The Hitchhiker's Guide to Fused Twins: A Review of Access to Digital Twins In Situ in Smart Cities

Grübel, Jascha; Thrash, Tyler; Aguilar, Leonel; Gath-Morad, Michal; Chatain, Julia; Sumner, Robert W.; Hölscher, Christoph; Schinazi, Victor R.

Published in:
Remote Sensing

DOI:
10.3390/rs14133095

Licence:
CC BY

Link to output in Bond University research repository.

Recommended citation (APA):

Academic Editor: Gwanggil Jeon
Received: 20 April 2022
Accepted: 21 June 2022
Published: 27 June 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Smart Cities already surround us, and yet they are still incomprehensibly far from directly impacting everyday life. While current Smart Cities are often inaccessible, the experience of everyday citizens may be enhanced with a combination of the emerging technologies Digital Twins (DTs) and Situated Analytics. DTs represent their Physical Twin (PT) in the real world via models, simulations, (remotely) sensed data, context awareness, and interactions. However, interaction requires appropriate interfaces to address the complexity of the city. Ultimately, leveraging the potential of Smart Cities requires going beyond assembling the DT to be comprehensive and accessible. Situated Analytics allows for the anchoring of city information in its spatial context. We advance the concept of embedding the DT into the PT through Situated Analytics to form Fused Twins (FTs). This fusion allows access to data in the location that it is generated in in an embodied context that can make the data more understandable. Prototypes of FTs are rapidly emerging from different domains, but Smart Cities represent the context with the most potential for FTs in the future. This paper reviews DTs, Situated Analytics, and Smart Cities as the foundations of FTs. Regarding DTs, we define five components (physical, data, analytical, virtual, and Connection Environments) that we relate to several cognates (i.e., similar but different terms) from existing literature. Regarding Situated Analytics, we review the effects of user embodiment on cognition and cognitive load. Finally, we classify existing partial examples of FTs from the literature and address their construction from Augmented Reality, Geographic Information Systems, Building/City Information Models, and DTs and provide an overview of future directions.

Keywords: Digital Twin; Smart Cities; Internet of things; sensor networks; remote sensing; Fused Twins; immersive analytics; embodiment; visualisation

1. Introduction
Can Digital Twins (DTs) [1,2] really make cities smart? The pressure to transform cities into Smart Cities is mounting amongst urban policymakers worldwide [3]. Initially, the hope was that smartification would be instilled in cities by simply creating citywide Internet of Things (IoT) installations and remote sensing applications [4–6]. The idea behind this
thinking is simple yet alluring: to have everything in the city attached to the Internet and make it proclaim its own state as if it were a smart object [7] remotely observing everything [4,6]. Once observed and connected, everything should become measurable. In such a Smart City, every minute detail, if measured, can be included in the analysis. This would allow for the uncovering of new insights [8] and give stakeholders a deeper understanding of processes that were previously thought to be fully understood.

However, the complexity of the collected data, the observed processes, and the city itself requires a unifying view. A DT [9] provides a platform on which cities can become really smart by shifting attention from data collection to data use. Here, a digital representation of the city allows for modelling and predicting its ever-changing state via simulations [10–14]. Unfortunately, until now, the virtual and Physical Environments of cities remain insufficiently interconnected [15]. At best, DTs are remotely displayed in city administration offices to obtain insights [9,16]. However, DTs have the potential beyond administration to inform decision-making on all levels from citizens to politicians and planners [17]. To realise such a wide-ranging implementation of a DT for a Smart City, its data must be accessible in situ within its physical context. Situated visualisations [18,19] enable sensing and actuating in the environment with which it is combined.

We introduce the Fused Twins (FTs) (on a grammatical side note, FTs are always plural because they consist of both the Physical Twin and the Digital Twin and their interaction) paradigm (see Figure 1) to enable the conceptualisation of an accessible Smart City. We argue that there is a need to clarify the definitions of DTs, Situated Analytics, and Smart Cities with an emphasis on the visualisation and development of a practical interface. The FTs concept aims to embed the DT in its Physical Twin (PT), fusing them via Extended Reality (XR) applications [17]. Throughout this paper, we follow the working hypothesis that DTs combined with Situated Analytics enable lower cognitive load, enhanced cognition, and easier embodiment; and therefore, FTs can make Smart Cities more accessible to users ranging from managers and policy makers to citizens. Ultimately, only when data are accessible where it is most needed and understandable to those who consult it, can DTs fulfil the promise that IoT alone could not: to power Smart City developments by interlinking the physical and Virtual Environments.

![Figure 1. Comparison of a Digital Twin and Fused Twins.](image)

In both rows, the same lobby of a building is shown. (a) the DT of a sensor system with 390 nodes is visualised in Virtual Reality (VR) [20,21]. The white areas visualise the areas of observation and the coloured spheres encode the type of sensor that is located at the respective position in the building. (b) The FTs [17] are visualised in Mixed Reality (MR) [22] by embedding the DT in situ in its PT, thus fusing the two. The fused representation enables Situated Analytics [19] which facilitates the user’s understanding of analysis through spatial cues of the environment.
This review is structured in four main sections. To define the term FTs, we need a thorough understanding of Smart Cities, Situated Analytics, and DTs. The first three sections cover the necessary background to develop, understand, and apply the FTs paradigm. We dedicate Section 2 to gaining an understanding of how Smart Cities arose as a research field and why they converge towards DT research. Section 3 covers Situated Analytics as a novel data exploration approach for Smart Cities. The similarity in the implementation structure leads us to the following section on DTs. Section 4 elaborates on DTs. We provide a definition following [23], where a DT is constructed using five components (see Figure 2). DTs is a vague term that requires multiple clarifications in terms of composability (Section 4.1), servicisation (Section 4.2), and differences among cognate terms (i.e., functionally and historically related but slightly different terms [24]; Section 4.3). Finally, in Section 5, we cover the FTs concept in depth including the required steps to implement FTs and classify examples of de facto FTs in the recent literature. Section 6 provides an in-depth discussion on Smart Cities (Section 6.1), Situated Analytics (Section 6.2), DTs (Section 6.3), and the newly introduced FTs paradigm (Section 6.4).

Figure 2. Components of a Digital Twin. A DT is often thought of as the interaction between the PT and the DT [2], but any DT implements five core components, even if only partially. These components are the Physical Environment, the Data Environment, the Analytical Environment, the Virtual Environment, and the Connection Environment [23]. The PT is measured through the Physical Environment of the DT, and the raw data are stored in the Data Environment. The Analytical Environment provides simulations, services and automatisation. Users often only perceive the DT represented in the Virtual Environment as a model of the PT without being aware of the involvement of all the other components. The Connection Environment is the invisible glue that holds the different components together and enables composition with other DTs if properly implemented. Generated with http://draw.io.

Beyond the definition for FTs, we also provide several new points on a wide range of topics. These topics have often been covered shallowly in shorter research papers, which has left a great potential for confusion. In this review, we try to discern some of these, especially in the context of DTs.
2. Smart Cities

Smart Cities are an inherently technological approach to resolving urban issues that spawn economic, social, environmental, and governmental aspects [25] (see Figure 3). Smart Cities are constituted using benchmarking indicators that can be constructed from technological sources [26]. This technologist approach only provides a narrow perspective [27] that is often divided along disciplinary boundaries [28]. The weakness in the Smart City concept is juxtaposed by its necessity because technological advances are one of the few options to manage the complexity of modern cities [25]. However, at the same time, it must be acknowledged that there are other perspectives or “urban imaginaries” [29] beyond Smart Cities such as resilient cities [30], intelligent cities [26], and sustainable cities [31]. In the remainder of this section, we will expand on the evolution of “Smart Cities” from the nineteenth century all the way to its portrayal in contemporary urban discourse.

Figure 3. The Smart City with a Digital Twin Foundation. Smart Cities rely on effective data collection, analysis, and communication to work. A DT can provide the required processing of the data collection systems to formalise the PT. In a Smart City, a DT provides analysis for governance, management, and entertainment. While governance refers to directing the city’s future, management refers to the present state of the city, and entertainment refers to the use of the city. Generated with http://draw.io.

Throughout their evolution, advances in ICT have repeatedly helped cities to cope with growing complexity [32–34]. During the nineteenth century, industrial cities were faced with the management challenges of large industrial enterprises as well as new administration challenges due to unprecedented population growth. The need to process information and communicate it quickly over great distances led to the invention of the telegraph, which enabled the industrial city to keep growing [32,35].

This transformation, from a society in which communication was impeded by distance to a society in which distance is addressed using ICTs, has forever changed humanity and cities [35]. On the one hand, critical obstacles to the growth of cities were removed with the integration of technologies that enabled businesses to flourish and municipalities to govern more effectively. On the other hand, futurists such as Toffler [36] and Negroponte [37] predicted that these technologies will lead to the “death of distance” because it is possible that improved telecommunications technologies will make cities, and space itself, obsolete. Their basic proposition was that electronics will eliminate the need for
face-to-face interactions and that the cities at their core facilitate those interactions and thus would themselves become obsolete [36,37]. Clearly, this is not yet the case. Major cities have become points of intense investment in ICT because the production, exchange, and dissemination of information were and still are critical to their function. Paraphrasing [38], the shift from goods handling to service and information industries has made cities into communication systems that are central to information flow in the global economy [38].

Today, the integration of ICTs into Smart Cities has become a complex man-made creation that is unparalleled in its cascading effects on every aspect of the city and its citizens [32,33]. Two innovations in the 20th and 21st centuries have enabled this evolution: the rise of the Internet infrastructure for networked computers and mobile personal computing. In the words of [33], “The democrat[isation] of computing power that started with the PC in the 1970s and leap[t] onto the Internet in the 1990s is now spilling out into the streets [33]” which can be interpreted as the rise of IoT as part of the Smart City landscape.

Contemporary ICTs have restored the importance of space to the previous discourse on urban telecommunication [39]. Everything and everyone in cities generate geo-taggable information which can reveal the position of a multitude of stationary or moving objects in real time [39,40]. Geo-tags enrich space with contextualised electronic content and are a fundamental dimension of a sort of return to space, or spatial turn, to use the expression coined by urban geographer Soja [41]. This spatial turn of digital technologies has been reinforced by the proliferation of electronic interfaces and wireless communications, which realise the notion of Ubiquitous Computing or Ambient Intelligence [33,39,42–45] and ultimately underpin a DT.

Sometimes visible, but more often hidden, countless chips and sensors allow objects, people, and vehicles to be located; consumption levels and transactions to be recorded; and temperatures, pollution levels, population densities and flows to be measured [46–48]. The possibility of transforming every Thing in the built environment into something more “smart” appears more feasible than ever before. Low cost, miniature sensors enable everything to sense and be sensed. The ability to perform analytics over the data collected has become ubiquitous, enabling Things to “think”, while wireless communication technology enables Things to “talk” to one another as part of the Internet of Things (IoT) [33,37,39,40]. However, as noted earlier, this dream of smartification has not been realised due to the increasing complexity of cities. It increasingly looks as if a DT is required to provide the necessary abstractions to enable the Smart City.

Ambient Intelligence has taken on many forms that resemble the functionality of DTs, but they remain disconnected and often disjoint. This leaves it up to the observer to connect the dots, forming the meta-analysis on the city level. Examples of Ambient Intelligence range from trash cans and parking lots to streetlights. Smart waste disposal containers in Barcelona are able to know when they have reached their maximum capacity and must be serviced [49]. Smart parking uses sensors to know if a parking lot is free or not. Citizens can then be directed to available parking in the London borough of Westminster [50]. In the Future City Glasgow project [51], streetlights can be equipped to sense and analyse what is happening below them. Sensors collect data on temperature, humidity, noise, and movement, and interpret the data into a functional “information-scape”. The streetlights can then respond in real time by adapting their light intensity or by communicating the interpreted data to the nearby police station, reporting any ongoing crime.

There are also even more ambitious approaches towards constructing a DT. The Spanish city of Santander with 180,000 inhabitants installed 20,000 sensors to monitor temperature, luminosity, ground humidity, and available parking slots [52]. Paris’s 120,000 trees have each been equipped with a radio-frequency identification (RFID) chip [53]. In a large number of towns, vehicular traffic is monitored in real time through technology ranging from induction loops, to wireless sensors, to video cameras. In Singapore, this monitoring enabled the introduction of one of the first dynamic pricing systems, which consists of making drivers paying in proportion to the desirability and, more importantly, the congestion levels of thoroughfares [39]. This new capacity of cities to detect,
measure, and record what is happening within them is truly unprecedented in their history [33,39,40] and yet requires new infrastructure to manage the sheer complexity [54,55]. In the words of [56]: “If aerial photography revealed the skeletal structure of the city during the nineteenth century, the revolution brought about by contemporary ICTs is likely to reveal its circulatory and nervous systems—as it reveals the “real-time city”. For the first time in history, cities can be understood as a whole, the way biologists see a cell—instantaneously and in excruciating detail, but also alive [56].”

