International Conference on Smart Infrastructure and Construction 2019 (ICSIC):
Driving data-informed decision-making

DeJong, Schooling and Viggiani
ISBN 978-0-7277-6466-9
https://doi.org/10.1680/icsic.64669.187

DEVELOPING A CITY-LEVEL DIGITAL TWIN – PROPOSITIONS AND A CASE STUDY

L. Wan, T. Nochta and J. M. Schooling*

Cambridge Centre for Smart Infrastructure and Construction, University of Cambridge, UK
** Corresponding author*

ABSTRACT Inspired by the UK National Infrastructure Commission's ambition of building a "digital twin" for the national infrastructure, this paper reviews some of the latest progress on digital twin development in the UK, particularly the recently published report "The Gemini Principles" by the Centre for Digital Built Britain. The paper further puts forwards six propositions on the city-level digital twin agenda, which stem from an ongoing research project on smart cities at the Cambridge Centre for Smart Infrastructure and Construction. It is argued that the development of a digital twin is progressive, where the digital twin as a technical apparatus and the social system that develops, operates and uses the digital twin must co-evolve together. On the one hand, our existing knowledge about cities and infrastructure would define how well we could build and evaluate the digital twin; on the other hand, thanks to artificial intelligence, the digital twin could potentially produce new knowledge that complements ours. In the context of developing smart cities, the paper points out that the digital twin is not a context-free technology that can be simply purchased and installed onto existing governance systems. To develop a purposeful digital twin application, the insights and changes from the governance perspective are essential. The paper also introduces the research plan for developing a digital twin pilot for the Cambridge sub-region, which serves as a case study to test the proposed propositions.

1. Introduction

We are living through the "4th Industrial Revolution", moving from a period of relative data scarcity to an era of 'digital abundance'. The sheer size of investments and efforts for building "smart cities" across the globe has epitomized the inspiration for future better cities in the digital era. In the UK, the National Infrastructure Commission (2017) has outlined an ambitious plan of building a national "digital twin" of UK infrastructure, which aims to demonstrate UK's global leadership in smart infrastructure. The digital twin brings together individual infrastructure models and the interdependencies between them through a federated system, thus representing a coordinated approach for planning, constructing and managing national infrastructure. The potential benefits of the national digital twin are apparent – optimising use of resources, reducing service disruption, increasing resilience and boosting quality of life for citizens, but how to materialise such benefits through effective policy making remains a big challenge ahead.

The relationship between smart cities and the digital twin lies in that the latter, though first coined in the manufacturing industry (Grieves, 2014), has been perceived as the new and perhaps the ultimate technological apparatus for "smartening" cities. The concept of large-scale digital twins for the built environment has found widespread favour recently in the UK, which is enabled by the fast development of digital infrastructure as well as the UK's policy initiative to digitalize the infrastructure and construction sector (HM Government, 2017). While the concept of the city-level digital twin is gaining momentum in the UK with several notable pilot projects under construction (e.g. Exeter, Bristol, Newcastle, Milton Keynes) or being considered (e.g. Cambridge), the varying approaches and the less explicit policy outcomes of these experiments imply a serious knowledge gap around the city-level digital twin, in terms of what it is, what/whom it is for and how to build/use/manage/regulate it to deliver better policy outcomes. This paper represents an early study to address the gap.

The paper is structured as follows. Section 2 reviews some recent developments of the digital twin concept in the UK. Section 3 proposes a series of theoretical propositions that address the current knowledge gap on digital twins in the context of city and infrastructure planning and management. Section 4 introduces an ongoing research project – a city-level digital twin experiment for the Cambridge sub-region, which demonstrates how the proposed propositions are being applied and verified through policy experimentation. The Section 5 concludes with implications for city and infrastructure planning.

