise that all the emergent (and non-emergent) behaviours observed in digital twin outputs are contained in the categories of *known knowns*, *known unknowns*, and *unknown knowns*, shown in Fig. 5 (b). The *unknown unknowns*, shown in Fig. 5 (b) *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. 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.

4.3 How can emergent behaviours be predicted using a digital twin?

Emergent behaviour 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. 6. Dynamic assembly can be interpreted in several ways, but here we use the idea of creating "connectors" such that the resulting connections lead to interactions between the components with the aim of reconstructing emergent behaviour. To demonstrate this concept we show a specific example of a dynamic assembly method.

We will consider a digital twin to be made up of a series of digital *components* such as models, agents etc. The components will be combined in such a way that they *reconstruct* the time-evolving behaviour of the physical twin. In order to show how two models can be combined to reconstruct a behaviour, we

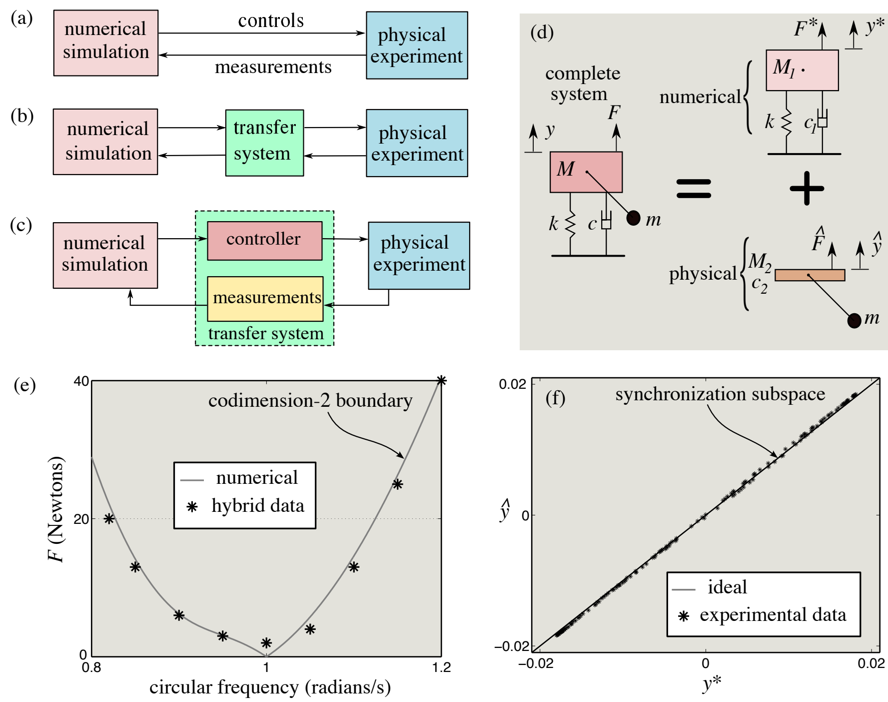

Figure 7: Schematic diagram showing (a) the underlying concept of hybrid simulation, (b) the use of a transfer system, and (c) the two main components in the transfer system. In (d) a mass-spring-damper-pendulum system is shown (labeled as “complete system” on the left of the subfigure”) which is decomposed (via component-based reductionism) into the addition of two subsystems (labelled as “numerical” and ”physical” on the right of the subfigure) Gonzalez-Buelga et al. (2005). The two subsystems are “virtually joined" at the pendulum pivot joint, using the controller to impose force equilibrium and compatibility of displacements. The output of the numerical subsystem is y^* and the output of the physical system is $\hat{y}$. The combined hybrid numerical-physical system is used to reconstruct a nonlinear codimenson-2 bifurcation boundary. This behaviour is only exhibited by the combined system, as shown in (e). The control algorithm is configured to ensure that $\hat{y}$ tracks y^*, and that if they are synchronised as closely as possible then the hybrid system will reconstruct the required behaviour to some level of fidelity. The “synchronisation subspace” for the test in subfigure (e) is shown in subfigure (f). For full details of these and related results see Gonzalez-Buelga et al. (2005); Kyrychko et al. (2006); Gonzalez-Buelga et al. (2007); Gawthrop et al. (2009).

consider the methodology used in the field of *hybrid simulation*. Note that we have deliberately chosen this method because, unlike agent-based-models (and related techniques) this enables *complicated heterogeneous components* (in this case two dynamical systems) to be joined together.