This new understanding of real-time urban dynamics is not (yet) centralised, nor is it ensured to be the property of city governments and planners. If serviced properly and arranged as DT components, it may remain distributed in nature with a low barrier to entrance and can be accessed from nearly everywhere at nearly any time. The ongoing rise in personal mobile computing, especially smartphones, is transforming how individuals perceive and interact with the urban space that surrounds them. Situated Analytics powered by DT infrastructure would provide citizens, tourists, and other users unprecedented insights into the city to help them in their every day life and in planning how to use the city in the future. Using DTs to power Smart Cities will make analytical tools available to everybody and will make them accessible in a spatial context that will democratise the understanding and use of the city. To realise this vision of a Smart City and deal with the complexity that is already plaguing current implementations of Smart Cities, DTs offer a bedrock. The composability of separate DTs into a city-wide DT, underpinning the Smart City, is a promising approach (see Figure 4).

Figure 4. Composition of Digital Twins to describe higher order complex systems. Multiple DTs are composed into higher order systems such as cities. Here, an open stationary city DT is composed of different open or closed and mobile or stationary DTs such as building DTs, transportation DTs, and infrastructure DTs. Each of the DTs contains complex nested DTs of different types such as for healthcare or industrial settings. The complexity level can be arbitrarily increased through nesting. Generated with http://draw.io.

3. Situated Analytics

Augmented Reality (AR) combines computer-generated visual content in 3D with the real world by superimposing virtual content onto a view of the real world [57,58]. The principles of an AR system are three-fold, combining real and virtual elements, registering in three dimensions, and, importantly, providing interactions in real-time [57]. There are at least two categories of approaches to produce AR. First, screen-based approaches, or “see-through AR”, can be divided into the optical combination of a screen with the user’s perception of the real world or 2D displays that replay a live recording of the real world and superimpose virtual content [57,59]. Second, projector-based approaches, or “spatial AR”, consist of projecting the virtual content directly onto the real world, benefiting from
projection mapping algorithms and hardware [59–61]. The last decades have refined see-through AR and, today, the main lines of AR systems are either based on the HoloLens for optical systems and ARKit or ARCore for display-based AR on smartphones and tablets. With the systems maturing to the point where it is both easy to use and easy to generate content, AR applications are going mainstream [62] and find applications across all industries [63], commerce [64], entertainment [65], education [66], and medicine [67].

As AR becomes more than a gadget, it opens the door for immersive analytics [68] of data and information either in situ or remotely (see Figure 5). Whereas the systems to display AR have been developed over the last decades, the visualisation of information in this new media is still in its infancy and yet rapidly evolving [69]. For example, traditional visualisation on 2D screens and print is a well established field with many well known representations and toolboxes, and high literacy among producers and consumers. However, most spatial 3D visualisations are still hand-coded and often imitate 2D visualisations. 3D visualisations also offer the opportunity for interaction and embedding in the augmented space [70,71], possibly easing cognitive load and enhancing cognition [72] through externalisation [73].

![Figure 5. Immersive Analytics compared to Situated Analytics.](https://miro.com) The real world is schematically represented in grey. AR content is shown in blue. On the left side, a user is displaying a Smart City environment (blue model) in non-situated AR. Interactables are highlighted in white, and several pop-ups showing information (white squares) are opened. Immersive Analytics make use of 3D space to visualise content but can be independent of the spatial context surrounding them. On the right side, a user (in yellow) is walking through the Physical Environment of the Smart City. In Situated Analytics, interactables are AR content that is displayed in situ for the Physical Environment that they describe. Similarly to (a), pop-ups (blue squares) are showing content to the user in AR space. Generated with https://miro.com and https://rhino3d.com with royalty-free models from https://cgtrader.com.

The development of AR technologies also enables new types of interaction with digital content. Specifically, embodied interaction is a concept that reconnects the fields of tangible and social computing and argues that interaction is fundamentally physical and social. Thus, failing to consider these aspects for design interaction would be a fallacy [74,75]. Embodied interaction is defined as “the creation, manipulation, and sharing of meaning through engaged interaction with artifacts” and places the responsibility of meaning-making on users rather than designers [74]. The level of responsibility for meaning-making at the object’s level is driven by different definitions of affordances across research fields that have yet to be reconciled [76]. There is an epistemological difference between affordances as properties of the object or interaction [77] and affordances as properties of the users’ mental representation [78]. From both Normanian [78] and Gibsonian [77] perspectives, AR can be viewed as deeply anchored in the physical world and therefore demonstrate strong affordances for embodied interaction [57]. Moreover, AR supports a higher amount of
embodiment compared to 2D displays by enabling high levels of sensorimotor engagement and immersion [79].

For example, anchoring and embedding visualisations in the physical 3D world can offer tangible ways to explore one’s physiological data [80] and, to some extent, one’s own “personal” DT. For example, Tobe is an AR tangible humanoid DT to explore one’s physiological data such as heart rate and breathing patterns and was used in a social context to let two users relax together via data observation [81]. Similarly, Teegi is an AR tangible humanoid DT used to explore one’s brain activity via electroencephalogram visualisations [82]. Non-humanoid representations have also been explored. Inner Garden is an augmented sandbox embodying one’s physiological state, representing breathing patterns as waves and cardiac coherence as weather [83]. In the wider context of embodied interaction, users’ bodies have also been explored as canvases for digital content [84], for example, by co-locating display and interaction on the users’ hands [85] or utilising embodied metaphors, such as mimicking scissors with one’s hand to cut through digital content [86]. Tests of embodied AR approaches have also been initiated for data exploration in the context of geographical data for visually impaired users [87–89].

The integration of users’ bodies and their physical context has been conceptualised during interaction with digital content. Towards that end, different physicalities of embodiment have been described: body-centred, object-centred, and environment-centred embodiment [90,91]. We offer examples of embodied DTs and FTs following this categorisation in Figure 6. These categories themselves include direct-embodied, enacted, manipulated, surrogate, and augmented approaches [90,91]. The direct-embodied approach considers the users’ bodies as primary constituents of cognition and therefore the body state is used to convey information. In the enacted approach, bodily actions are considered a form of enacting knowledge through movement. The manipulated approach uses the body to directly manipulate a representative object. The surrogate approach uses a physical avatar that performs an interaction with the object. Finally, in the augmented approach, the body is used to manipulate an avatar to interact with the larger environment. The impact of the physicality of embodiment on interaction and cognition is not yet clear. However, these considerations are of importance because a mismatch between activity design and type of embodiment can result in decreased movements, persistence, and learning outcomes [92].

When discussing the spatial context of data visualisation and data embodiment, we also ought to consider the role of the users’ bodies in the interaction. Often, the users’ bodies are solely considered as physical entities (i.e., Körper) utilised to push buttons and perform actions, rather than feeling entities (i.e., Leib) [93]. Novel approaches such as somaesthetic appreciation design [94,95] describe the manner in which users’ bodies may be integrated in the design process as well as in the experience with digital content and data. Empirical evidence suggests that bodies’ representations impact the sense of embodiment [96,97], as well as the types of gestures performed to interact with a system and the time spent observing and interacting with data [98]. In this regard, the challenge of FTs extends beyond meaningfully anchoring virtual data onto a physical context to include the design of the interaction with the augmentation from an embodied, somaesthetic perspective [74,94,95].

Two further problems arise when visualising embodied information in augmented space. First, most traditional authoring tools are focused on 2D visualisations, and new toolkits such as DXR [99] and RagRug [100] have only provided the basis to display 2D visualisations of data in 3D. Second, the benefit of the third dimension is difficult to grasp and even more difficult to implement. Most modern immersive analytics actually rely on embedding 2D displays into the real world [17,99–101]. Visualisations in 3D are typically derived from the previous 2D media, but there is a trend towards exploring new possibilities. For example, multiple 2D visualisations can be arranged to use the third dimension to filter information and visualise this filtration [102].
Figure 6. Examples of Embodied Digital Twins and Fused Twins. We observe the difference between non-situated AR (top row) and situated AR (bottom row) for anchoring content at the body of the user, an object, or the immediate environment. The distinction between body-, object-, and environment-centred AR is inspired by the classification of different physicalities of embodiment offered by [90,91]. With a body anchor, the DT is anchored on the users’ bodies, similar to the directly embodied approach. With an object anchor, the DT is anchored on an object with which the user can interact, similar to the manipulated approach. With an environment anchor, the DT is anchored in a room, similar to the augmented approach. In the situated environment (implementing the FTs paradigm), the relationship between the DT and the representation is homologous such that the body, object, or environment is used to anchor its own DT in a one-to-one correspondence. However, in the non-situated environment, this relationship is non-homologous because the mapping from the DT to the representation is transformation instead of a one-to-one correspondence.

Situated visualisations are less represented in the literature. Recently, there has been some defining work that characterised the types of enhancements that immersive analytics may provide through enhanced vision and visual synesthesia [72]. Addressing the question of how situated visualisations might improve cognition is an ongoing research area. The supposition has been put forth that cognition may be improved by enhanced attention, enhanced numeracy, enhanced recall, enhanced comparison, and enhanced prediction [72]. Moreover, theoretical frameworks such as embodied, situated, and socially situated cognition support the idea that cognition is deeply rooted in its physical and social context [103–105]. However, there are no empirical evaluations yet because, so far, researchers have not focused on the embodiment of the user in the analytics or understanding how the spatial context can reduce cognitive load or enhance cognition.

Some research is starting to address the impact of the third immersive dimension on visualisation possibilities [106]. For example, one small study demonstrated that a toolkit to create new immersive visualisations helped to keep users engaged for longer and produce more elaborate immersed visualisations. We have a clear understanding of how 2D visualisation improves cognitive load through externalisation [73], but for AR, the evidence is less clear [107–111]. Previous work has proposed the following mechanisms for how 3D visualisations can improve cognition by reducing task switching with visual cues [112,113] and context-sensitive displays [114]. Moreover, recent work also explores the potential of AR to reduce negative split-attention effects in multimedia activities [85,115,116]. Split-attention effects [117] occur when two related sources of information are spatially or temporally distant. However, split attention is not always negative. For example, indexing (i.e., connecting multiple data representations through pointing gestures) is investigated as a mechanism for sense-making [118]. The distance between information sources can also be reduced using AR, in particular the FTs paradigm, because it co-locates physical and virtual entities in real time [57,115,116,119,120]. In addition, overwhelming displays of information have been shown to decrease performance [121]. There is no definite answer as to how immersive analytics ultimately improves users’ performance, but theoretic arguments keep this research area active.
Furthermore, the spatial context in the augmented space has often been used as an empty canvas [99] instead of a source of information [122]. Most commonly, this spatial context has been used for geographic representations to show the built environment either in the past through remote sensing or as planned in the future [122]. The subfield of Situated Analytics [19] combines data with its spatial context to achieve a form of virtual data physicalisation [123] and embodied interaction [70,74,75,85,86] that enables humans to maximise the use of all senses to perceive the data (see Figure 7). Indeed, information on cities can be embedded in the situated context with meaningful 3D representations that users can explore. For example, situated AR has been used to teach historical events by anchoring the relevant data in the actual physical context of these events [124]. The real world can become the canvas to help users make sense of the information in situ [125].

Figure 7. Situated Analytics [19,126] within a DT context. Situated Analytics use the Digital/Physical Divide to define the generative steps in situ in immersive analytics in the physical world (bottom row) or digital world (top row). There is a remarkable similarity between the components of DTs (blue boxes) and the generation of Situated Analytics (white boxes). The mapping is visually nested to be highlight. An observer in the physical world sees the physical referent or PT (in the DT context). The observer also sees the physical presentation of information (Virtual Environment of a DT). In the digital world, the raw data are received from the referent, which covers the processes of gathering data by the Physical Environment and storing data by the Data Environment. The Visualisation Pipeline produces a visual output (Analytical Environment of a DT). Lastly, the data are placed into the physical presentation (Virtual Environment). Generated with http://draw.io.

4. The Digital Twin

Before we explore how FTs can make DTs more accessible, we will review DTs, their historical background, and their relations to other adjacent technologies and the wide range of definitions that accompany them. Historically, the conceptualisation of DTs originated from the industrial branch of IoT developments in 2003 [1]. The concept was loosely defined at the time but already included the most important components: a physical object, a virtual object, and their connections [2]. Throughout this section, we will expand this understanding to reach our definition of a DT based on composability and the five components, the Physical Environment, the Data Environment, the Analytical Environment, the Virtual Environment, and the Connection Environment (see Figure 2).

Historically, there was a lag in the adoption of DTs because the technologies necessary to effectively implement a DT were not yet developed [8]. In the 2010s, the DT concept outgrew its industrial roots by capturing cyber–physical integration as well as representing the physical world virtually [127,128]. As DTs grow in size from factories and cities [129] to the globe [130], remote sensing with drones and satellites plays an increasingly important role to keep DT representations up to date but also offer an opportunity for interactions...
with the city through interactive mobile display technology [61]. For a full review on current DT applications, we refer to [54]. Here, we will elaborate the definitions of DTs to form a common basis from where we can develop the concept of FTs.