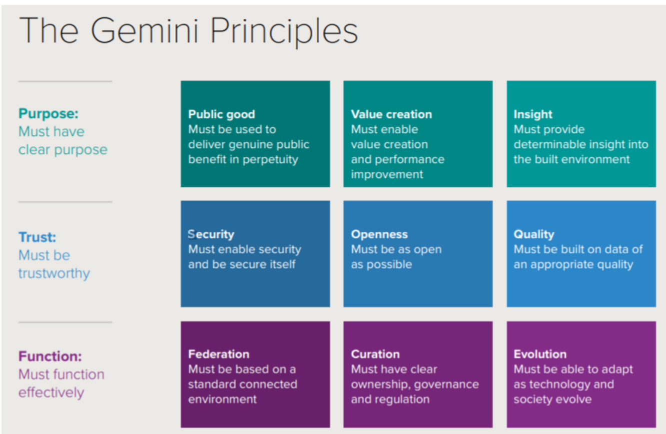
2. Literature Review

A digital twin is a digital representation of something physical. The UK National Infrastructure Commission (NIC)'s timely report "Data for Public Good" (2017) points out the great potential of utilizing the data and digital technology to support better-informed decisions and proposes to develop a national digital twin, intended as a national resource for improving the performance, quality of service and value delivered by assets, processes and systems in the built environment.

To address the policy recommendations of the NIC report, the Digital Framework Task Group (DFTG) has been established, as part of the Centre for Digital Built Britain (CDDB) based in Cambridge, to steer and guide the successful development and adoption of a digital framework for infrastructure data and the digital twin technology. The DFTG has recently published the "Gemini Principles" (2018) which sets out high-level guidelines for the national digital twin (see Figure 1).

The first principle, i.e. purpose, highlights the essential link between the digital twin apparatus and the policy outcomes that the digital twin is expected to serve – the digital twin is the *means* to achieve better-informed decisions, rather than the *end*. This principle addresses one of the outstanding issues observed in early smart city initiatives, where technology is deployed as a one-size-fits-all solution without considering the specific social and political context of the problem of interest. The purpose principle not only justifies the need for a digital twin, but also, together with the associated context-specific constraints, informs how digital twins should be designed, developed and operated in order to achieve the purpose. The second principle, namely trust, is timely proposed as it is in the wake of recent data misuse and privacy breach incidents (e.g. the Cambridge Analytica case). To build a trustworthy digital twin, a good balance between openness/transparency and security needs to be explored through a combined perspective of both governance through/of technology. The third principle, function, suggests that the digital twin, as a federated system of models, needs to function effectively in support of its purpose, and be able to adapt to future changes of technology and society.

Figure 1 The Gemini Principles for developing digital twins of the built environment (Centre for Digital Built Britain, 2018)



The data-driven digital twin requires a sound information management framework to guide the design of its data architecture. The "Gemini Principles" proposes an information value chain which illustrates the process of data being translated into intelligence for decision making (see Figure 2). In principle, as data is collected, processed, interpreted and used to support decision making, the data volume would decrease while the data value would increase.

A similar while earlier version of the diagram has been presented in an industry paper by Mott MacDonald (Wildfire, 2018) (see Figure 3).

Figure 2 The information value chain: showing the connection between data and better decisions that lead to better outcomes (Centre for Digital Built Britain, 2018)

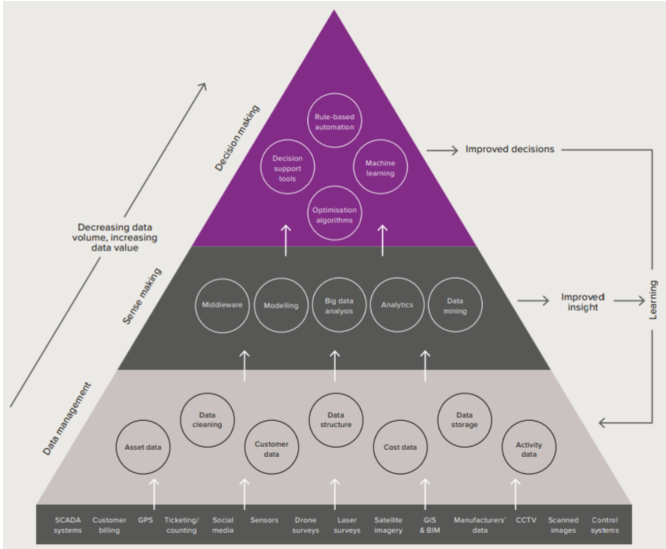
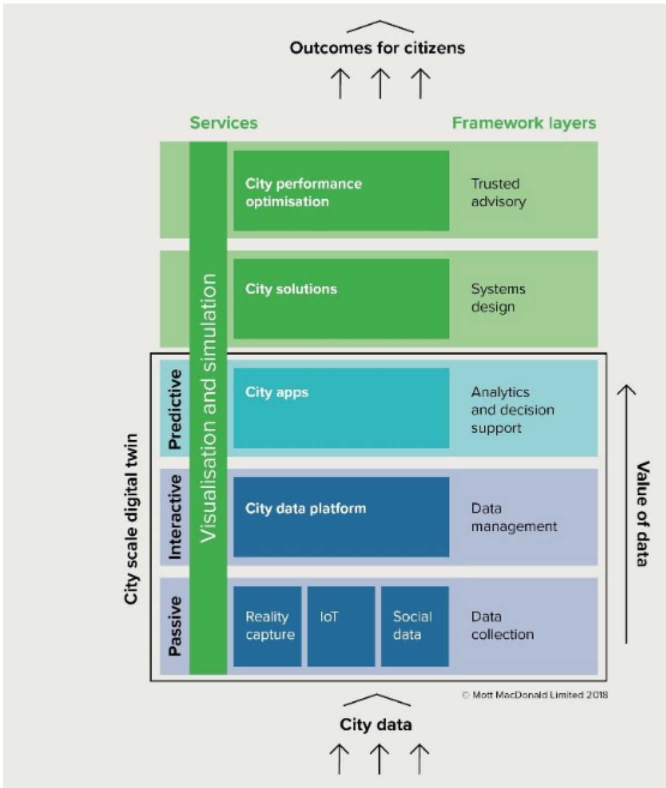