Hybrid simulation is a technique where a physical experiment and numerical simulation are combined using control and data acquisition hardware, typically in *real time* — see for example Wallace et al. (2005); Carrion (2007); Carrion et al. (2009); Chen and Ricles (2009); Gao et al. (2013); Tsokanas et al. (2021) and references therein[49]. The concept is shown schematically in Figure 7.

In Fig. 7 (a) the basic idea of hybrid simulation is shown, where a numerical simulation and a physical experiment are combined in real-time using control algorithms and measurements. In the case of most physical experiments, a *transfer system* is required to achieve this as shown in Fig. 7 (b) and (c)[50]. Note that in this case the transfer system is the connector. Furthermore, the objective in hybrid simulation is to get the transfer system to act like an *identity transformation* between the two systems being connected. In other words, the characteristics of the connector (transfer system) should not distort the interaction between the two models. In Fig. 7 (c) the two key parts of the transfer system are shown.

To consider a specific (although simple) example of hybrid simulation, in Fig. 7 (d) we show a mass-spring-damper-pendulum example originally developed in Gonzalez-Buelga et al. (2005). The complete system (on the left of the subfigure (d)) is the mass-spring-damper-pendulum system. The idea is that the nonlinear part, in this example the pendulum, is 'difficult' to model, and is therefore taken to be the physical experiment part (labelled as 'physical' on the right of subfigure (d))[51]. The remaining linear part, the mass-spring-damper is modelled numerically (labelled as 'numerical' on the right of subfigure (d)).

During the hybrid simulation, the output of the numerical system, y^*, is used as the setpoint in a control algorithm that controls the input to the physical system so that $\hat{y}$ tracks y^*. At the same time the physical force, $\hat{F}$, from the experiment is measured and feedback to be applied in the next computation of the numerical model. Delay compensation schemes are used to remove the effects of latency in the control and measurement hardware, and ensure that the numerical and physical systems are properly synchronised.

Minimising the error between $\hat{y}$ and y^*, shown in subfigure (f), allows the hybrid numerical-physical system to *reconstruct* dynamical behaviour of the complete system. In this example, the combined hybrid numerical-physical system is used to reconstruct a nonlinear codimenson-2 bifurcation boundary, as shown in subfigure (e).

In this type of hybrid simulation, we only numerically model the parts that are relatively easy (e.g. the linear part in the example). The physical part of the hybrid simulation (in this example the pendulum), is represented by the data from measurements used directly[52].

It is important to notice that the complete system output, y only "exists" during the hybrid simulation. Or in other words, complete system outputs only exist whilst the control algorithm is working to connect the two systems together (a process we call dynamic assembly) such that $\hat{y} \rightarrow y^* \rightarrow y$ (and $\hat{F} \rightarrow F^* \rightarrow F$). Without the control system connection, the output of the two systems would not dynamically assemble the output of the combined system (e.g. $\hat{y} \neq y^* \neq y$) — we refer back to the Carlo Rovelli quotation at the start of Section 3.3. As a result, emergent behaviours in digital twins will have the

[49]This and related techniques are also known by numerous other names such as hybrid testing, hardware-in-the-loop, real-time dynamic substructuring, and pseudo-dynamic testing.

[50]For some electrical and electronic engineering applications a transfer system is not used, and this is typically referred to as hardware-in-the-loop testing.

[51]Of course the simple pendulum is not that difficult to model, but the concept was developed for systems that have very complicated behaviours, like the failure of a concrete or masonry structural component during an earthquake.