DTs have been “hyped” according to Gartner as one of the top ten most promising technological trends of the previous decade [8,131–134]. Beyond the hype, across many different definitions, DTs fuzzily describe any software aiming to closely represent something in reality [135]. These qualities are in line with common and older definitions such as Cyber–Physical Systems (CPSs) [136], Ubiquitous Computing (UC) [46], and Mirror Worlds [137], which begs the question of whether we can (and should) draw a line between DTs and these concepts. Here, we argue that DTs should be approached as an analogy to a mental model that helps to structure information and actions and to attach them to the real world through appropriate mechanisms such as the FTs concept.

In this review, we next situate DTs in the context of their predecessors. DTs are a specialised form of Cyber–Physical System (CPS) [136] because they focus on the accurate representation and manipulation of processes in the real world with a model. DTs are similar to Ubiquitous Computing [46] because they attach computations to real-world objects, but without the FTs paradigm, DTs are ultimately not Ubiquitous Computing because the computations are detached from the embodiment of the represented object. Fusing the PT and DT [17,101] joins CPS properties with the PT to produce truly Ubiquitous Computing as envisioned by Weiser [46].

Whereas Ubiquitous Computing has no clear requirements for the kind of computations to be performed, there is a common theme across different definitions of DTs with respect to the need for simulations to predict a system’s response to critical events. The aim to predict often goes beyond simulating known events towards understanding deviations between the actual observations and the predictions to indicate unknown issues [138,139]. Furthermore, DTs are often expected not only to report on the state of the system, but to also actively manipulate the system through actuators [2,54,139]. The tension between the accuracy of DT predictions and computational capacity is captured by the idea of Mirror Worlds [137], which perfectly imitate reality similar to a mirror. Early ideations of the DT therefore required the underlying model to support similitude [138], while similitude requires the implementation of a multi-physics, multi-scale, probabilistic, ultra-fidelity simulation that mirrors the PT based on historical data, real-time sensor data, and a physical model. The “ideal DT” requires similitude to become as close as physically possible to mirroring the real world. Indeed, DTs need to model the mutual dependency of hardware and software to tightly couple the PT and the DT, producing high-quality simulations. Until recently, most simulations were feasible only for certain sub-components of complex systems that relied on scientific knowledge and real observations to produce a more accurate offline simulation. Therefore, the requirements of similitude have only been satisfied in specialised cases such as with the components of flying vehicles [138]. New computational methods increase simulation performance such that events in the DT can be simulated fast enough for predictive purposes in the real world [140]. However, even with respect to aircraft, there are only plans to produce a DT of a whole vehicle [141,142] which are expected to be completed in the coming decade. Regarding the sheer scale of cities, the Mirror Worlds that would underlie a Smart City seem to be a distant dream. Due to the simulation complexity of large-scale systems such as cities, researchers have since called for a less rigorous model definition to make DTs relevant to real-world applications [9].

In other words, stepping away from ideal DTs to “practical DTs” can allow for more practical applications. Easing any of the core aspects of similitude allows for the implementation of many useful concepts such as analytical assessment, predictive diagnosis, and performance optimisation [8], even in large-scale systems. Limiting the functionality of DTs to a synchronous representation of reality can also allow for the monitoring of complex production processes, adjust these processes, and possibly optimise them [143]. Ideally, the model underlying a DT is complex enough to enable a system to continuously react to dynamic changes in physical space [144], including unforeseen/unexpected
changes [145]. DTs also open the door to continuous design improvement through data feedback loops [146] enabling data assimilation into a model [147,148]. These incomplete but practical DTs can become especially useful if they are composable to cover more features of the PT. Next, we will introduce specialised, practical DTs which answer questions on particular topics and can be composed to resemble ideal DTs.

In our definition, we side-step the discussions surrounding ideal DTs by focusing on structural features of practical DTs. We refine the definition of a DT based on earlier work [8] as a five component system: the Physical Environment, the Data Environment, the Analytical Environment, the Virtual Environment, and the Connection Environment (see Figure 2). In the Physical Environment, IoT systems [47] and remote sensing units [149] try to capture the PT and provide interactions with the PT through actuators. The Data Environment contains the DT’s memory and is usually located somewhere in the cloud but may also be stored locally or at the edge [150,151]. For the Analytical Environment, the processes and models of the DT are computed [152], often in the cloud, and insights and predictions are communicated back to the Data Environment. The Virtual Environment makes the DT accessible to human users and possibly other machines [17,101,153]. Finally, the Connection Environment provides an Application Programming Interface (API) to transfer data as required between the components. We consider two ways of designing the Connection Environment, (1) from a service perspective, where the Connection Environment is not an independent component but refers to the APIs exposed by the different components, and (2) from a compatibility/hypervisor layer perspective where the Connection Environment is an independent component that mediates between the other different components in the system. Each component could be implemented in many different ways, and making a sensible choice may be overwhelming. This review gives an overview to guide newcomers through the most important components and decisions that are required to implement the FTs concept with a DT.

To understand DTs, it is necessary to comprehend the function and goal of the core components in relation to the literature. Specifically, we connect our definition of the Physical Environment to research on IoT and remote sensing, as well as our definitions of Analytical and Virtual Environments to research on simulation and interaction with DTs. Our definitions of the Data and Connection Environments are not broadly covered in previous research because previous implementations have used case-specific data solutions and were not focused on interoperability.

The Physical Environment component captures the PT and associated interactions via IoT and remote sensing infrastructure. The composition of IoT stack is regularly repeated but rarely analytically decomposed with respect to its trade-offs. For most applications, the stack is designed with the purpose of underlining where and why a specific new technology is needed [154–157]. Technically, this is a useful approach to help experts situate a technology in the literature. However, these stacks typically make many unspecified assumptions that we disentangle here. We start with the notion that the DT perspective can be mapped to aspects of IoT and remote sensing. Following the semantic decomposition of the term IoT [158], Things themselves are not actually in the Internet, but information about them is perceived either with attached sensing units or with remote sensing and then shared via the Internet. The DT can be used as the necessary representation of the perceived Thing in the Internet.

Previous resource-oriented and service-oriented architectures have represented IoT hardware as practical DTs without using this particular nomenclature [157,159–164]. In contrast to IoT definitions, the DT perspective does not necessitate the digital information to be shared globally as may be implied by the term Internet. The potential local scope of DTs is reflected in their origins in cyber–physical systems [2,136] where the application scope is usually limited to an industrial production site. Nonetheless, the DT perspective provides us with a useful definition of the properties that a Thing should have when represented digitally. A broad range of example IoT applications have been addressed in previous publications [7,45,47,165–167] and can be generalised to DT solutions.
The Analytical and Virtual Environments in DTs are difficult to precisely define because their development is multi-disciplinary and reaches beyond the simulation perspective. These paradigms have been advanced by independent research in the areas of information science, production engineering, data science, and computer science [8] as well as electrical engineering [168]. A DT is a highly complex system that combines infrastructure, simulation, data processing, human–computer interaction, and services [8]. These approaches give rise to another understanding of DTs as an advanced digitalisation of a system that extends representation of the PT with operations. This produces a rich digital model that has many desired properties such as scalability, interoperability, extensibility, and high fidelity [169]. To arrive at a digital model, digitalisation in a space can also be understood as a virtualisation process [170]. Here, the system is represented in 3D space via modelling and remote sensing. First, behaviours are extracted and knowledge representations are created. Second, the interactions between its constituent elements are modelled and integrated into the knowledge representations. Third, operations and their effects are modelled. Lastly, everything is assembled in a simulation to make the system match the PT as close as possible. Because digitalisation/virtualisation only creates the Analytical Environment (see Figure 2), most research has clearly focused on the analysis and not interaction with the DT. Most research on DTs also follows the simple physical–digital split into PT and DT components [2,135], but this split often fails to capture the difference between the Analytical Environment for computation and the Virtual Environment for interaction (but see [23]). The FTs paradigm offers a new perspective that highlights the unique contribution of the Virtual Environment, resulting in a more intuitive use of the DT in nearly every application domain.

There may also be issues with many definitions of DTs because of the unrecognised difference between a universal ideal DT and a partial practical DT. Ultimately, universal ideal DTs are unachievable because of the need to integrate the complex requirements of similitude and to comply with the many definitions across different literatures. Therefore, any practical DT is only partially implemented compared to ideal DTs. Confusion regarding this distinction has led to a heated debate and has resulted in the differentiation of DTs according to their degree of completeness from digital models and Digital Shadows to (ideal) DTs [152,171], see Figure 8.

A digital model is an independent representation of a real-world system that may be used in simulations but that is generally not fed real-time data. According to our interpretation, a digital model reflects the implementation of the Analytical Environment. A Digital Shadow is the assembly of data from the real-world system for descriptive purposes. By our interpretation, a Digital Shadow reflects the Physical Environment of the DT. Thus, our definitions can shift the discussion on Digital Shadows and digital models from whether something is a DT to what component of a DT is implemented. While a completely implemented, ideal DT at the urban scale is probably out of scope for now, it is feasible to implement and compose partial practical DTs that are capable of performing important tasks for Smart Cities. To elucidate IoT, DTs, and Smart Cities, the present paper adopts the FTs perspective because it unites three views on DTs that are often confounded: Why to generate the DT (i.e., business case for the Analytical Environment), how to generate the DT (i.e., technological implementation of the Physical Environment), and how to use the DT (i.e., end-user interaction of the Virtual Environment).
4.1. Composition of Digital Twins

Composed DTs offers an integrative perspective to understand and address the complexity of the PT by using multiple abstractions and performing different tasks on each abstraction (see Figure 9). Towards this end, DTs are composable through their Connection Environment allowing both for the nesting and recombination of components. This approach allows us to re-conceptualise any sensors in the Physical Environment as virtually belonging to all associated composites of the DTs. Composites of the DTs can be produced to offer previously unavailable information via new virtual sensors [172] or virtual sensor networks [173,174].

For DTs to be composable, interoperability between components is the most important feature [169,175–178]. An example composition of the different components is shown across associated DTs in Figure 10. The Physical Environment contains sensory equipment within arbitrary boundaries to capture properties of the PT, which is immutable and equally anchors all overlapping DTs. The Data Environment is a heterogeneous composition of storage architectures that best befits the data requirements. The Analytical Environment is usually defined per task per DT. In addition, it may also have feedback loops to other DTs via the Data Environment. Outputs of the Analytical Environment are fed back into the Data Environment as virtual sensors to provide the information for display and interaction at a later time. The Virtual Environment forms a local Mirror World [137] that allows for the visual inspection of information in the DTs. The scope of the Virtual Environment is limited to the current composition of DTs and allows for access to and interaction with both the data from the Physical Environment and the data from the Analytical Environment. The Connection Environment manages access rights between components and ensures that users (both humans and other DTs) have the appropriate access rights to use the respective environments, addressing concerns regarding trust, security, and privacy [178–192].

Figure 8. Digital Shadows and Digital Models in the context of Digital Twins. The left side demonstrates the relations between a Digital Shadow, a digital model and a DT [152]. The right side demonstrates the mapping to our definition. The Digital Shadow corresponds to the Data Environment and relies on the Physical Environment to collect data. The Engineering, Digital Model and their DT simulation correspond to the Analytical Environment. The experimentable DT is akin to the Virtual Environment. In our definition, all blocks are components of the DT, whereas in the other definition, the DT is a piece in the diagram. The reduction of the DT to a simulation removes the HCI from the system and requires terms such as Digital Shadow and Digital Models to differentiate DTs. Under our definition, these terms reflect different aspects of a DT and map to different components. Generated with http://draw.io.
In practice, DTs are applicable to all fields in which IoT has been applied [47,165,193], including industry [159,194], health care [43,44,166,195–198], urban management [48,199,200], vehicular application [201], and mobile end-user applications [202]. However, the composability of DTs suggests that, by extension, these applications should be composable, too. For example, DTs in industry, health care, and transportation could exchange vital information to
improve their effectiveness, but, to the best of our knowledge, this has not been accomplished in practice. To understand how DTs of these varied systems could be composed in the context of true Smart Cities, we categorise each application by its system type and location type to form a taxonomy of DT contexts (see Table 1).

Table 1. Taxonomy of Digital Twin Context Types.

<table>
<thead>
<tr>
<th>System Type</th>
<th>Open</th>
<th>Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location Type</td>
<td>Stationary</td>
<td>Cities, Economies &amp; Countries</td>
</tr>
<tr>
<td></td>
<td>Mobile</td>
<td>End users, goods &amp; vehicles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial processes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Body (health) &amp; machines (internal)</td>
</tr>
</tbody>
</table>

The system type of a DT is either open or closed [203]. Open systems are characterised by an exchange with the wider environment in which they are embedded, whereas closed systems do not directly exchange with the environment and are usually observed in isolation. In addition, closed systems usually supervise well defined processes, whereas open systems try to characterise ill defined processes. Obviously, no system is entirely isolated, but industrial processes or health applications can be considered relatively closed systems with usually well defined interactions with the environment. In contrast, cities and traffic [178] have no clear limit to what is entering from the surrounding environment, and congestion may be a side effect of this ill defined boundary.

