ECAI 2020
G.D. Giacomo et al. (Eds.)

doi:10.3233/FAIA200435

Towards the Development of Societal Twins

Dan Birks[1,4] and **Alison Heppenstall**[2,4] and **Nick Malleson**[3,4]

Abstract.[1] A digital twin is a virtual data-driven replica of a real-world system. Recently, digital twins have become popular in engineering and infrastructure planning where they provide insights into complex physical systems or processes. Yet, to date, considerably less research has explored how digital replicas of social systems - representing the decisions, behaviors and interactions of individual people, and, in turn, their emergent outcomes - might be developed and integrated with those of physical systems. In this position paper we discuss the need for such societal twins, what they might look like, and set out key challenges that will need to be overcome if these benefits are to be realised.

1 INTRODUCTION

In recent years a number of engineering and infrastructure planning applications have been transformed by the development of 'digital twins' [1] – virtual replicas of physical systems or processes that incorporate rich and diverse data. The rationale for constructing these models is simple – real world systems are complex, difficult to understand and expensive to manipulate and iterate – thus considerable benefit can be realised by using digital twins to identify possible problems, test potential solutions and increase system efficiencies and effectiveness. Yet our world is an imbricated backcloth of interdependent physical and social systems, and while the development of digital twins has generated significant advances in the study of physical systems, considerably less research has explored how societal twins – simulating the decisions, behaviours and interactions of individuals that make up our social systems – might be developed, validated and used to support real world decision making.

At the same time there is an increasing acceptance amongst social science scholars that understanding and developing solutions to wicked problems [2] such as inequality, poverty, crime and economic underdevelopment requires a holistic perspective that acknowledges the complex interrelationship of the physical and social typically beyond both the purview of historically disparate academic departments and traditional analytical approaches. For example, how does the cost of public transport impact on the ability of different socio-economic groups to access public services within their locality?

The recent proliferation of 'big data' provides the opportunity for digital twins of social systems to be constructed that have the potential to answer these questions. These data – that include real-time footfall counters, environmental sensors, traffic detectors, alongside a range of social media streams – remain variable in quality and individual-level precision, but offer new and richer insights into the behaviors of previously hidden groups as well as providing an extra level of validation for groups whose behavior is already well understood. Embedding these new forms of data within digital twins will present a number of challenges but also provide unique opportunities to focus on understanding the bottom-up mechanisms that drive complex systems such as cities, rather than concentrating purely on aggregate outcomes and thus, 'smoothing over' important lower-level dynamics [3]. In this position paper we discuss what a societal twin might look like, highlight some potential applications in supporting evidence-based decision making, and set out key challenges that will need to be overcome if these benefits are to be realised.

2 TOWARDS SOCIETAL TWINS

In the field of urban analytics there is growing discussion of digital twins of cities [4] – see for example projects in Rotterdam [5], Singapore [6], and Rio de Janeiro [7]. These systems provide access to real-time information about the city – such as the volume of traffic on road networks, local temperature and air quality information, public transport utilisation, etc. – and are ultimately moving towards the provision of entire digital realisations of physical urban infrastructure. Where many systems are lacking, however, is in their representations of the individual *people* whose actions ultimately drive the evolution of the cities they inhabit. Put simply, if we have digital twins of physical infrastructure – *cities* – we need to populate them with realistic synthetic populations of *people*, i.e. societal twins.

For example, if changes are made to the transport network – planned or otherwise (i.e. in response to an emerging problem) – how might such alterations affect commuter behaviors and subsequently how the spatial and temporal distribution of potential victims of crime, which in turn should inform the allocation of crime reduction resources. In effect, the physical systems that cities are composed of are not only intricately linked with each other, but are also linked to the dynamic behaviors of the people who utilise them. Without an adequate representation of these synergies, we cannot hope to build accurate and ultimately useful models of cities. A societal twin, coupled to existing physical digital twins, could benefit a wide range of societally important areas. Here we propose agent-based models (ABM) supported by new advances in techniques for model calibration and validation, provide a natural framework for the creation of a societal twin.

3 AGENT-BASED SOCIETAL TWINS?

While there are likely multiple ways that synthetic societies can be constructed to populate our digital twins of cities, ABM offer

[1] School of Law, University of Leeds, UK, email: d.birks@leeds.ac.uk
[2] School of Geography, University of Leeds, UK, email: a.j.heppenstall@leeds.ac.uk
[3] School of Geography, University of Leeds, UK, email: n.s.malleson@leeds.ac.uk
[4] The Alan Turing Institute, British Library, London, UK

considerable promise. Agents offer a direct analogy to real world populations, allowing researchers to model the individual actions and interactions of heterogenous, autonomous decision makers (both amongst themselves and with their environment), and the aggregate societal outcomes that emerge as a result. In constructing these models, agent behaviors can be informed by both social science theory and the best quantitative and qualitative evidence we have about human decision making. Taking this approach offers a white-box method for understanding historical societal system states and, where appropriate calibration and validation has taken place, may offer means to forecast future system states – providing a computational laboratory in which interventions (both social and physical) can be prototyped free of logistical and ethical constraints. It is under this rationale that ABM has already been applied successfully in a wide range of relevant domains, from modelling the short-term behaviors of crowds of people [8], the impacts of localised street network configurations on crime risk [9], and longer-term adoption to new technology [10]. However, with a handful of notable exceptions [e.g. 11], the vast majority of models are either applied to hypothetical situations or, where more closely constrained by real world scenarios, operate on tightly constrained problems. None have attempted to model the *general* behavior of a population and its interactions with physical infrastructure over a spatio-temporal scale that would be appropriate for a societal twin. Key questions then become: what do we need to make an agent-based societal digital twin a reality, and in what ways would it differ from those previous applications of ABM in the social sciences?