[52]Although we do need to consider other effects of fidelity, such as any signal corruption that could occur in the measurement system.

property of *dynamically (or operationally) dependent outputs* meaning that the output(s) *only exist* whilst the digital twin is *operational*, e.g. meaning "live" or "switched on".

It should be noted that in some digital twin designs, models (or other digital objects) are either *selected* and/or *ensembled* together. For example in Edington et al. (2023), two models out of a choice of three were ensembled together with weighted coefficients. This type of ensembling (as it is currently implemented) *is not designed* to simulate emergent behaviours because there is no interaction between components in the digital twin. Furthermore, it should be noted that if the desired outputs are time dependent quantities of interest, such as velocities or accelerations, then the model interaction needs to be carried out in *real-time* in order to properly represent those variables[53].

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.[54] We have shown in the example from Fig. 7 that one example of an output is to choose a quantity of interest (QoI). In the study of differential equations, an important underlying concept is the idea of *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[55].

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 behaviour of the interconnections between components in the digital twin. It might be possible that existence and uniqueness of digital twin outputs (e.g. the *reconstructed* behaviours) could be assessed using 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

[53]There is a related field of pseudo-dynamic testing that operates at speeds less than real-time and estimates the dynamic variables — see for example Jung and Shing 2006.

[54]Note the similarity between the hypothesis H1 listed above in Sect. 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 many be many solutions, and therefore allowing for *nonuniqueness*.

[55]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.

meant by 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 behaviours — 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

5.1 Summary and conclusions

In this paper we have described the philosophical context for the digital twin concept. This began, in Section 2, with a selected introduction to the philosophy of modelling, and a discussion of the role of knowledge in model making. This selected introduction enabled us to consider how a philosophical purpose for a model could be defined, 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. Broadly speaking this philosophical approach aligns with the model dependent realism concept from Hawking & Mlodinow's Grand Design (Hawking and Mlodinow, 2010) — although some qualifications to this are given below.

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 world-view. In particular, we discussed the importance of modelling emergent behaviours, that cannot be captured in a reductionist paradigm. Importantly, it is *interactions* that lead to emergent behaviours, and these have to occur dynamically — depending on the exact context, we note that the environment might also influence the emergent behaviour.

Some of the limitations of systems engineering (overly reliant on input-output processes and black-boxes) and complexity science (mainly focuses on deterministic interactions of 'simple' agent-based models) prompted a review of the more recent role of artificial intelligence research. Here the development of 'intelligent' agents has been a distinguishing feature, with techniques such as reinforcement learning. Such intelligent agents have the capacity to both learn from data and take actions in real-world environments (although they also have limitations relating to reliance on learning over knowledge). It was also noted that symbolic-based AI methods, although currently out-of-fashion, offers the potential of combined learning and reasoning using so-called *neuro-symbolic* AI methods. The last part of our review was a discussion on human biases and the effects of phenomena such as confirmation bias and philosophical tribalism. This is often an under-rated, or even neglected, factor but we consider it to be highly relevant to the current context of digital twins.

In Section 4 of the paper we presented a proposed foundation of a philosophical framework for digital twins. This foundation consisted of 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 behaviours 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.

Next we draw together some further comments not captured directly in the other parts of the paper, which leads to some open research questions.

5.2 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 practice, adopting model dependent realism 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. like making observations). The issue of realism vs. 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 behaviour(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 world-view which is focused entirely on explaining the physical behaviour of the Universe we live in. It could be considered that the *“direction-of-fit"* is one-way. Or 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 contextualised using a more detailed philosophical analysis such as that proposed by Douglas (2013) which distinguishes between between internal consistency (a minimal criteria) and external consistency (an ideal desiderata, presuming a 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. This is an area for future research development.

2. *Emergence is counter-parsimonious:* As was described in Section 4.2, a digital twin will only be able to exhibit behaviours 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’s 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’s only capable of reinforcing your own (or an inherited) prejudicial expectation. The exact relationship between emergent behaviours 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 interoperation 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 in a systematic way.