The location type of a DT describes the relation between the unit of observation and the sensors as active or passive sensing [204]. In stationary systems, the sensors are fixed in the environment and observe throughput. In mobile systems, the sensors are moving around either with the observed unit or independent of it. They may also capture environmental conditions around the observed unit such as in human navigation [205] and health monitoring [206].

When composing different types of DTs, it is important to understand their characteristics. Mobile DTs may only temporarily compose with Stationary DTs, while Closed DTs often form units within Open DTs. Open Stationary DTs often form the overarching type of DT that contains other types of DTs, whereas Mobile Closed DTs usually describe atomic units in composed DTs. Typically, Mobile Closed DTs can be part of Mobile Open DTs when a person has a Body Area Network (BAN) [207] for health care applications and a Personal Area Network (PAN) [208] to interact with the environment. In this example, the PAN will provide a wide range of interactions that can be composed with the Open Stationary DT of a Smart City, whereas the BAN will only share limited information. Similarly, a Closed Stationary DT for a Smart Factory within Open Stationary DT of a Smart City will only share limited information.

4.2. Servicisation of Digital Twins

While the word DT suggests some form of unitary entity, it would be more accurate to understand DTs as context-aware representations of Things in the cloud composed of services that act upon the abstraction. Services can range from interacting with the environment to representing the sensed and forming the bedrock of ubiquitous computing [46]. Basic services again can be composed into higher-order services to address more complex issues.

With digital representation by the DT, it is possible to turn everything into a service—or servicing [209–212]. In other words, the customer buys the temporary use of a service associated with a product instead of the product itself. The cost of developing and maintaining a DT’s infrastructure justifies this approach because it is more efficient to reuse resources rather than to duplicate functionalities [212]. Therefore, DTs can be freely composed of services that conform to the requirements of the Connection Environment.

The three most common themes in “as a Service” (*aaS) are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Newer types of services
can be subsumed in these categories, including Sensing as a Services (SenaaS) [213,214], Semantics as a Service (SemaaS) [176], and Robots as a Service (RaaS) [215]. Notably, Infrastructure, Platform, and Software “as a Service” form an interesting interaction with the DT components because they can provide abstractions for particular components (see Figure 2). From an abstraction perspective, each replaces the lower layers of the stack with an interface that hides implementation complexity. In this section, we will use the example of weather services as part of a Smart City to explain how IaaS, PaaS, and SaaS can map to the components of a DT (see Figure 11).

**Figure 11. Mapping of the “as a Service” paradigm to the components of a Digital Twin and the weather services example.** This example application is an alert for a nearby weather event. Presented from left to right, the PT forms the Physical Environment of the DT and is captured by local sensors and networks and remote sensing as IaaS. The Data Environment of the DT serves as a PaaS. The Analytical Environment of the DT models the weather and calculates the proximity to the predicted event as a SaaS. The Virtual Environment is implemented as a smartphone application that provides an alert on an abstract Dashboard. The Connection Environment spans across all services and environments providing an API that allows for one environment to be replaced by another. This makes the DT both modular and composable. Generated with [http://draw.io](http://draw.io).

### 4.2.1. Infrastructure as a Service (IaaS)

A DT requires the creation of a Physical Environment as the representation of the PT. The Physical Environment provides access to and measurement of the PT using infrastructure to collect raw data (see Figure 2). Creating and maintaining the DT infrastructure is the most physically laborious part of developing a DT. This severely limits the number of entities that can afford to create a DT on a city scale. However, provided as a service, it may be possible to finance the required infrastructure in aggregate and reuse it across multiple DTs. When creating a DT, the IaaS approach allows for the reuse of a service provider’s communication network, sensor network, and remote sensing units rather than having to build one’s own system from scratch. Therefore, IaaS enables one to have alternative competing and composable DTs based on the same or different data source(s). There are also more specialised infrastructure services that are available such as Sensing as a Service (SenaaS) [213,214], Remote Sensing as a Service (RSaaS) [216], and Robots as a Service (RaaS) [215].

In our weather service example (see Figure 11), the sensors, network, and remote sensing units in a Smart City could be servicised. The city government or a private provider could implement a policy to maintain such an IaaS to make digital services in a city more accessible, enabling the creation of different city DTs. This would allow new entrants to easily develop their own DTs that provide large-scale digital services without having to build physical hardware with redundant functionality [210].

At the same time, IaaS may raise some concerns regarding security. The IoT used for IaaS consists of billions of locatable, addressable, and readable devices [191]. There are three
avenues by which security in IoT must be enforced: information, physical infrastructure, and management practices [189]. Security breaches from unexpected behaviour may arise from data collection and management via privacy protection problems, heterogeneous network authentication, and access control problems [189–191,217]. The fundamental challenges for all of these topics are questions of trust: how to give whom access for what purpose and for what duration. Over the last decade, the IEEE and IETF have laid the foundations to build a security infrastructure along the IEEE 802.15.4 standard [218,219] and the IETF LLN stack [208]. These efforts include sensor network security [184], wireless sensor network security [183], and dedicated secure protocols such as CoAP/Lithe [220,221], 6LowPSec [222], and RPL [223]. Ultimately, such as with the Internet, security will be an ongoing process that must be attended to at all times.

4.2.2. Platform as a Service (PaaS)

PaaS extends access to hardware by simplifying interfaces. PaaS usually packages digital representations into a Data Environment and thus reduces the complexity of accessing the data. Platforms vary largely and are usually tied to one of the five IoT application areas, industry [159,194], health care [43,44,166,195–198], urban management [48,199,200], vehicular applications [201], and end-user applications [202]. The concept of PaaS can be vague, and its use may sometimes be difficult to differentiate from IaaS or SaaS. In the context of a DT, a PaaS ought to provide only access to data as a Data Environment in contrast to an IaaS that collects raw data from the physical world as a Physical Environment and a SaaS that derives higher-order data from previously collected data as a Analytical Environment.

A DT probably cannot rely on a single PaaS to source all of the required data. More often, a DT will mix and match PaaS sources [224] to fulfil its task. The Data Environment is more complex than simply making data available and needs to be curated, indexed, and integrated into a knowledge graph based on context-awareness [156], ontologies [225–227], and middleware [228–230]. In reality, IaaS and PaaS are often mixed and sometimes offer a subset of services that should be associated with SaaS [231].

In our weather service example, the PaaS stores the information of the weather-related sensor data from sensing units throughout the city, including their spatiotemporal context. Specifically, the PaaS does not include any analytical components but offers a raw data source upon which higher-order data services may be built. A clear distinction is necessary between services to disambiguate their different responsibilities in the DT and ensure composability.

4.2.3. Software as a Service (SaaS)

According to the classical SaaS literature, SaaS is often exposed directly to an end user [202] and ought to follow a singular use case. However, in the context of the DT, SaaS could be mapped to both the Analytical Environment and the Virtual Environment. Following the “as a service” paradigm, we suggest limiting the use of SaaS to the Analytical Environment in DTs to reduce redundant functionality and to enable more servicisation in end-user applications (see Figure 2). Conventionally, one may consider SaaS as built upon PaaS, but the relationship between the Data and Analytical Environments may actually be bidirectional. The results of the analyses may be written back to the Data Environment before being forwarded to the Virtual Environment for user interaction. Thus, there may be an ambiguity that makes it difficult to separate SaaS from PaaS. The difference lies in whether one considers the results of an analysis to be newly computed as a SaaS or retrieved from memory as a PaaS.

In our weather service example, the classification of the predictive weather data to an end user could be qualified as SaaS if it is being calculated directly. However, if the predictive weather data are provided as a component of a Smart City environmental observation system, it would be qualified as a PaaS. In general, a SaaS conducts some analytical tasks and represents the Analytical Environment of DTs because it generates the models to mirror reality in the DT.
4.3. The Cognates of Digital Twins

DTs are a relatively new perspective that shares a common but amorphous root with a variety of other perspectives. While this root may be difficult to trace precisely and may differ from the intent of the author, it may have originated with various thinkers such as Wiener [136], Weiser [46], or even Tesla ([232], p. 122). Given this complexity, we borrowed the term “cognate” from linguistics [24], which refers to words with a formal resemblance or a common historical root. The confusing overlap among DT cognates is driven by developments across a diversity of research backgrounds such as information science, production engineering, data science, and computer science [8]. The different components of the DT are prioritised differently in each field, which results in similar yet slightly different meanings. For instance, the cognate term IoT is dominating the public sphere according to search word trends (see Figure 12). However, as hype gives way to pragmatism, simply having the physical infrastructure necessary for the creation of DTs is insufficient by itself. The two cognates IoT and DTs are often confused, but IoT may form the basis of DTs in a similar manner that the Internet forms the basis of the Web. Although it is too early to determine, search word trends may eventually place DTs or another cognate as the ultimate concept to describe these systems. The following sections clarify the nuances associated with DTs and their cognates across different disciplines.

![Figure 12. Google search trends by search term in web searches. Interests are averaged with LOESS across months from 2003 to early 2022. In general, the Internet of Things has been falling since 2017, and Digital Twin has been steadily rising. This year, Digital Twin has become the most searched term among the cognates other than Internet of Things. Terms for underlying technologies and previous system names are stagnating or declining in use. Data collected from https://trends.google.com on 8 February 2022.](image)

4.3.1. Cyber–Physical Systems

A Cyber–Physical System (CPS) represents physical objects as virtual objects. Users can interact with either physical or virtual objects to change the state of both. A DT is a special instance of a CPS that includes a model of the PT for predictive operations.

The term DT is relatively new because the original idea of digital representation was first described in detail by Wiener [136] in the field of control theory. Later, the term Cyber–Physical Systems (CPSs) Theory [233,234] was coined to describe a physical object, homologous to a “Thing” [158] in the IoT context and the PT in the DT context. This physical object has a cyber correspondence, which is homologous to the “Digital Twin” [2]. A priori, CPSs have no predetermined (physical) scope and are an idealised form to describe possible interactions with entities [45,235]. Therefore, the CPS literature is rarely linked directly to DTs, remote sensing, or the IoT (but see [236,237]). CPSs are increasingly used to capture social data as well [238]. Their high level of abstraction makes CPSs a suitable paradigm to theorise concerning DTs, and findings in the CPS literature are very relevant to DTs.
4.3.2. (Wireless) Sensor (and Actuator) Networks

(Wireless) Sensor (and Actuator) Networks (WSN/SN/SAN/WSAN) provide sensing and actuating infrastructure that processes the signals to and from the physical world to enable a CPS or DT.

A DT cannot be conceptualised without having first digitised the PT. While the DT is usually presented as agnostic with respect to its underlying hardware, the DT cannot be constructed without it. Creating this representation requires sensing to comprehend the physical world and actuating to change the physical world [168,193]. Sensor networks ground any Cyber–Physical System in the physical world by providing readings thereof. For each device, questions of how to implement “autonomicity, scalability, energy efficiency, heterogeneity in terms of user equipment capabilities, complexity and environments” [239] need to be answered. Sensor networks had been researched long before the origination of the terms IoT or DT. However, contemporary research focuses more on how sensor networks are seen as a component of an IoT system [240].

In a sensor network context, there are active and passive devices that can be placed in the environment. Passive devices called Radio Frequency Identification (RFID) harvest energy from the transmission signal via induction to return a signal [241]. The signal can be used as an identification of an object in order to trigger a software-based service or to determine its location depending on the RFID reader [242]. Typically, RFID cannot return environmental information (but see [243]). Active devices are usually battery powered and have actual computational infrastructure, sensory equipment, and possibly actuators to interact with the environment. Whether active or passive devices are required depends on the task. A DT will often require a combination of both active and passive devices to efficiently represent the PT and the processes therein. A comparison of tasks and the technology that best captures them can be found in Table 2.

Table 2. Task Classification for Active and Passive Devices.

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Active</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessing</td>
<td>Accessing information on an object linked to the device.</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Actuating</td>
<td>Triggering an action of the object linked to the device.</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Locating</td>
<td>Identifying the location of an object linked to the device.</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Sensing</td>
<td>Measure characteristic of the object linked to the device.</td>
<td>✔️</td>
<td></td>
</tr>
</tbody>
</table>

Without using the term DT, RFID implements a similar abstraction because a “physical object is accompanied by a rich, globally accessible virtual object that contains both current and historical information on that object’s physical properties, origin, ownership, and sensory context [241]”. RFID helps to embed access to knowledge in the environment and can make access to the DT context-driven [241]. For example, if users and environment are both tagged, the appropriate DT services can be called to facilitate the users’ task. At the same time, the system can track such interactions to help model activity for the DT. However, RFID tags could reveal sensitive information, and it is important to manage access control [244].