Figure 3 Smart City: concept framework (Wildfire, 2018)



It should be noted that the diagram in Wildfire (2018) is a conceptual framework for smart cities, rather than specifically for a city-level digital twin. Comparing these two diagrams yields some interesting findings. First, the Gemini Principles (as in Figure 2) explicitly includes a "learning" process (see the annotation on the right side of the pyramid), which represents a feedback loop between the upper-level decision making and the lower-level data management and processing. In practice terms, it involves a post-implementation evaluation of the "improved" insight in terms of its role in driving informed decisions, the outcome of which then informs how data should be better collected, processed and presented. By contrast, Wildfire (2018) does not include a feedback loop and implicitly assumes a linear transition from data analytics to city solutions. We argue that the feedback loop is critical for developing a city-level digital twin. Given a specific policy question, a knowledge gap may appear between what data and digital technology can offer (for engineers and analysts) and what evidence and insights are required for informed decisions (for policy makers) – it thus entails an iterative process to accumulate a body of knowledge and establish a new workflow that links data analytics to actual policy outcomes.

Second, one aspect that the Gemini Principles has not broached but Wildfire (2018) has is delineating the functional boundary of digital twin over the information flow. Wildfire includes the data collection, management, analytics and decision support within the remit of a city scale digital twin, while opts out city-level solution making and optimisation. The Gemini Principles however does not present such a functional boundary, perhaps because there is no consensus yet on this topic. In manufacturing industry where the term 'digital twin' was originally coined, the functional boundary of a digital twin explicitly reflects the functional structure of the subject thus tends to be well delimited (e.g. an engine, an aircraft or an assembly line). Human may be included as endogenous actors in some cases (e.g. workers on an assembly line) but is rarely the main subject of the twin model. However the upscaling from an engineering digital twin to a city-level digital twin is not straightforward as the functional boundary of cities as a complex, hierarchical system is not yet fully understood.

Third, there seems to be an underlying assumption that a city-level digital twin is exclusively good for cities and citizens. Learning from early smart city initiatives, technology investment may lead to negative social impacts. One notable example is the case of increasing unequal access to city infrastructure: the term "splintering urbanism" (Graham & Marvin, 2001) has been introduced to describe the exclusionary and fragmentary effects of privatizing infrastructure provision. Do we know exactly what the digital twin can do (either good or bad) and can the digital twin provide us what we need? We will continue the discussion on this point shortly in the next section.

In the academia, Batty (2018) provides an insightful discussion over some fundamental questions on the digital twin concept. He argues that the digital twin, as an exact mirror of the physical system, represents a distinct approach compared with conventional urban system models, where abstraction remains a core element to enable realistic construction and operation of models. He also argues that "a true digital twin running in real time is no different from the system itself and this poses the question as to how the digital twin can be used to learn about the system and used to explore, simulate and test new designs if it is the system itself" (Batty, 2018, p. 818). A philosophical enquiry into this question is beyond the scope of this paper, but it addresses a profound question - do we have enough knowledge of our cities to develop a digital twin to the same level of complexity as observed in reality, which is ever increasing thus unlikely to be fully understood.