4. *Validation of digital twins:* Some comments on 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[56]. 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. In the example (e.g. results in Fig. 7), the system was simple, and therefore we *knew in advance* what behaviour to expect, and could therefore validate the hybrid result quite easily[57] (e.g. the validation between a numerical computation and the hybrid system results is shown in Fig. 7 (e)). 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.

5. *Logic vs 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 philosopical discussion about how logic, learning and probability interrelate (see for example the discussion in Belle (2022)). This is an interesting topic, that we 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' vs 'bottom-up' approach to creating a digital twin?[58] 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.

Next we consider areas for future research development, that haven't already been mentioned above.

[56] 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).

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

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

5.3 Future directions for research

1. *Human-factors;* is a topic we have only mentioned very briefly in this paper. However, it is one of the most important for future understanding of digital twins. Broadly this area of research includes the topics of (i) the role of humans in designing and building digital twins (partially discussed in Sections 2.1 and 3.5), (ii) how human users interface with digital twins & act on the outputs they receive, and (iii) digital twins that include humans in some way, for example in medicine or social systems that include human behaviours. 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, DTs raised a (relatively) narrowly circumscribed set of ethical, legal, and societal issues (e.g. safety compliance). However, as DTs 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 (i.e. principles 19 and 20: *learning* and *insights* obtained from the DT) 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 DTs, 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 DTs within society (e.g. unintended behaviours caused by model drift, data privacy and security violations). There is a lot to explore here, and the presence of a robust conceptual framework could provide a systematic means for grounding and evaluating the myriad normative issues associated with DTs. For this reason, while we have acknowledged the existence of some of the associated issues (e.g. presence of bias), we leave this as a topic for future research.

3. *Methods for dynamic assembly:* In practical terms, one of the most interesting areas for future research is methods that enable dynamic assembly of the 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 behaviours. There are already techniques, such as agent-based modelling 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 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. In fact, the method used in Section 4.3 was ‘borrowed’ from another application domain, but essentially relied on real-time control to create the interaction. The scope for new developments in this area is significant.

4. *The role of knowledge:* relates to human factors listed above, but is so important that is warrants a separate discussion point. In particular the role of knowledge and insight, in supporting subsequent actions taken, such as decisions. One way in which 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 rigour what knowledge means in a DT context, particularly linking to topics such as knowledge representation, inference, model interpretability. See for example the discussions and examples in Pan et al. (2023); Olsson and Axelsson (2023); Akroyd et al.

(2021); Li et al. (2021). Much of the existing work relates to ontologies (e.g. see for example (Nguyen, 2022; Akroyd et al., 2021; Singh et al., 2020; West, 2011)), there is scope for more investigation based on the ideas of causality, and more general epistemology.

Acknowledgments The authors would like to acknowledge the support of The Alan Turing Institute Research and Innovation Cluster for Digital Twins (TIRC-DT). In addition, the authors would also like to thank the DTNet+ Leadership Team members, the TRIC-DT Co-Directors, Mark Girolami, Ben MacArther, Rebecca Ward, Tim Rogers, Nikos Dervilis, Matthew Bonney, Xiaoxue Shen, Ziad Gauch, Matthew Tipuric, Prajwal Devaraja, Saeid Taghizadeh for insightful discussions about these topics.

Funding Statement This work was funded by the Alan Turing Institute Research and Innovation Cluster for Digital Twins. The authors would also like to acknowledge the funding support of UKRI via the grants EP/Y016289/1 and EP/R006768/1.

Competing Interests The authors declare that they have no conflict of interest.

Data Availability Statement There is no new data or code associated with this current work.

Ethical Standards The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Author Contributions DW created the initial framework and drafted the initial version of this paper. All authors have contributed to the content, read and approved the final manuscript.

References

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

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

Adams, K. M. and Meyers, T. J. (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.

Area, I., Fernández, F. J., Nieto, J. J., and Tojo, F. A. F. (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, J. S., Kinard, D. A., Kraft, E. M., Marks, I., Matlik, J. F., Fischer, O. J. P., Sangid, M. D., Seal, D., Tucker, E., and Vickers, J. (2020). Digital twin: Definition & value—an aiaa and aia position paper. *AIAA: Reston, VA, USA*.