In contrast to passive devices, active devices usually have local computational power but may be constrained, which should be classified according to usability. Resource constraints on active devices can be grouped into six categories: battery, communication range, bandwidth, memory storage, processing capability, half duplex communication [178]. The IETF has classified devices based on hardware capabilities [245], and each class requires different operating systems (OSs) to function optimally [246,247]. They range from Class 0 devices to Class 2 devices. Class 0 devices have a tailored OS that provides no functionality beyond the required features. In contrast, Class 1 devices and above may already run a generalised if bare OS fulfilling the criteria of a small memory footprint, support for heterogeneous hardware, network connectivity, energy efficiency, real-time capabilities,
security, and opportune scheduling policies [246,247]. Active devices can be fully integrated into DTs and can be managed from within the DT, whereas passive devices currently only provide anchor points for DTs.

4.3.3. Smart Objects

Smart Objects are physical objects that have embedded computational power to provide interactions for its virtual object to make it more interesting and may represent the fusion of the PT and DT.

A DT can be understood as a highly accurate model of a physical system [2,9,135], whereas a Smart Object (SO) offers interesting interactions with a particular object through embedded or remote computing power [7]. Both DTs and SOs represent a CPS [234]. It could be argued that the concept of an SO either incorporates or encapsulates the concept of a DT. However, a DT is usually required to mirror the role of an actual object in a larger context, whereas SOs only provide end users with interactions with the particular object [7]. Nonetheless, many SOs will represent the underlying physical object sufficiently closely to use them interchangeably with a DT. Assuming that an SO qualifies as a DT and that the SO can offer physical access to its digital representation, it can be considered a dual embodiment for both the object (i.e., PT) and the DT, providing a physical instance of the FTs paradigm (see Figure 9).

However, a DT may also describe an entity that is composed of multiple SOs, other Things, processes, and high-order entities. These entities may entail buildings, cities, or regions [4,248–250] modelled through remote sensing [54,129,251]. Moreover, a DT may also be formed of Things that are immaterial or at least incomprehensible in the physical world such as phantom traffic jams [252]. These can be considered purely virtual objects in contrast to virtualised objects [253], which would have a mapping back to a physical object. Composition allows for the formation of even-higher-order virtual objects that form the basis of a DT in a complex system such as the processes or aggregates representing a city [10,254], traffic flow [255], or production line [23]. Furthermore, a DT may also be produced exclusively by remote sensing without having local sensors attached to the PT [54]. A DT is thus a more encompassing concept than an SO.

4.3.4. Ubiquitous and Pervasive Computing/Ambient Intelligence

Ubiquitous Computing is a paradigm that expects computation to be available for all objects to form a CPS, whereas Pervasive Computing specifies the computation to occur at the object or close to it. SO and Ambient Intelligence are instances of Ubiquitous Computing in CPSs in which the sensing and actuating capacity is directly embedded into the object that is described by the corresponding DT.

Ubiquitous and Pervasive Computing are usually considered interchangeable. The former is an older term to link computations to any object in the physical world [46]. The latter is a younger term and is associated with placing computing power within (or close to) any object in the physical world [256]. Without loss of generality, we will only refer to Ubiquitous Computing in the remainder of the present paper. For a DT, any data collected in the physical world must be represented in “cyberspace” and thus requires computation. Many Ubiquitous Computing applications pre-date the IoT paradigm but have been subsumed by it (see [257]).

The closeness between the Thing and computations performed on sensor data from the Thing has often resulted in them being regarded as equivalent to each other [47]. Similarly, the concept of a DT is often used as a synonym for both the PT (homologous to the Thing) and the Physical Environment with WSAN (homologous to computations on sensor data from the Thing) from which the DT is created. Ubiquitous Computing offers us a lens to distinguish between PT and DT by considering the difference between computations on the state of an object and the actual object. SOs [7,258–260] are a typical example where the distinctions between PT, DT, sensors, actuators, and computations are blurred. SOs are expected to have awareness of their environment and have some model of the ongoing
activities, policies, and processes around them. At the same time, SOs should represent functions, rules, and flows that allow for interactivity [7,258]. This formulation clearly indicates that the SOs have a Physical Environment in which they are embedded, an Analytical Environment to model their behaviour, a Data Environment to maintain their state, and a Virtual Environment in which their interactivity can be triggered. When the computations on the object or Thing become unrecognisable to humans because they are “woven into the background”, as predicted by Weiser [46], the term Ambient Intelligence becomes more relevant [42–45]. Making a clear distinction between a Thing, an SO, and Ambient Intelligence is probably neither possible nor desirable.

4.3.5. Internet of Things

The IoT is an amalgam of competing technologies with the goal of digitally representing the physical world. It interconnects WSANs to be combined into local CPSs that in turn compose larger DTs.

The idea behind IoT was first hyped [261–263], then overestimated, and then slowly became practical [264]. The term IoT was probably coined towards the end of the last millennia [265], but its roots are older, albeit unnamed. The functional similarity of CPSs [136] and Ubiquitous Computing [46], as well as stipulations from early thinkers such as Nikolai Tesla ([232], p. 122), make its historical provenance a question of choosing a remotely related cognate. Following Atzori and colleagues [232], we take the probable origin of the term “Internet of Things” by Ashton in 1999 [265] as a starting point to limit ourselves to cases where the general idea of IoT was already the goal of research.

Nearly anything remotely related to Information and Communication Technologies (ICT) [264] is included in modern IoT research [47,165,247,266–277] because IoT promises to be the invisible glue to digitally represent the world and offers new means of analysing and understanding [46]. In part, this may be due to the well known term Internet that immediately allows people to associate it with global information infrastructure [158]. On the other side, a Thing is vague enough to label any entity that can be distinguished by some measure [158]. The Intranet of Things in a WSAN has been transformed through ICT [277–279] and cloud-centric processing [268,280,281] to the IoT. This loose description is what allowed the IoT to soak up many kinds of ICT developments for the last decade (see Figure 12).

While the term IoT has recently become the most common term for the linkage between the virtual and the real, that was not always the case. Previously, a set of competing ideas (e.g., CPS, WSAN, and Ubiquitous Computing) co-created the ecosystem that today is generously labelled IoT and which laid the hardware foundation for DTs. These ideas were closely related but ultimately based on different assumptions because they originated from different fields. According to public perception, the terminology has settled on IoT even as the underlying technology is still quickly diversifying and expanding. With current trends, DTs appear to be a concept that is set to overtake the IoT in terms of public perception (see Figure 12).

Another issue is that IoT is a buzzword that is difficult to grasp because it is overloaded with meaning. Simply tapping into the literature is overwhelming and, at the same time, highly specific. The many technologies peddled under the umbrella-term IoT would be more distinctive if labelled specifically according to their function. While the technological umbrella-term is now well established, if ill defined, there is still a lingering sense that different fields of application are uniquely distinct and require special attention. Research and development is thus heavily directed towards five distinct real-world applications that promise the biggest gains: industrial application, urban management, health care, vehicular applications, and end-user applications [47,165,202]. This ingrained belief regarding the distinction leads to difficulties in synthesising common knowledge across applications. Overcoming this barrier could help the interdisciplinary fields of IoT and DTs in many ways. Not only would it enable learning from each other to clarify their theoretical backgrounds,
but it would also make the use of IoT as a basis for DTs more accessible. Indeed, a clearly communicated and unified approach to IoT would lower the barriers of entry tremendously.

Weiser [46] postulated the goal of Ubiquitous Computing to make the computations disappear in the mundane of the environment. To obtain services that hide in plain sight, three categories of design considerations have been brought forward that need to be considered for IoT to work: fundamental properties of a Thing, network communication properties, and real-time interaction properties. The fundamental properties of a Thing include existence, sense of self, connectivity, interactivity, dynamicity, and sometimes environmental awareness [178,191,239]. Much research has also been conducted on properties to make network communication more efficient, including energy, latency, throughput, modular design for improved scalability, topology, security of information, safety [42,178,239], deployment, mobility, cost, efficient resource usage, heterogeneity, communication modality, infrastructure, coverage, network size, and lifetime [178,239]. Real-time requirements in industrial and healthcare settings have brought forth another set of properties that need to be considered to enable real-time interaction with IoT networks: predictable delays and latencies, congestion-free communication, accurate time synchronisation, real-time assurance, interoperability with existing infrastructure, and prioritised communication [178].

The historic development IoT and the evolution of definitions has been addressed in detail by [232]. The work identified three generations of IoT development with different foci that have resulted in the current diffusive definitions (see Table 3). An additional new generation is identified in this review paper that exemplifies the most current trends of DT research in IoT that were not reported in [232]. Notably, while these four generations focus on distinctive features, they are effectively contemporary because these generations are still actively researched. To make sense of these concurrent generations, it is best to understand each generation as a necessary condition to start work on the following generations.

Table 3. Waves of IoT generations.

<table>
<thead>
<tr>
<th>Generation I</th>
<th>Generation II</th>
<th>Generation III</th>
<th>Generation IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earliest Mention Technologies</td>
<td>1999 RFID M2M</td>
<td>2010 Inter-networking Web of Things</td>
<td>2011 Cloud computing ICN</td>
</tr>
<tr>
<td>Focus</td>
<td>Object integrations Networking</td>
<td>Constrained devices</td>
<td>Semantic IoT Centralisation</td>
</tr>
<tr>
<td>Source</td>
<td>[232,265]</td>
<td>[232,282]</td>
<td>[232,268]</td>
</tr>
<tr>
<td>2012 Fog Computing DTs</td>
<td>Opportunistic IoT Decentralisation own definition,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[2,283,284]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first generation is driven by the rise of RFID tags [265] to digitally distinguish Things. At this time, IoT was often taken to be synonymous with RFID. This generation is tightly coupled to Ubiquitous Computing [46], Machine-to-Machine (M2M), and CPSs as the RFID technologies ultimately derive from these research fields. However, by the late 2000s, most of these technologies had been tagged IoT. The second generation is driven by interconnecting Things instead of just centralising the information flow. It focuses on the network component of IoT devices and realistically brings the Internet into the IoT [282]. Control of information becomes more central as well, reinforcing the connections to its roots in CPSs [136]. The third generation is driven by a re-centralisation around cloud technologies that matured meanwhile [268,280,281]. This generation focuses on processing data and generating knowledge. The fourth generation is driven by a second wave of decentralisation due to time sensitivity of applications, fog computing [283], and the rise of DTs [2] and servicisation [212]. As questions of processing have been addressed, the newest generation brings more efficiency to IoT. IoT is often mentioned as a multiplier that makes other technologies more attractive [232], contributing to the diffusion of the term.
4.3.6. M2M

Machine-to-Machine (M2M) communication allows Things and network components to interact automatically with each other without human actions, thus enabling automated formation of CPSs or DTs on different scales.

The complexity of an IoT system underpinning a DT requires much communication between each involved Thing and the network infrastructure to support a control system \[136,234,285\]. These communication protocols have been aggregated and named M2M communication. M2M is an ongoing research field and is usually required for all other IoT-related and DT-related activities \[286–289\]. Essentially, M2M describes how IoT components interact with each other and can be understood as part of the implementation of the Connection Environment of a DT. As such, M2M, IoT, and DTs cannot be meaningfully separated, although they do not constitute the same concept. Before IoT became the defining word for ICT, M2M was a contender for being the umbrella-term for these technologies.

The number of devices connected to IoT today surpasses the number of devices manually connected to the Internet. Thus, it is impossible to manually manage the communication between devices by connecting them or maintaining them \[178,290–292\]. This aspect triggered the development of the M2M concept in the early 2000s \[288,293\] followed by standardisations such as oneM2M \[153\] in the 2010s. At the centre of M2M activities are the protocols that are used to exchange information between devices. Protocols describe predefined procedures of how to initiate, maintain, and reestablish communication \[153\]. While some protocols are high-level, they are often bound by the physical limitations of the communication capabilities of the constrained devices. The complexity of the classical standard Internet Protocol (IP) was quickly disqualified as a platform \[294\] because of the assumptions that the computing infrastructure was powerful, wired, and connected and that communication was a negligible overhead. However, constrained devices have been highly optimised in terms of energy consumption, size, and costs such that communication has become the major bottleneck \[295\].

The physical infrastructure to connect constrained devices to the network is usually referred to as physical/PHY \[181\], and the data link/MAC address is used to uniquely identify single units \[296–298\]. These terms originate from the OSI model and refer to the physical and data link layers between devices \[299\]. Originally, a link would be a copper or fibre cable. However, wireless transmission has adopted the same terminology, so PHY usually refers to some kind of antenna that uses a pattern of signals to encode information \[251\]. To enable a more efficient use of bandwidth and potentially allow for more devices, Cognitive Radios allow secondary users to use communication channels while primary users are idle \[287,300,301\]. The main trade-off for the PHY connection is between data rates, connection distances, and energy consumption. Currently, technology can be approximately grouped into four application areas: Body Area Networks (BAN), Personal Area Networks (PAN), Local Area Networks (LAN), and Wide Area Networks (WAN) (see Table 4 for an overview).