3. Propositions

Based on the brief literature review, we introduce six propositions about the city-level digital twin in this section, which are interim research findings of the ongoing "Digital Cities for Change (DC^2)" project at the Centre of Smart Infrastructure and Construction, University of Cambridge.

Proposition 1: The development of a digital twin is a progressive process; the digital twin (as a technical system) and the social system that develops, operates and uses the digital twin must co-evolve with each.

Our first proposition relates to the intellectual relationship between the digital twin and the social system that builds, operates and uses it. It raises fundamental questions in terms of what it takes to build a digital twin and what a digital twin can eventually offer the society, in intellectual terms. On the one hand, our existing knowledge about cities and infrastructure would define how well we could build and evaluate the digital twin; on the other hand, thanks to artificial intelligence (AI), the digital twin could potentially produce new knowledge that complements or even surpasses ours. Therefore there is a possibility in theory that the digital twin, equipped with advanced artificial general intelligence, could tell us how to plan and manage our cities before we know how. The issues relating to data/AI ethics are not the focus of this paper, but even from a technical perspective, we should not get too excited by the alleged prospect that the digital twin could solve city and infrastructure problems "automatically". The Proposition 1 suggests that, if our understanding of cities is not comparable with the digital twin, we would not be able to design and evaluate the digital twin, let along regulate it through pre-emptive measures.

Proposition 2: the advancement of data science, particularly the machine learning techniques, will complement existing theories of cities and infrastructure and jointly contribute to the essential knowledge for developing digital twins.

Conventional modelling is usually based on established theories, such as physics, engineering, economics, geography and urban planning, which are constructed by humans. The emerging data science, particularly the machine learning technique, represents a new approach for producing knowledge that is comparable with human intelligence in some respects.

In terms of the implications for digital twins, it supports the wide claim that the digital twin, once properly built up, could execute complex control functions automatically based on real-time data feed (e.g. real-time traffic management). The application of AI can also help researchers to identify patterns from enormous amount of digital data generated from cities and infrastructure systems. However, the development of artificial general intelligence (AGI), though making great strides on some threads (e.g. playing computer strategy games and assisting radiological imaging diagnosis), is still at an early stage and remains distant from being able to solve real-world policy problems as complex as traffic congestion. It may be a more productive approach to combine the power of AI with conventional modelling for one shared purpose – establishing a body of knowledge on the interconnections and interdependences among city functions and infrastructure systems. On the one hand, the knowledge informs the design and use of the digital twin; on the other hand, the digital twin can co-create such knowledge with human agents.

Proposition 3: A city-level digital twin does not necessarily mirror the city in geometric terms; and the spatial/temporal resolution of the digital twin should be informed by the purpose it serves.

This proposition raises a practical question regarding how to make the digital twin of the built environment useful before it reaches its final development stage, i.e. an exact mirror of the physical system. It may relate to a common (mis)understanding that a digital twin of the city must look like the city geometrically – it probably explains why most of the current digital twin products are based on a high-resolution 3D models. It should be noted that the paper does not reject the idea of city-scale 3D models; in fact we acknowledge that developing an interoperable master database of the built environment through 3D geospatial modelling would be a key enabler for digital twins. However it does not necessarily mean that a city digital twin can only be functional if the whole city is digtally mirrored in geometric terms. Some urban models represent the physical systems and processes in cities using mathematical expression (Anas, Arnott, & Small, 1998; Fujita, Krugman, & Venables, 1999; Jin, Echenique, & Hargreaves, 2013) - they are also important building blocks for the city-level digital twin.

For a city-level digital twin to be useful and feasible, the data inputs and the associated level of detail and rate of change should be defined according to the purpose. Those specifications are not necessarily static, and can be adapted at a later stage in order to accommodate the likely changes of purpose and technology improvement. It suggests that not all digital twins have to aim at real-time, nor the finest spatial unit of analysis. For city and infrastructure planning, the resolution of a digital twin model should be informed by the scale/rate of change of the policy question. For example, appraisal for major infrastructure investment usually considers long-term (10-20 years if not longer) impacts across the wider region, which implies predictions per second and per square meter may be less relevant. For policy questions that involve a wide range of spatial and temporal scales (e.g. tackling traffic congestion involves both real-time traffic management and long-term strategic planning), it is important to develop an appropriate strategy at the outset to mediate the contrasting dimensions. It is likely that those questions would require a federation of digital twin models that are interoperable and interactive, rather than a single model.