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman & Hall.

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.

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

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2014). *Hierarchical modeling and analysis for spatial data*. CRC Press.

Barber, D. (2012). *Bayesian reasoning and machine learning*. Cambridge University Press.

Barlow, J. S. (2002). *The cerebellum and adaptive control*. Cambridge.

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

Bedau, M. A. and Humphreys, P. (2008). *Emergence: Contemporary readings in philosophy and science*. MIT press.

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

Belle, V. (2022). Logic meets learning: From aristotle to neural networks. *Neuro-symbolic artificial intelligence: The state of the art*, 342:78.

Berman, G. and Izrailev, F. (2005). The fermi-pasta-ulam problem: fifty years of progress. *Chaos*, 15(1):15104.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning by Christopher M. Bishop*. Springer Science+ Business Media, LLC.

Blockley, D. I. (1980). *The nature of structural design and safety*. Ellis Horwood, Chichester.

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.

Botín-Sanabria, D. M., Mihaita, A.-S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., and Lozoya-Santos, J. d. J. (2022). Digital twin technology challenges and applications: A comprehensive review. *Remote Sensing*, 14(6):1335.

Boyd, S. P. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

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

Bucciarelli, L. (2003). *Engineering philosophy*. DUP Satellite; an imprint of Delft University Press.

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, J. L. (2023). The role of complexity for digital twins of cities. *Nature Computational Science*, pages 1–8.

Carrion, J. E. (2007). *Model-based strategies for real-time hybrid testing*. University of Illinois at Urbana-Champaign.

Carrion, J. E., Spencer, B., and Phillips, B. M. (2009). Real-time hybrid simulation for structural control performance assessment. *Earthquake Engineering and Engineering Vibration*, 8:481–492.

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.

Chen, C. and Ricles, J. M. (2009). Improving the inverse compensation method for real-time hybrid simulation through a dual compensation scheme. *Earthquake Engineering & Structural Dynamics*, 38(10):1237–1255.

Chouard, T. (2016). The Go files: AI computer clinches victory against Go champion. *Nature*.

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

Churchman, C. W., Ackoff, R. L., and Arnoff, E. L. (1957). *Introduction to operations research.* Wiley.

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, U. A., 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.

Crisfield, M. A. (1997). *Non-linear finite element analysis of solids and structures.Volume 2:Advanced topics*. John Wiley.

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.

Dale, K. I., Pope, E. C., Hopkinson, A. R., McCaie, T., and Lowe, J. A. (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 (togaf) and the us department of defense architecture framework (dodaf). *The Open Group*, 4.

Dauxois, T. (2008). Fermi, pasta, ulam, and a mysterious lady. *Physics today*, 61(1):55–57.

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

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

Dingli, A. and Farrugia, D. (2023). *Neuro-Symbolic AI: Design transparent and trustworthy systems that understand the world as you do*. Packt.

dos Santos, C. H., Montevechi, J. A. B., de Queiroz, J. A., 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.

Edington, L., Dervilis, N., Abdessalem, A. B., and Wagg, D. (2023). A time-evolving digital twin tool for engineering dynamics applications. *Mechanical Systems and Signal Processing*, 188:109971.

Edson, M. C., Henning, P. B., and Sankaran, S. (2016). *A guide to systems research: Philosophy, processes and practice*, volume 10. Springer.

Enders, M. R. and Hoßbach, N. (2019). Dimensions of digital twin applications - a literature review. In *Proceedings of Twenty-fifth Americas Conference on Information Systems*.

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

Estefan, J. A. 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.

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, M. H., Chiementin, X., Chaari, F., Bolaers, F., and Haddar, M. (2020). Digital twin-driven machine learning: ball bearings fault severity classification. *Measurement Science and Technology*.

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

Fradkov, A. L., Miroshnik, I. M., and Nikiforov, V. O. (1999). *Nonlinear and adaptive control of complex systems*. Kluwer: Dordrecht.