<table>
<thead>
<tr>
<th>Scope</th>
<th>Abbreviation</th>
<th>Technologies</th>
<th>Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body Area</td>
<td>BAN</td>
<td>BLE, Zigbee</td>
<td>[207,302–305]</td>
</tr>
<tr>
<td>Personal Area</td>
<td>PAN</td>
<td>UWB, Bluetooth, WiFi</td>
<td>[306,307]</td>
</tr>
<tr>
<td>Local Area</td>
<td>LAN</td>
<td>WiFi, LTE</td>
<td>[307,308]</td>
</tr>
<tr>
<td>Wide Area</td>
<td>WAN</td>
<td>LoRaWAN, Sigfox, NB-IoT</td>
<td>[20,309–314]</td>
</tr>
</tbody>
</table>

The large standardising bodies, IEEE and IETF, approach protocols in constrained environments based on expansion of the IEEE Standard 802.15 \[315,316\] and the IETF LLN protocol \[294,295\] to handle IP. In addition, there are technological niches that have not been fully covered by those standardising organisations such as the Low Power Wide Area
Networks (LPWAN) where protocols vary and are highly incompatible, representing a low level of industry cooperation [309–312]. However, several consortia have developed around the different core technologies associated with each protocol, and only subsequent years will show how prevalent each WAN technology has become.

The commonality among all protocols is that they are simplified to move data from the constrained devices to some kind of translation unit that transfers the data into the IP protocol, making it widely available on the Internet [317–319]. This translation is necessary to cope with the reduced computing capacity of constrained devices as well as questions of how to transmit data and scale local IoT networks for use in a DT. A difficulty arises for these protocols because the translation unit of constrained devices not only sends data (such as with sensors) but may also receive data (such as with actuators). This aspect requires a bidirectional translation protocol that can uniquely address every node from the outside [320].

4.3.7. Computation-on-the-Wire/Dew/Mist/Fog/Cloud and Edge Computing

Computation-on-the-wire computes all necessary transformations on data at the best possible point in the network according to the subsidiarity principle to deliver Ubiquitous Computing for CPSs such as DTs.

The trend to locally compute, distribute, and store data in end-user devices and local clusters rather than centralised data centres [283,321] is a reaction to the third generation of IoT’s cloudification [232]. At the same time, the trend away from cloudification is the foundational moment for the fourth generation of IoT. Furthermore, sensors generally pose a trade-off between computational power, latency, cost, accuracy, energy consumption, and security [42,191,322,323]. To balance these trade-offs, additional computational power was introduced physically closer to the sensors, which also limited data transmission and reduced the load on the network. The placement of computational units in the physical world to provide the necessary computing capacity for Ubiquitous Computing produced a set of hydrologically inspired terms to describe different trade-offs. The division of remote computing into cloud [186,236,324–328], fog (also edge or grid) [150,151,283,322,327–340], mist [341,342], and dew [341,343–346] is an attempt to follow this principle (see Figure 13). While the analogy may have been taken a bit too far, this simile is quite adept at equating the distance from the observer to the water droplet and the distance from the device to the computation location. At the same time, the high level of differentiation and the unclear boundaries between these concepts—sometimes used interchangeably—makes the discussion more complicated. For example, fog, mist, and dew computing are often used interchangeably across papers, and edge and grid computing often fit partial definitions of the hydrological terminology.

Another, less metaphorically laden, view focuses on computation-on-the-wire to describe the same phenomena [347]. This perspective assumes that data need to be manipulated between the data generation point (i.e., a Thing at a physical location) and the data consumption point (i.e., a user inquiring about something related to the PT via its DT). To be the most cost-effective, the computation is required to take place as close to the generation and consumption points as sufficient and as far away as necessary. Therefore, the computation-on-the-wire concept is essentially the application of the subsidiarity principle. Taken from political science, subsidiarity implies that a central authority should only perform tasks that cannot be performed more locally [348]. The principle is most famously applied in the European Union and should be taken as the guiding principle for IoT as well. On a more abstract level, it implies that actions regarding an entity should be taken (a) as close as possible to an entity and at the same time (b) as far as necessary to successfully execute the action. Applied to computation-on-the-wire and DTs, this principle implies that computations for DTs should occur (a) as close to a device, a requester, or the PT as possible and (b) as far into the cloud as necessary to successfully compute the output. The concept elegantly joins the hydrological simile and the underlying trade-offs by applying the subsidiarity principle [348].
Remote sensing has two meanings. First, it provides observations recorded by a distant measurement device with some form of camera. Second, remote sensing refers to distant access to a local measurement device. Both can be processed to represent the Physical Environment around a CPS (such as a DT) and to possibly interact with the Physical Environment.

Remote sensing can be easily misunderstood as either exclusively meaning distant measurements based on some kind of camera [349] or using remote access to sensing units [350,351] to measure local properties from afar such as in agriculture [352], healthcare [353], Smart Cities [354], or everywhere [355]. Traditionally, remote sensing exclusively refers to camera-based, often airborne, approaches [349]. However, the secondary meaning of remote sensing as a shorthand for “remote access to local sensing units” has become prevalent across nearly all applications that do not apply traditional remote sensing. Whereas there is a clear difference in the meanings of remote and sensing in both interpretations, recent applications combining the two such as sensor networks on Unmanned Aerial Vehicles (UAV) often linked to ground stations and the Internet [356,357] may make the difference irrelevant in the long term. More recently, for the context of the DT, the precise sourcing of information from the Physical Environment is not important. Instead, the accuracy of the measurements must be reported to allow for a decision-making process that accounts for uncertainty. Throughout this review, we understand remote sensing in both its meanings as the main source for information from the Physical Environment. Furthermore, we assume that any form of traditional remote sensing can be made available in real-time as a virtual sensor in an IoT network [149,358,359].

5. The Fused Twins Concept

DTs require visualisation to unfold their potential. Naturally, it is the Virtual Environment [20,21,101] of a DT through which information be accessed. To make sense of
the different forms of visualisations, we use the Reality–Virtuality Continuum [360], see Figure 14 to classify a broad range of human interface devices.

![Figure 14. Devices to access Digital Twins on the Reality–Virtuality Continuum [360]. On the Continuum, FTs span instances of Extended Reality (XR) input modalities and media from Augmented Reality (AR) via Mixed Reality (MR) to Augmented Virtuality (AV). Typical media include tablets and phones but also holographic displays and smart objects. Reality-embedded access to DT focuses on abstract Dashboards, whereas virtuality-embedded access focuses on remote analysis of the DT. The FTs spectrum covers direct interaction with the PT that is fused with augmentations from the DT either displayed via AR on tablet, MR on HoloLens, or haptically through AV smart objects. Generated with http://draw.io.](image)

Commonly used implementations of DTs make use of visualisation types on both ends of the continuum which includes the Dashboard and Virtual Reality (VR)-based digital models. Dashboards [361] reside on the reality end of the spectrum because they display information independent of the physical world on screens which display numeric visualisations (note that reality here is the continuum instance in which the user is located and not the medium through which the application is experienced). The visualisation is not embedded directly in reality and is accessed in a non-immersive context. On the other end of the spectrum, remote monitoring [362,363] is located in the virtuality end and can be implemented both as VR and non-local AR. The information usually has a digital model that is displayed in the virtual space with no direct relations to the physical space in which the data are consumed.

The FTs concept captures the situation where the DT is queried from within the PT allowing for a fused in situ representation. A medium is used to connect the PT with the DT through the display of Situated Analytics [100]. Possible media range from tablet-based (AR) and HoloLens (MR) to SOs (AV). AR and MR devices can be used to overlay immersive analytics of the PT with the DT. In contrast, SOs [7] are interactables in the environment that allow users to manipulate the state of both the PT and the DT.

Despite the similarities, FTs are different from situated visualisations [19]. FTs require a fully fleshed DT behind the situated visualisation that includes all five types of environments and the ability to compose the underlying DT with other DTs that may be spatially or logically connected. FTs use the situated visualisation as an entry point to interact with all systems that are connected to that particular point in the PT to provide access to the whole PT through the DT. In contrast, many situated visualisations focus on local applications such as smart switches for IoT devices [364]. However, at first sight, the situated visualisation is often indistinguishable from FTs for end users. Nonetheless,
continuous interaction with the tool eventually reveals the limited functionality. Thus, any situated visualisation can be expanded to implement the FTs paradigm by integrating it into a DT model.

This section will outline the DT Requirements to implement the FTs paradigm. The particular features of the main environments are discussed, but the Data Environment and the Connection Environment are not discussed in detail. The Connection Environment has not been implemented as interoperability between DTs or their components is as of yet uncharted territory. Consequently, the Data Environment is usually a proprietary implementation that requires no further discussion here. An open standard for DTs would open up the research on the Connection Environment and the Data Environment. Lastly, while discussing the Virtual Environment, we reflect upon current situated visualisations that are close to implementing the full FTs concept.

5.1. The Physical Environment of the Digital Twin

The Physical Environment encapsulates the infrastructure to cast the Digital Shadow [152] of the PT (see Figure 8) into the Data Environment. When assembling an IoT or remote-sensing system for a DT, it is important to be able to determine the dimensions which matter most for the functioning of the DT and choose an architecture accordingly. Most research on architecture has been done in the context of IoT. However, as time progressed, architectures focused more on how to deal with the data and consequently today, these architectures are also used in DT cognates and DTs themselves.

International institutions have developed IoT reference architectures [365–367]. Architectures are usually presented as layers. However, multiple architectures exist with a similar number of layers but different functional goals. Going into detail of how each layer is supposed to work sometimes contradicts other architectures with the same number of layers (see Figure 15). On the one hand, there is no unified architecture due to the sheer endless number of application scenarios across different branches of research. On the other hand, architectures are sometimes compatible, and we provide a mapping of the different components in Figure 15. These architectures may appear as fully fleshed DTs at first glance but on closer inspection will reveal themselves to be mostly concerned with the Physical Environment and the Data Environment. However, depending on the overall goal of the architecture, the upper layers may actually contain features of the Analytical Environment and the Virtual Environment. IoT architectures are the basis for most current DT applications and will probably morph into fully fleshed DTs in the coming decade.

Architectures aim to be somewhat self-explanatory and ideally each layer covers exactly one functionality on transforming the data. Realistically, layers help conceptually to assign tasks to different parts of the real system. However, most implementations rely on a messier, less clean interaction that takes place between the components. To choose the layer, it is typical to consider aspects of security, energy consumption, data collection density (spatial and temporal), maintenance, availability (real time vs long-term evaluation), interactivity (explicit reaction versus implicit reaction), and heterogeneity [42,178,191,239]. Lastly, management of the infrastructure and security concerns relates to every layer and is often represented as vertical columns to the side of the layers [47,165,247], but they are not shown in Figure 15 for brevity. When comparing this to our DT design, it is those management layers that map to the Connection Environment that would ultimately allow to compose DTs given appropriate permissions.

While the number of architectures is infinite, we will focus on a few of the most common approaches in more detail here. The simplest architecture divides the world into two layers, the Physical layer and some abstract layer such as Digital [2] or Cyber [234,236]. Mapping these realms to our DT model, the Physical covers the Physical Environments and the Digital/Cyber covers the Data, Analytical, and Virtual Environments. The Connection Environment maps most closely to the Management and Security columns in previous architectures but the focus is more on interoperability while also satisfying management and security requirements. This architecture has no representation of how data are collected,
transported, transformed, and attributed in the DT. Nonetheless, it is a powerful vision that is easily understood despite the technological intricacies necessary to make such a vision possible. This perspective strengthens ontological approaches to describe processes and qualities of objects in such a way that allows for interaction. This concept is often applied in Industrial IoT (IIoT) where it is usually used as an end-user perspective on a Service-Oriented Architecture (SOA) [154].

**Figure 15. Comparison of Digital Twin Architectures.** Data flows from the bottom (Thing) to the top (User). The functionality of each layer is displayed in comparison to other architectures where possible by spawning each layer vertically such that it contains all functionality that similar layers in other architectures provide. For example, Presentation and Session in the OSI Reference Model are subsumed in Application in the IETF LLN Stack. The labels at the top provide the name of the models, and the labels at the bottom references the origin of the respective model. From left to right: OSI Reference Model, Variation of IETF Internet Model [316], Digital Twin [2], CPS [234], ETSI M2M [165] and SENSEI [368], SoA [154], 5C [155], and Context Aware [156]. Generally, management and security columns spawning all layers are not shown. Our own definition is mapped to these architectures in the last column. Generated with [http://draw.io](http://draw.io).

For practical implementations, often a three-layer architecture is used that describes the detailed work in the middle [165,368]. These architectures are already able to address how the DT arises and as such a third layer is brought to attention: the middleware. The physical world is sensed in a perception layer mapping to the Physical Environment. Then, the data are structured into semantically meaningful units in the middleware layer mapping to the Data Environment and the Analytical Environment and finally presented to end users in an application layer mapping to the Virtual Environment. This architecture belongs to the earlier models and is focused more on the process of how to create basic IoT infrastructure. The perception layer describes the process of data acquisition usually provided from sensors, actuators, and tags. The middleware layer aggregates the data and provides APIs to access necessary information. When comparing the three-layer model to the two-layer model, middleware is the hidden layer that converts the real world’s PT into its DT. However, most real systems stop short of complex analysis based on simulation and often limit themselves to presenting raw data or aggregates.