Proposition 4: To help solving complex policy challenges, a realistic use of a digital twin is to identify system-level risks and inefficiencies of development options and to foster cross-disciplinary/professional collaboration, as opposed to providing a singular model-based optimization.

Policy challenges, particularly those pertaining to cities (e.g. traffic congestion and air pollution) are cross-cutting thus require a coordinated approach. The digital twin, once purposefully built, provides a useful policy simulator to explore the strength and weakness of alternative policy options. Furthermore, the digital twin may reveal system-level risks and inefficiencies which are usually caused by the lack of coordination in policy making. The digital twin can provide a consistent evidence base across disciplines to facilitate policy debates and collaboration among urban professionals. System-level optimization, though being the explicit purpose of some digital twins in the engineering sphere, may not be an effective approach to those “wicked” urban problems. One of the reasons is that the developers of the digital twin, given their technical background, often do not have enough understanding of the non-technical factors (e.g. political and value difference, the hierarchies of decision making, institutional constraints, implicit trade-offs faced by policy makers). These non-technical factors are usually highly influential but tend to be very difficult to be incorporated in the model. A model-based optimization which does not consider such non-technical constraints would have little meaning. This then leads to the next proposition.

Proposition 5: The development of the digital twin, as a technical apparatus needs to incorporate insights from the governance perspective to be able to deliver the expected policy outcomes.

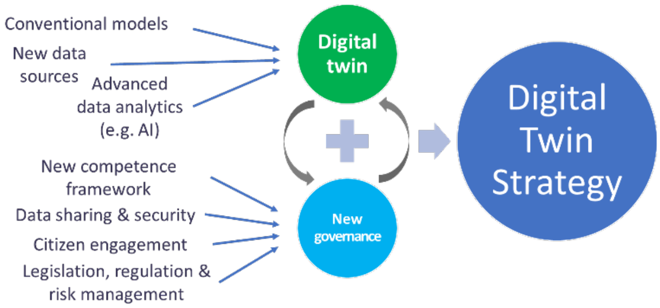
In order to make a real policy impact, the digital twin as a technological tool must be accompanied by a new governance framework to provide meaningful and constructive policy insights. Figure 4 illustrates how the digital twin and the governance framework come together to form a “digital twin strategy”. The digital twin strategy features a two-way interaction between the technology and the governance.

- The digital twin necessitates a new governance framework to be able to leverage its analytical power to inform decision making; it also acts as a technological driver of change in governance;
- The design and implementation of the digital twin should address the specific characteristics of the local governance system; pro-active policy measures

are required to regulate the use of the digital twin and mitigate the possible risks.

- The interaction between the digital twin technology and the governance framework must be designed purposefully to address specific local policy needs and deliver the aspired policy outcomes.

Figure 4 A Digital Strategy Combining Technology with Governance



The proposed digital strategy also suggests that the digital twin is not a context-free technology that can be simply purchased and installed over existing governance system. To develop a purposeful digital twin application, the digital twin must incorporate insights from the governance perspective from its design phase. The digital twin technology needs to be coupled with policy instruments to form a coordinated policy bundle in order to deliver the expected benefits.

Proposition 6: Digital twins do not necessarily provide direct solutions to city and infrastructure problems; and certain problems do not lend themselves to a digital solution either.

Although the digital twin opens up new possibilities for planning and managing our cities and infrastructure, it should not be seen as a panacea for all policy challenges. For city policy making, the city-level digital twin provides a useful simulation environment to test various policy options, which is rarely possible in real world. Nonetheless designing policies is likely to remain a task for humans in the foreseeable future, even if the artificial intelligence embedded in the digital twin may eventually be able to propose policy designs. For some urban challenges, such as social inequality and housing crisis, the digital twin tends to be less effective. This is because the digital twin, as a digital representation of the physical world, does not directly address the underlying socio-political causes. In fact, these issues may simply not lend themselves to a digital solution. The investment on digital technology and data, albeit being helpful in terms of collecting evidence, should not overshadow the efforts of investigating the inner mechanism and root causes of the issues.