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

Gao, X., Castaneda, N., and Dyke, S. J. (2013). Real time hybrid simulation: from dynamic system, motion control to experimental error. *Earthquake engineering & structural dynamics*, 42(6):815–832.

Garcez, A. d. and Lamb, L. C. (2023). Neurosymbolic ai: The 3 rd wave. *Artificial Intelligence Review*, pages 1–20.

Gawthrop, P. J., Neild, S. A., Gonzalez-Buelga, A., and Wagg, D. J. (2009). Causality in real-time dynamic substructure testing. *Mechatronics*, 19(7):1105–1115.

Gelman, A., Carlin, J. B., Stern, H. S., and Dunson, D. B. (2014). *Bayesian data analysis*, volume 3. CRC press.

Gershenson, C. (2007). *Design and control of self-organizing systems*. PhD thesis.

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, 1 edition.

Giere, R. N. (2019). *Scientific perspectivism*. University of Chicago press.

Glendinning, P. (1994). *Stability, instability and chaos*. Cambridge University Press.

Gonzalez-Buelga, A., Wagg, D. J., and Neild, S. A. (2007). Parametric variation of a coupled pendulum-oscillator system using real-time dynamic substructuring. *Structural Control and Health Monitoring*, 14(7):991–1012.

Gonzalez-Buelga, A., Wagg, D. J., Wallace, M. I., Neild, S. A., and Macdonald, J. H. G. (2005). Testing an autoparametric pendulum using a hybrid numerical experimental method. In *Proceedings of the 1st International Conference on Advances in Experimental Structural Engineering*, volume 1, pages 417–424.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Graesser, L. and Keng, W. L. (2019). *Foundations of deep reinforcement learning*. Addison-Wesley Professional.

Greiner, W. (1994). *Quantum mechanics, an introduction*. Springer-Verlag: Berlin.

Grieves, M. (2005). *Product Lifecycle Management: Driving the next generation of lean thinking*. McGraw-Hill Professional.

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

Grieves, M. W. (2019). Virtually intelligent product systems: Digital and physical twins. In *Engineering: Theory & Practice*. American Institute of Aeronautics and Astronautics.

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.

Hall, N. (1992). *The new scientist guide to chaos*. (Penguin).

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Han, J. and Kamber, M. (2022). *Data mining: concepts and techniques*. Morgan kaufmann.

Han, Y., Niyato, D., Leung, C., Kim, D. I., 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.

Haragus, M. and Iooss, G. (2010). *Local bifurcations, center manifolds, and normal forms in infinite-dimensional dynamical systems*. Springer Science & Business Media.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Hawking, S. and Mlodinow, L. (2010). *The grand design*. Random House Digital, Inc.

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, C. G. (2007). Complexity and philosophy. In *Complexity, Science and Society*, pages 117–134. Radcliffe Publishing.

Hillier, F. S. and Lieberman, G. J. (2001). *Introduction to operations research*. McGraw hill Companies, Inc.

Hirsch, M. W. and Smale, S. (1974). *Differential equations, dynamical systems , and linear algebra*. Academic Press.

Hirshorn, S. R., Voss, L. D., and Bromley, L. K. (2017). *Nasa systems engineering handbook*. NASA.

Holland, O. T. (2007). Taxonomy for the modeling and simulation of emergent behavior systems. In *Proceedings of the 2007 spring simulation multiconference-Volume 2*, pages 28–35.

Holland, T. (2018). *Foundations for the modeling and simulation of emergent behavior systems*. CRS Press/Taylor and Francis, Boca Raton, FL, USA.

Hoppitt, W. and Laland, K. N. (2013). *Social learning: an introduction to mechanisms, methods, and models*. Princeton University Press.

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, J. P., ANe, C., Larget, B., and Ronquist, F. (2008). A bayesian perspective on a non-parsimonious parsimony model. *Systematic biology*, 57(3):406–419.