The application layer offers useful services on top of the data. In a sense, those services are offered in the Virtual Environment of the DT (see Figure 2), but they heavily rely on the Analytical Environment to produce interesting services. Naming conventions for the three layers vary, but their basic functionality remains the same such as in ETSI M2M [165], SENSEI [368]. Additional layers always focus on very specific features and mostly subdivide the Application layer for Service-oriented Architecture [154], Resource-oriented Architecture [162] business applications [155], or Context Awareness [156].
5.2. The Analytical Environment of the Digital Twin—Models and Services

The Analytical Environment is often considered to be the DT itself [152] because it simulates the PT behaviour. However, throughout this review paper, we have outlined more specific tasks for the Analytical Environment such as running simulations, producing aggregations, and providing services. We also have elaborated on why a DT is more than just a simple simulation with real-world data. Nonetheless, it is the Analytical Environment that usually gives DTs and their cognates the “smart” label such as in Smart City.

The word smart generally alludes to the superiority of a (connected, digitised) Thing compared to its dumb (unconnected, analogue) counterpart. Since connectivity and digitalisation on their own do not make anything smart but only provide accessibility, the smartness implies easy interaction with the Thing, new insights into the Thing, and new applications of the Thing. To attain any of these goals, models are required to represent reality, predict the future, and monitor ongoing processes. For DTs, this means that services [369,370] are provided on top of the data through the Analytical Environment to enrich a PT’s functionality and simplify interactions. In general, most Things will require a DT to become smart.

However, the term smart has been misused. The previous description of smart is applicable to any meaningful service. The use of “smart” as a prefix for services has proliferated to the point where it has become superfluous. Following the above argument, the prefix has been attached to anything that could have a DT or has implemented features that are indistinguishable from a DT. In the simplest case, a Smart Object provides awareness, representation, interactivity, and identity [7,258]. Smart Objects are concrete implementations of a DT (and even the FTs paradigm), but they limit the Thing to be a small-scale object with a fixed use. Smart Objects are also a form of data physicalisation [123] that provides the opportunity to interact with data beyond a visual medium.

As objects increase in size and become places, the terminology adapts to Smart Environment [321], Smart Home [371,372], Smart Office [372], Smart Museum [373], Smart City [4,10,48,199,200,254,324,371,374–376], and even Smart Regions [377] and Smart Communities [248,378]. Conceptually, these Smart Places [379] do not differ from Smart Objects [7] but only generalise them to sets of Things. This scaling aligns with the composability of DTs and allows for these concepts to be nested.

However, there is a difference to be made between physical Things and non-physical Things such as processes. The Thing in this context is a non-physical entity but interacts with physical entities in a defined way allowing to define smart processes such as Smart Governance [380], Smart Manufacturing [381,382], Smart Construction [383], Smart Health Care Services [384], Smart Mobility [385], and Smart Experiences (for touristic activities) [386].

With everything labelled smart, it becomes more meaningful to talk about the kind of actions a Smart Thing can provide. Interestingly, it is these interactions that are also sought after in the immersive analytics context [68]. We locate actions on a scale of how independently they can react to external input:

1. Sensing: Only produce data;
2. Reacting: Use threshold on collected data to trigger action;
3. Interacting: Use additional input beyond collected data to change behaviour;
4. Analysing: Use a model to interpret collected data when choosing actions;
5. Proposing: Use a decision model to propose actions based on data models;
6. Deciding: Autonomously perform all necessary action to interact with data.

The action space of Smart Things is constantly expanded and we cannot enumerate everything here. However, all of these developments can be understood as the Analytical Environment of a DT. Research has covered many approaches to provide the Analytical Environment needed to transform data into insights (see Table 5).
Table 5. Research Areas for Digital Twin Analytics.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Description</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontologies</td>
<td>Make semantic sense of the data and produce knowledge and context</td>
<td>[156,225–227,387–389]</td>
</tr>
<tr>
<td>Middleware</td>
<td>Infrastructure to identify, process, store, secure and present information</td>
<td>[228–230,390,391]</td>
</tr>
<tr>
<td>Data preprocessing</td>
<td>Data aggregation and data fusion</td>
<td>[304,392–398]</td>
</tr>
<tr>
<td>Data mining</td>
<td>Cloud-based Big Data analysis on unstructured data</td>
<td>[399–406]</td>
</tr>
<tr>
<td>AI&amp;ML</td>
<td>Machine learning and deep learning</td>
<td>[396,407–409]</td>
</tr>
<tr>
<td>Blockchain</td>
<td>Security, trust, and privacy through smart contracts</td>
<td>[410–414]</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction for improved access</td>
<td>[70,85,86,260,415–417]</td>
</tr>
<tr>
<td>ABM</td>
<td>Agent-based modelling Objects</td>
<td>[418–421]</td>
</tr>
</tbody>
</table>

5.3. The Virtual Environment of the Digital Twin—Access, Interaction, and Representation through the Fused Twins Paradigm

The last environment we discuss provides the human–computer interaction [422,423] with the DT. When implementing the FTs paradigm for a DT, the main focus lies on AR interactions with the DT. AR is only beginning to arrive in the world of SOs, the IoT, remote sensing, and DTs but has already been accepted as a crucial feature in the near future [424]. AR is of interest for FTs as it explicitly supports embodied interaction: interaction that is not agnostic to its physical and social context [74,80]. AR can also reduce negative split-attention effects [117,119], and appropriate visualisations of data can lower cognitive load [73] and enhance cognition [72]. The expectation is that AR will offer a new avenue to simplify interaction with SOs, the IoT, and DTs [17]. The FTs paradigm is a formalisation of the interaction between AR and DTs in a situated context.

Because the combination of technologies is only emergent now, most published material towards this end is either a prototype [17,101,364,425,426] or a proposal for future research directions [427–430]. The main reason is that applying AR on top of the already complex combination of SO, IoT, remote sensing, and DT stacks requires many software systems to run smoothly that are currently only being prototyped.

The fusion of the digital representation and the physical entity lies at the core of the FTs paradigm [17] but also of the more general Situated Analytics [19]. Interestingly, other research on the topic of Situated Analytics has often implemented FTs to some degree but without the explicit mentioning of the DT aspect [364,425,426]. A notable exception is the “Corsican Twins” [101] implementation of an FTs platform with an explicit DT powering the Situated Analytics. Short papers on future research directions in Situated Analytics clearly point towards FTs-like implementations [427–430]. We summarise ongoing and future research that is already covering a wide variety of topics (see Table 6). Some of the research explicitly mentions the DT approach or a Situated Analytics approach but often simply uses the available technology to demonstrate that their problem can be addressed [428,430,431], implicitly developing a DT and applying Situated Analytics.

Table 6. Recent and Ongoing Research Applying the Fused Twins Paradigm.

<table>
<thead>
<tr>
<th>Application Type</th>
<th>Year</th>
<th>Physical</th>
<th>Data</th>
<th>Analytical</th>
<th>Virtual</th>
<th>Connection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure Management [432]</td>
<td>2009</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>City Information [18]</td>
<td>2009</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>History [431]</td>
<td>2013</td>
<td>*</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>City Information (concept) [433]</td>
<td>2014</td>
<td>*</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>Shopping [434]</td>
<td>2015</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>IoT device interaction [425]</td>
<td>2016</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>City Information [435]</td>
<td>2016</td>
<td>*</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Shopping [436]</td>
<td>2018</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 6. Cont.

<table>
<thead>
<tr>
<th>Application Type</th>
<th>Year</th>
<th>Physical</th>
<th>Data</th>
<th>Analytical</th>
<th>Virtual</th>
<th>Connection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Campus [437]</td>
<td>2018</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Evacuation [438]</td>
<td>2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine [439]</td>
<td>2018</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facility Management [440]</td>
<td>2018</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rehabilitation [441]</td>
<td>2019</td>
<td>✓</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>IoT device visualisation [426]</td>
<td>2019</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seat Selection [442]</td>
<td>2020</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Equipment Maintenance [443]</td>
<td>2020</td>
<td>*</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In situ visualisation [101]</td>
<td>2020</td>
<td>*</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education [444]</td>
<td>2020</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training &amp; Assistance [445]</td>
<td>2020</td>
<td>*</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firefighting [430]</td>
<td>2021</td>
<td>-</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General [17]</td>
<td>2021</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>IoT device interaction [364]</td>
<td>2021</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General [100]</td>
<td>2022</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * Partially implemented or insufficient information.

Since the FTs concept in particular and DT concept in general is not well defined yet, we analysed the literature for traces of it. To construct Table 6, we analysed the papers for the presence of implementations of any of the five components. For the Physical Environment, we considered it as fully implemented (i.e., we gave it a check-mark (✓) in Table 6) if active sensor data collection was reported, or we considered it as partially implemented (i.e., we gave it a star (*) in Table 6) if passive sensors or markers were reported (e.g., QR code stickers). For the Data Environment, we considered it fully implemented when explicit data collection and storage was reported (e.g., SQL databases or BIM models), and we accepted partial implementation if data were streamed but not permanently stored. For the Analytical Environment, we considered it fully implemented if there was an element of only analysis, and we considered it partially implemented if simple data transformation was performed. For the Virtual Environment, we considered it fully implemented if some digital model was reported (e.g., 3D models of rooms or machines) or partially implemented if a digital representation could be inferred but was not explicitly reported (e.g., spatial representation of labels). For the Connection Environment, we considered it fully implemented if it was possible to add or compose elements in different environments such as adding new sensors, changing the Virtual Environment, or using different analysis and partially implemented if it was theoretically extendable or if the reports hinted at extendability of a single component (e.g., adding more sensors).

Across all partial implementations of the FTs paradigm, we find the recurrent theme of how to tightly couple the PT and the DT. All DTs require an effective representation of the PT to convey meaningful information to the user. For the built environment, this commonly involves Building Information Models (BIM) [17,428,430,437,446], or City Information Models (CIM), and the Geographic Information System (GIS) [428,433]. In a city context, usually, BIM/CIM is used for the local model, whereas GIS provides a global reference frame. However, if BIM/CIM is not available, remote sensing offers methods to produce placeholder models [440]. Both GIS and BIM/CIM have a long research tradition and have been used in a variety of ways in industry, academia, and governance. It could be argued that they provided DTs before the word was created. However, in the DT literature, a consensus has built up that they rather provide a digital model [152,171] and are at best the basis for a DT.

The reason that these opinions have developed is that BIM/CIM and GIS are often seen as a rather static representation of DTs and that, for dynamism, it is required to continuously measure from sources such as IoT [17,425,428,430]. These measurements are
often called the Digital Shadow [152,171]. Additionally, the raw data provided by sensors is often classified as Big Data [447,448]. Visualising the data becomes nearly impossible without analytical tools. This highlights the strong interaction between the Analytical Environment and the Virtual Environment. At the same time, a DT is characterised by the bi-directional interaction between the PT and DT. Therefore, there is an expectation for actuators [449] to be able to be triggered from within the Virtual Environment to implement the full functionality of a DT.

Beyond an accurate model, data acquisition, and actions, FTs require a mapping between the virtual world and the physical world. Many approaches are available, and they commonly identify features in the physical world that can be used as anchors for the virtual world. These location services are based on different technologies such as Bluetooth-based [364], image-marker-based [101], vision-based [17,425], or signal-based [426]. Consequently, it is possible to display content in the Virtual Environment as an overlay in physical world. Depending on the medium, it is possible to interact with the augmented content through poses, gestures, mid-air interfaces, tangibles, touch, voice, face, or gaze [450–452]. Immersive and Situated Analytics [19,68] provide the research field in which most of the visualisations that will be used in FTs will be developed.

Implementing FTs requires some expansion of an existing DT. In the Virtual Environment, services for the digital model, the human–computer interaction, and the location model must be integrated to be able to effectively display the measurements from the IoT and remote sensing in situ. The services first locate data in relation to the BIM/CIM [437,446] and then locate the BIM/CIM in the physical world through Computer Vision and GIS [453]. The result is displayed in AR (see Figure 16) and offers Situated Analytics to the observer.

**Fused Twins Torus Stack**

![Fused Twins Torus Stack](http://draw.io)

**Figure 16. Fused Twins Torus Stack.** Generalised technology stack for the FTs paradigm based on [17]. The stack implements a full DT and expands the Virtual Environment to accommodate the FTs. Data flow is shown through the arrows. The stack becomes a torus because the chain of technologies and environments requires circular links of the data. Changes to a DT in the Virtual Environment flow back to the PT through the Data Environment. Generated with http://draw.io.