4 KEY CHALLENGES

Uncertainty: Social systems are complex and adaptive. Consequently, social simulations typically exhibit orders of magnitude greater degrees of freedom than traditional analytical approaches. If societal twins are to be used to support real world decision making, we need a much better understanding of the uncertainties associated with underlying models. For example, is poor model performance a result of an inadequate agent behavioral representation or otherwise poor model conceptualisation (in which case the model will never perform well) or do the problems arise because the system is inherently uncertain and even a perfect model will diverge from reality (in which case a 'real time' model, discussed shortly, might be necessary).

Social System Data: Societal twins require access to appropriate data to ensure they reflect real world social processes. While increasingly large quantities of data describing fundamental properties of populations are collected – do the same data exist for social processes that exist within our cities? If so, at what resolution? And for whom are they representative? Moreover, how best can they be incorporated into our models? While we have access to diverse data, bringing disparate sources together and analyzing them to identify social processes remains a key challenge. Nevertheless, recent work has seen ABM linked to other methods including statistical and machine learning for the purposes of model behaviour calibration, optimisation and evaluation.

Real-Time Modelling: Whilst many of the questions that we will ask of our digital twins will have relatively long time scales (e.g. 'how will the development of city *y* vary over the next decade in response to *x*'), there will be occasions where much more up to date insight could be valuable. Could a digital twin, backed by an agent-based societal model, estimate the behaviour of the underlying system in *real-time?* This is analogous to meteorology, where regional and global weather models run in real-time and are constantly updated with the latest available data from environmental sensors. This real-time updating results in much more accurate short-term predictions. Although work has begun to explore these opportunities [e.g. 12] they are a long way from being able to simulate real human systems.

Ethics of Societal Twins: The development and use of societal twins also present a range of significant ethical challenges. How do we ensure that data used to construct the twin are sufficiently unbiased and representative of real-world societal processes? Is it appropriate to support the provision of resources to select communities using the model while acknowledging uncertainty? How do we manage seemingly competing goals of utilising rich individual level data while ensuring appropriate privacy? These and other questions will need to be answered co-operatively with experts in diverse fields and through active engagement with the communities these models seek to support.

ACKNOWLEDGEMENTS

AH is funded by an ESRC-Alan Turing Fellowship (ES/R007918/1) and NM through the European Research Council Horizon 2020 research and innovation programme (grant agreement No. 757455).

REFERENCES

[1] F. Tao and Q. Qi, *Make more digital twins*. Nature.com, (2019).
[2] B.W.Head, *Wicked problems in public policy*. Public policy, 3(2), 101, (2008)
[3] M. Batty, *Agents, cells, and cities: new representational models for simulating multiscale urban dynamics*. Environment and Planning A, 37:1373–1394. (2005)
[4] M. Batty, *Digital twins*, Environment and Planning B: Urban Analytics and City Science, vol. 45, pp. 817--820, (2018).
[5] European Innovation Partnership on Smart Cities and Communities, *Rotterdam's Digital Twin Redefines Our Physical, Digital, & Social Worlds.* www.eu-smartcities.eu/news/rotterdams-digital-twin-redefines-our-physical-digital-social-worlds (2019)
[6] National Research Foundation, Prime Minister's Office Singapore, *Virtual Singapore,* www.nrf.gov.sg/programmes/virtual-singapore
[7] R. Kitchin, T. Lauriault and G. McArdle, G. *Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards*. Regional Studies, Regional Science 2, 1, 6–28, (2015)
[8] D. Helbing, P. Molnár, I.J. Farkas and K. Bolay. *Self-Organizing Pedestrian Movement*. Environment and Planning B: Planning and Design 28, 3, 361–83 (2001)
[9] D.J. Birks and T.P. Davies, *Street Network Structure And Crime Risk: An Agent-Based Investigation Of The Encounter And Enclosure Hypotheses*." Criminology 55, 4, 900-937 (2017).
[10] H. Zhang, Y. Vorobeychik, J. Letchford, and K. Lakkaraju. *Data-Driven Agent-Based Modeling, with Application to Rooftop Solar Adoption*., 513–521. Proceedings AAMAS '15. (2015).
[11] J. Parker, and J.M. Epstein. *A distributed platform for global-scale agent-based models of disease transmission.* ACM Transactions on Modeling and Computer Simulation (TOMACS), 22(1), 2, (2011)
[12] J.A. Ward, A.J. Evans, and N.S. Malleson. *Dynamic Calibration of Agent-Based Models Using Data Assimilation*. Royal Society Open Science 3, no. 4 (2016). https://doi.org/10.1098/rsos.150703.