Hughes, A. J., Barthorpe, R. J., Dervilis, N., Farrar, C. R., 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*. OUP Oxford.

Ioannidis, J. P. (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*.

Jensen, H. J. (2022). *Complexity science: the study of emergence*. Cambridge University Press.

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.

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, A. J. (2000). *Game theory: Mathematical models of conflict*. Elsevier.

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*.

Jung, R. Y. and Shing, P. B. (2006). Performance evaluation of a real-time pseudodynamic test system. *Earthquake Engineering & Structural Dynamics*, 35(7):789–810.

Kapteyn, M. G., Knezevic, D. J., 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.

Kauffman, S. A. (2000). *Investigations*. Oxford University Press.

Kelleher, J. D., Mac Namee, B., and D'arcy, A. (2020). *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. MIT press.

King, A. C., Billingham, J., and Otto, S. R. (2003). *Differential equations*. Cambridge.

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

Kitcher, P. (2011). Science in a democratic society. In *Scientific realism and democratic society*, pages 95–112. Brill.

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.

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.

Kuersten, A. (2023). Health Digital Twins, Legal Liability, and Medical Practice. *The American Journal of Bioethics*, 23(9):66–69.

Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press.

Kutz, J. N. (2013). *Data Assimilation Methods*, pages 539–555. Oxford.

Kuznetsov, Y. A. (2004). *Elements of Applied Bifurcation Theory*. Springer.

Kyrychko, Y. N., Blyuss, K. B., Gonzalez-Buelga, A., Hogan, S. J., and Wagg, D. J. (2006). Real-time dynamic substructuring in a coupled oscillator-pendulum system. *Proceedings of the Royal Society A*, 462(2068):1271–1294.

Lanza, R. B. (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.

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

Lehmann, E. L., Romano, J. P., and Casella, G. (1986). *Testing statistical hypotheses*, volume 3. Springer.

Leslie, D. (2021). The Arc of the Data Scientific Universe. *Harvard Data Science Review*, 3(1). https://hdsr.mitpress.mit.edu/pub/ln5hu6vq.

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*, pages 1–18.

Li, M., Vitányi, P., et al. (2008). *An introduction to Kolmogorov complexity and its applications*, volume 3. Springer.

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. *arXiv preprint arXiv:2212.05937*.

Linger, A. (2006). *The Rock Warrior's Way*. Desiderata.

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

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, R. Y. (2021). A review of digital twin in product design and development. *Advanced Engineering Informatics*, 48:101297.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton university press.

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, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Madni, A. M., Madni, C. C., and Lucero, S. D. (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*.

Marcus, G. F. (2003). *The algebraic mind: Integrating connectionism and cognitive science*. MIT press.

Marques, T. A. (2021). On how statistics is used and abused to find truth in science. *Kairos. Journal of Philosophy & Science*.

Marsh, H. W. and Hau, K.-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The journal of experimental education*, 64(4):364–390.

Martin, M. W. and Schinzinger, R. (2008). *Engineering ethics*. McGraw-Hill.

McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

Meadows, D. H. (2008). *Thinking in systems: A primer*. chelsea green publishing.

Melesse, T. Y., 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*, pages 9–12.

Miller, A. M., 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.

Mitcham, C. (1998). The importance of philosophy to engineering. *Teorema: Revista internacional de filosofía*, pages 27–47.

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford university press.

Moon, F. C. (1987). *Chaotic vibrations: An introduction for applied scientists and engineers*. John Wiley: New York.

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*. London: Reg. Soc. Præss.

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

Niederer, S. A., Sacks, M. S., Girolami, M., and Willcox, K. (2021). Scaling digital twins from the artisanal to the industrial. *Nature Computational Science*, 1(5):313–320.

Nilsson, N. J. (2009). *The quest for artificial intelligence*. Cambridge University Press.

Oberkampf, W. L. and Roy, C. J. (2010). *Verification and validation in scientific computing*. Cambridge University Press.