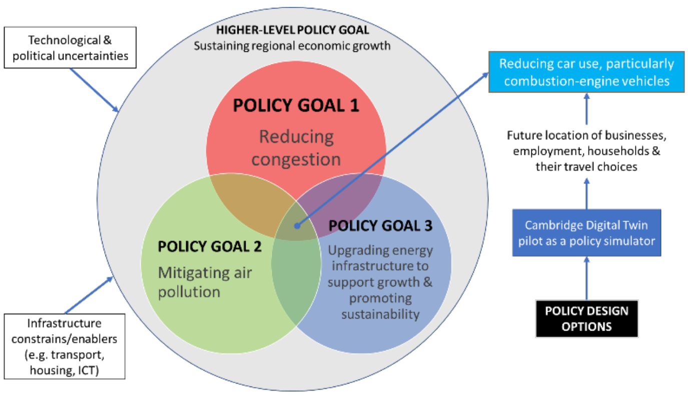
4. Case Study of the Cambridge Sub-region

To empirally test the above propositions through a case study, the research team is currently developing a digital twin pilot for the Cambridge Sub-region in collaboration with the local authorities. The Cambridge Digital Twin (CDT) pilot project is also supported by the Centre for Digital Built Britain (CDBB). In this section, the purpose and the technical approach of the CDT pilot project will be introduced. (The CDT pilot is expected to end by August 2019, which means the final research findings are not yet available as we prepare the paper.) The propositions discussed earlier underpin the design of the Project and it is expected that the Cambridge Digital Twin case study would provide empirical evidence for verifying the Propositions and generate transfereable knowledge on city-level digital twin development.

The purpose of the Cambridge Digital Twin (CDT) is to bridge the policy silos of transport, housing, environment and energy in the Cambridge sub-region. These policy themes have been jointly selected by the local authorities and the research team as cross-cutting policy issues with local urgency. It is clear that those policy themes involve varying spatial-temporal scales as well as a wide spectrum of stakeholders.

In terms of time horizon, the proposed CDT prototype will be focused on the medium-to-long term policy scenarios (potentially by 2031, in line with the time span of the local plans). In terms of modelling scope, the CDT prototype aims to quantify some of the interdependences among transport, air quality, housing and energy infrastructure in the Cambridge sub-region in relation to some background changes and uncertainties (see Figure 5). It has been identified through workshop and interviews that the future location of businesses, employment, households and their travel choices is the pivot variable set that connects the selected policy themes. In terms of travel behaviour, the CDT pilot will be mainly focused on commuting trips (e.g. journeys to work) which is a key determinant of peak-time travel demand.

Figure 5 Policy interdependences among transport, air quality and energy infrastructure in the Cambridge sub-region



The technical approach for developing the Cambridge Digital Twin pilot involves three inter-connected tasks.

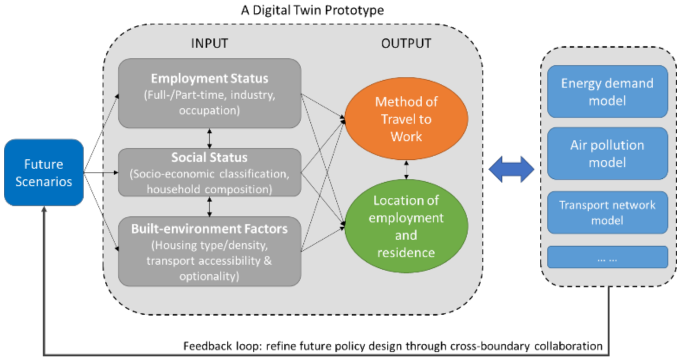
TASK 1: Understanding the past trends. Investigating how people travel to work in relation to their residence location, workplace, housing tenure, socio-economic classification, nature of employment in terms of full-/part-time and industry, and public transport options. The relationship between explanatory variables and the dependent variable (how people travel to work) will be identified through statistical analysis and then represented through urban system modelling. The model developed will be able to articulate the observed pattern of journeys to work in the study region.

TASK 2: Scenario design. A series of the future scenarios will be developed and tested to highlight the impacts of infrastructure constraints and uncertainties on future location of businesses, employment, households and their travel choices. Given the exploratory phase of the CDT development, the scenarios are expected to be highly contrasting, rather than minor tweaks of the status quo. To test the CDT, it is expected that the policy makers and analysts could design their own scenarios and run the CDT by themselves. To facilitate the co-development, the CDT will employ a web-based platform with a back-end model. A visualized web interface is being designed to take user-specified scenario inputs. One possible scenario example regarding the increase of teleworking is introduced below.