### 6. Discussion

To understand FTs, we have reviewed three distinct but connected literatures. First, this review has focused heavily on how to understand DTs today. The fuzziness around their definition has contributed to the continued vagueness of nearly any ICT being labelled
or associated with IoT and DTs. However, with the rising practicability of DTs, no aspect of life will be beyond digitalisation for good and bad. Second, we looked at AR technology as a means to interact with information and Situated Analytics in particular to cover the interaction with a DT in situ of its PT. This combination is forming the basis to fuse the twins. Third, we explored the tightly connected topic of Smart Cities which often has been a canvas for testing ICT, IoT, DT, and AR applications. We expect the FTs concept to be applied often in this context, even under a different name or no name at all. We have enumerated a large number of FTs applications that partially qualify already today to give an idea of what is to come.

6.1. Smart Cities Discussion

Smart Cities are still a far way from being smart. In particular, because most “smart services” are not accessible. DTs offer an organisational paradigm that could make the complex Analytical Environment of cities accessible by flattening the hierarchy from the user perspective. The DT abstracts the heterogeneous technological structure of a city into semantically meaningful units that can be queried for information extracted from the underlying networks. At the same time, the DT also offers an opportunity to overcome the heterogeneity of sources from IoT to remote sensing without loss of accuracy. Lastly, DTs building on top of heterogeneous sources can represent abstract concepts that have no physical counterpart such as the livability and sustainability of a city [454] or the spatial movement of traffic jams [252].

A fully implemented DT would impact the management, planning, and use of a city. It would allow its economy to become more efficient, markets to be more transparent, processes to be explorable and problems possibly to be more addressable. If a DT of a city is made accessible to the general public, it could even serve as a participatory tool and integrate the city’s population. At the same time, concerns of privacy and security will not be easily resolved and a delicate balance between gains and losses has to be struck to realise DTs of cities.

6.2. Situated Analytics Discussion

AR combines real and virtual elements registered in three dimensions and provides interaction in real time. Despite decades of development, immersive analytics are still in their infancy, and no standard user interfaces have been established [125,455]. Recent reviews have further found that, for Situated Analytics, there is still a lack of an overarching definition, data representation is not unified, and the target audience is not defined [125]. Broad adoption is only now within reach, and the future will bring many new developments.

Interestingly, the description of how to assemble a Situated Analytics system maps closely to our layout of a DT (see Figure 7). There is a synergy between AR and the FTs concept: the FTs concept offers one of the most promising applications of AR, while AR supports an implementation of FTs that reduces split-attention effects and cognitive load. In this context, the FTs paradigm combining a DT of Smart Cities with Situated Analytics in AR integrates the physical and social aspects of the DT and thus aligns with the theories of embodied interaction, embodied cognition, situated cognition, and socially situated cognition [74,103–105]. Whether Situated Analytics live up to the expectation of our working hypothesis will require more rigorous research on the impact of embodiment, cognitive load and enhanced cognition. Nonetheless, we argue that the FTs concept may become the largest application of Situated Analytics to be seen, used, and even embodied by most people on the planet regardless of formal verification.

6.3. Digital Twins Discussion

We have reviewed the history of the DT to disambiguate the concept from its cognates. There is a fine line between DTs and CPSs, M2M, WSANs, SOs, Ubiquitous Computing, IoT, computation-on-the-wire, and Remote Sensing. In the context of this review, we have shown that most cognates are actually underlying technological approaches or perspectives
that constitute components of the DT. The sole exception is a CPS which is a more general approach with the DT being a specialised instance thereof. We believe that DTs have the potential to subsume current cognates similar to how the IoT has been previously considered the most encompassing term in this area. At the same time, each cognate carries an important meaning that is worth maintaining, and we strongly recommend to understand these differences in order to gain a deeper understanding of DTs.

Most DTs “in the wild” are prototypes that were created as novel approaches to particular problems but whose theoretical background is underdeveloped beyond the definition of an ideal DT with similitude that is impractical for nearly all applications. Indeed, most research on practical DTs lacks theoreatisation beyond the assembly of the system. Consequently, a debate arose on the differences between digital models, Digital Shadows, and DTs. This is beginning to change with the theoretical work of [8] which first developed a more sound differentiation of DT components. We follow the initial ideas from [23] but diverge in some details. We generalise the theoreatisation into five general components (see Figure 2) that we rename the Physical, Data, Analytical, Virtual, and Connection Environments of the DT. Our theoreatisation allows us to resolve the differences between models, shadows, and twins to partial implementations of the components of DTs.

Ideally, these environments work in isolation and are interoperable to enable the composition of DTs on the level of components as well as nesting. Furthermore, our definition covers most DTs and DT-like applications by assuming partial or implicit implementations of its components. Throughout this review, we have gathered evidence that cognates, other DT models, Smart Cities, and even Situated Analytics have theoretical correspondences to the components in our DT model.

At the same time, it is difficult to assemble the different environments. The serviciation of IoT has provided an interesting foundation upon which DTs could be quickly assembled with very little overhead. However, IaaS, PaaS, and SaaS are not sufficiently separate in the real world and, under current implementations, would require a Connection Environment that leans towards a hypervisor to compensate for the unreliable interfaces between services. It would be preferable to develop DTs and their components in a more standardised way via isolated and interoperable environments. If IaaS, PaaS, and SaaS were optimised to strictly map to DT components, then DT components “as a service” could democratise the use of DTs by enabling people to quickly bring DT services together. Two open questions that remain are whether the services should be provided in a centralised or decentralised manner and whether the services are provided as a government service or a market product. These services could be more accessible with an open standard for DTs that clearly define how the different components interact.

6.4. Fused Twins Discussion

Without appropriate human–computer interaction, the data in the DT will remain insurmountably complex and thus incomprehensible. The composability of DTs underpinned by interoperable IoT and remote sensing will enable a smoother interaction between people and the environment. In this context, the FTs concept will allow for an intuitive interface to access data, information, knowledge, and understanding where it is most needed and useful. The FTs Paradigm attains this with the AR/MR/AV interaction between the DT and the PT to identify opportunities for simple interactions in the complex data in situ where the data are generated and collected. It is paramount to align the user interfaces, data formats, general interoperability, and creative services that eventually converge towards real Smart Cities for its users.

The idea underlying FTs has been around for some time, albeit under a different or no specific name. DTs and their cognates Ambient Intelligence and Smart Objects, as well as sometimes the more general Things, can all be understood as different embodied FTs. At the same time, FTs and Situated Analytics are very similar, which opens the question of whether a new term adds clarification. We do believe that a new term is necessary because we are currently observing the confluence of two topics, and the new terms expand beyond...
each individual conceptualisation. These two topics are the aforementioned research in AR that does not necessarily rely on DTs (but may possibly implement FTs without mentioning it) and the expanding applications of DTs in nearly every research area and field. The necessity of visualising the DT, and the capability to do it in augmented space rather than on a display, ushers in a new era in which data can be made visually and spatially accessible at the time of creation and across the consumption. FTs play a special role in this context where the data are fused to its spatial context which differs from classical AR which merely uses the real-world space as an empty canvas devoid of any context (in most cases).

FTs are not unique to Smart Cities, but they are especially applicable in Smart Cities. Some application areas for FTs include healthcare and industry. For example, in healthcare, human DTs have been proposed to improve individual diagnosis and care. Monitoring patients with FTs allows practitioners to locate medically relevant information in the human body. This may be especially helpful in the context of surgery during which representations of bones and other tissues may be superimposed to facilitate the surgeon’s decision-making. In an industrial context, the inner workings of machines can be overlaid onto the real machines and augmented with analytical information. While healthcare and industrial application are most useful for specialists, FTs in a Smart City context may be an accessible form of FTs for the average citizen. For example, the Augmented Reality maps from Google already embed navigation information in the physical city and arguably employ a DT in the background, making it a simple and accessible example of FTs. Similarly, in the context of tourism, museums have created DTs of specific exhibits but could be expanded to include the entire exhibition to create an FTs. In the future, FTs could also be expended to communicate city planning initiatives to the public.

7. Conclusions

In this review paper, the relevant technologies and terminology behind the FTs concept for interacting with the DT of Smart Cities have been presented. We argue that FTs constitute a new phenomena beyond Situated Analytics because they explicitly use a DT and therefore stand for a richer interaction. Suitable examples of FTs applications across different literatures have been highlighted (see Table 6). Nonetheless, a stable and applicable out-of-the-box solution to fulfil the promised FTs for Smart Cities remains to be developed and validated to answer our working hypotheses. DTs also remain to be well specified, which may cause delays in developing FTs. Here, we provided a thorough review of DTs and their cognates to address open points and a concrete definition that can be applied in retrospect to nearly all DTs. Compared to previous research that employed Dashboards and other interfaces for Non-Situated Analytics, FTs represent a special application of the Virtual Environment of a DT to provide Situated Analytics. The key technologies behind the FTs Paradigm for DTs are IoT, remote sensing, BIM/CIM, GIS, and location services. The IoT, remote sensing, BIM/CIM, and GIS are used to assemble a DT and provide analytics, and location services are used to place the Situated Analytics in the PT. Each of these technologies on its own is well developed, but only now are they being combined into prototypes that can be classified as FTs. While many researchers have theorised about similar systems that could be built, few FTs full-fledged systems have been implemented. Given the plurality of fields that have converged on this general idea, we expect that the FTs paradigm will be a common theme in future technologies and may eventually be as common as a smartphone today.

**Funding:** This research was funded by ETH Zurich grant number ETH-15 16-2 and J.G. was supported by an ETH Zurich Doc.Mobility Fellowship.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank Johanna Beyer and Zhu-Tian Chen for the feedback on the AR section. All diagrams were generated with templates from http://draw.io. And don’t forget your towel.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Abbreviations**
The following abbreviations are used in this manuscript:

- **IoT** Internet of Things
- **DT** Digital Twin
- **PT** Physical Twin
- **FTs** Fused Twins
- **XR** Extended Reality
- **CPS** Cyber–Physical System
- **UC** Ubiquitous Computing
- **SO** Smart Object
- **WSN** Wireless Sensor Network
- **SN** Sensor Network
- **SAN** Sensor and Actuator Network
- **WSAN** Wireless Sensor and Actuator Networks
- **MR** Mixed Reality
- **AR** Augmented Reality
- **VR** Virtual Reality
- **AV** Augmented Virtuality
- **IaaS** Infrastructure as a Service
- **PaaS** Platform as a Service
- **SaaS** Software as a Service
- **SenaaS** Sensing as a Service
- **SemaaS** Semantics as a Service
- **RaaS** Robots as a Service
- **RSaaS** Remote Sensing as a Service
- **RFID** Radio Frequency Identification
- **OS** Operating System
- **ICT** Information and Communication Technologies
- **M2M** Machine-to-Machine
- **BAN** Body Area Networks
- **PAN** Personal Area Networks
- **LAN** Local Area Networks
- **WAN** Wide Area Networks
- **LPWAN** Low Power Wide Area Networks
- **BIM** Building Information Model
- **CIM** City Information Model
- **GIS** Geographic Information System
- **UAV** Unmanned Aerial Vehicles
References


27. Mattern, S. A city is not a computer. *Places* J. 2017. [CrossRef]


150. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge computing: Vision and challenges. IEEE Internet Things J. 2016, 3, 637–646. [CrossRef]


Remote Sens. 2022, 14, 3095


211. Pleypys, A.; Heiskanen, E.; Mont, O. European policy approaches to promote servicizing. J. Clean. Prod. 2015, 97, 117–123. [CrossRef]

212. Agrawal, V.V.; Bellos, I. The potential of servicizing as a green business model. Manag. Sci. 2016, 63, 1545–1562. [CrossRef]


222. Glissa, G.; Meddeh, A. 6LoWPanS: An end-to-end security protocol for 6LoWPAN. Ad Hoc Netw. 2019, 82, 100–112. [CrossRef]


298. Sotenga, P.Z.; Djouani, K.; Kurien, A.M. Media access control in large-scale Internet of Things: A review. IEEE Access 2020, 8, 55834–55859. [CrossRef]


305. Miraz, M.; Ali, M.; Excell, P.; Picking, R. Internet of nano-things, things and everything: Future growth trends. Future Internet 2018, 10, 68. [CrossRef]


346. Gushev, M. Dew computing architecture for cyber–physical systems and IoT. *Internet Things* 2020, 11, 100186. [CrossRef]


352. Triantafyllou, A.; Sarigiannidis, P.; Bibi, S. Precision agriculture: A remote sensing monitoring system architecture. *Information 2019*, 10, 348. [CrossRef]


Remote Sens. 2022, 14, 3095 51 of 54


386. Kontogianni, A.; Alepis, E. Smart tourism: State of the art and literature review for the last six years. *Array* 2020, 6, 100020. [CrossRef]


424. Chettri, L.; Bera, R. A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. *IEEE Internet Things J.* 2019, 7, 16–32. [CrossRef]


453. Liu, X.; Wang, X.; Wright, G.; Cheng, J.C.; Li, X.; Liu, R. A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS Int. J. Geo-Inf.* 2017, 6, 53. [CrossRef]