Okashah, L. A. and Goldwater, P. M. (1994). Unknown unknowns: modeling unanticipated events. In *Proceedings of Winter Simulation Conference*, pages 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)*, pages 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. *arXiv preprint arXiv:2306.08302*.

Pathria, R. and Beale, P. D. (2011). *edition 3. Statistical mechanics*. Amsterdam. Elsevier.

Popa, E. O., 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.

Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*. routledge.

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

Ríos, J., Hernández, J. C., 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*, pages 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, K. D. (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.

Rozenblat, A. and Fernández-Villacanas, J. (2023). Complexity science for digital twins. *Preprint*.

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

Särkkä, S. (2013). *Bayesian filtering and smoothing*. Cambridge University Press.

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, K. J. (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.

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

Sepasgozar, S. M., Khan, A. A., Smith, K., Romero, J. G., 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, C. T., and Yeom, C. (2021). City digital twin potentials: A review and research agenda. *Sustainability*, 13(6):3386.

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

Silver, D., Huang, A., Maddison, C. J., 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, J. A., 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.

Smith, R. C. (2013). *Uncertainty quantification: theory, implementation, and applications*, volume 12. Siam.

Snow, C. P. (2012). *The two cultures*. Cambridge University Press.

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

Sozen, M. A. (2002). A way of thinking. *EERI Newsletter*, 36(4).

Strogatz, S. H. (2019). *Nonlinear Dynamics and Chaos*. Perseus Books Group.

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.

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, C. M., 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, B. D., Todd, M. D., 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.

Thompson, J. M. T. and Stewart, H. B. (2002). *Nonlinear dynamics and chaos*. John Wiley: Chichester.

Tigard, D. W. (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, A. S., Adhikari, S., and Chakraborty, S. (2023). Probabilistic machine learning based predictive and interpretable digital twin for dynamical systems. *Computers & Structures*, 281:107008.

Tsokanas, N., Wagg, D., and Stojadinovic, B. (2021). Robust model predictive control for dynamics compensation in real-time hybrid simulation. *Recent Advances and Applications of Hybrid Simulation*, pages 57–73.

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., and Spottswood, S. M. (2011). Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011.

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

Van de Poel, I. and Goldberg, D. E. (2010). *Philosophy and engineering: An emerging agenda*, volume 2. Springer Science & Business Media.

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

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

Vincenti, W. G. et al. (1990). *What engineers know and how they know it*, volume 141. Baltimore: Johns Hopkins University Press.

Vining, G. (2013). George box: Scholar, scientist, and statistician.

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

Wagg, D. J. and Neild, S. A. (2015). *Nonlinear vibration with control - For flexible and adaptive structures*. Springer, 2nd edition.

Wagg, D. J., Worden, K., Barthorpe, R. J., and Gardner, P. (2020). Digital Twins: State-of-the-Art and Future Directions for Modeling and Simulation in Engineering Dynamics Applications. *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, 6(3). 030901.

Walden, D. D., Roedler, G. J., Forsberg, K., Hamelin, R. D., and Shortell, T. M. (2015). *INCOSE systems engineering handbook*. John Wiley.

Waldrop, M. M. (1993). *Complexity: The emerging science at the edge of order and chaos*. Simon and Schuster.

Wallace, M., Wagg, D., and Neild, S. A. (2005). An adaptive polynomial based forward prediction algorithm for multi-actuator real-time dynamic substructuring. *Proceedings of the Royal Society of London A.*, 461(2064):3807–3826.

Wanasinghe, T. R., Wroblewski, L., Petersen, B., Gosine, R. G., James, L. A., De Silva, O., Mann, G. K., and Warrian, P. J. (2020). Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. *IEEE Access*.

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

Watson, M. D., 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.

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*, pages 49–64.

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

Weissert, T. (1999). *The genesis of simulation in dynamics: pursuing the Fermi-Pasta-Ulam problem*. Springer-Verlag New York, Inc.

West, M. (2011). *Developing high quality data models*. Elsevier.

Worden, K., Staszewski, W. J., and Hensman, J. J. (2011). Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111.

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.

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*.

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.