- "Prevalence of teleworking" scenario: assuming that a significant amount of jobs (defined by industry and occupation level) can accommodate workers working remotely at community-based co-working centres. Key scenario inputs: the quantity and distribution of remotely-accessible jobs, and the location of the co-working centres.

TASK 3: Linking the CDT with other models or existing workflow in local authorities. This step is critical as it tests to what extent the CDT may foster cross-boundary collaboration in policy making. One of the current policy conundrums revealed from the workshop and interviews is that, to tackle air pollution, the most effective measure would be removing the pollutant from the source – either reducing the use of combustion-engine vehicles or replacing them with electrical vehicles (EVs); the former measure requires a joint effect with the transport team, combining price mechanisms and alternative mobility solutions to reduce car use; the latter entails the coordination with the energy team, who may experience difficulty of financing the upgrade of local energy infrastructure. In this case, the CDT application will be connected to existing local models on transport, air quality and energy, and together serve as a shared evidence and knowledge base, upon which a coordinated policy bundle can be explored. The key model input and output, and the proposed connections with existing models are summarized in Figure 6.

Figure 6 Technical approach for developing the Cambridge Digital Twin prototype



5. Conclusion

The digital twin technology represents one of the latest technical trends for smartening our cities. However the roadmap to develop a city-/national-level digital twin is not yet clear. This paper reviews some of the latest progresses on digital twin development in the UK, including the recently published report "The Gemini Principles" by the Centre for Digital Built Britain, which provides a high-level guidance for digital twin development. Inspired by the "Gemini Principles", the paper puts forwards some propositions on the city-level digital twin agenda, which stem from an ongoing research project on smart cities at the Centre for Smart Infrastructure and Construction. The paper also introduces the research plan for developing a digital twin prototype for the Cambridge sub-region, which serves as a case study to test the proposed propositions.

In the UK several large-scale digital twin projects are taking place across various sectors and locations, with more in the pipline. It is expected that more and more empirical evidence coud be gathered to established a shared body of knowledge for leveraging and regulating the power of the digital twin technology. As discussed in the first proposition, the development of a digital twin is a progressive process, where the digital twin (as a technical system) and the society who develops and operates the digital twin must co-evolve together. The concept of digital twin may be a topical idea whose time will pass, but the knowledge developed behind it and by it will bring us ever forward.

6. Acknowledgements

This article is developed based on the work and discussions of the "Digital Cities for Change" project at the Centre for Smart Infrastructure and Construction, Department of Engineering, University of Cambridge. The Project is funded by The Ove Arup Foundation and receives additional financial support from the Centre for Digital Built Britain. The authors would like to thank the the project's Steering Group as well as the officials from the local councils in Cambridge for their insightful comments and feedback.

7. References

Anas A, Arnott R, & Small KA (1998) Urban spatial structure. *Journal of Economic Literature*, *36*(3), 1426–1464.

Batty M (2018) Digital twins. *Environment and Planning B: Urban Analytics and City Science*, *45*(5), 817–820. https://doi.org/10.1177/2399808318796416

Centre for Digital Built Britain (2018) *The Gemini Principles*. https://www.cdbb.cam.ac.uk/Resources/ResoucePublications/TheGeminiPrinciples.pdf (accessed 26/04/2019)

Fujita M, Krugman P & Venables AJ (1999) The Spatial Economy. *Economic Geography*, Vol. 24, pp. 245–261.

Graham S & Marvin S (2001) *Splintering urbanism: Networked infrastructures, technological mobilities and the urban condition*. Routledge.

Grieves M (2014) *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*. http://innovate.fit.edu/plm/documents/doc_mgr/912/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf (accessed 26/04/2019)

HM Government (2017) *Industrial Strategy: building a Britain fit for the future*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf (accessed 26/04/2019)

Jin Y, Echenique M & Hargreaves A (2013) A recursive spatial equilibrium model for planning large-scale urban change. *Environment and Planning B: Planning and Design*, *40*(6), 1027–1050. https://doi.org/10.1068/b39134

National Infrastructure Commission (2017) *Data for the public good*. https://www.nic.org.uk/publications/data-public-good/ (accessed 26/04/2019)

Wildfire C (2018) How can we spearhead city-scale digital twins? www.infrastructure-intelligence.com/article/may-2018/how-can-we-spearhead-city-scale-digital-twins (accessed 26/04/2019)