The Alan Turing Institute

Towards ecosystems of connected digital twins to address global challenges

Successes and recommendations from The Alan Turing Institute's AI for science and government programme

The Alan Turing Institute

About The Alan Turing Institute

The Alan Turing Institute is the national institute for data science and artificial intelligence (AI). Established in 2015, we are named in honour of Alan Turing, whose pioneering work in theoretical and applied mathematics, engineering and computing laid the foundations for the modern-day fields of data science and AI. Headquartered at the British Library in London, we partner with organisations across government, industry, academia and the third sector to undertake world-class research and innovation that benefits society.

About the ASG programme

Established in 2018, the AI for science and government (ASG) programme is a five-year, £38.8 million research programme at The Alan Turing Institute that aims to deploy data science and AI to address significant societal challenges. Funded through UK Research and Innovation's (UKRI's) Strategic Priorities Fund, and delivered in partnership with the Engineering and Physical Sciences Research Council (EPSRC), the programme brings together researchers from diverse disciplines to tackle problems including climate change, health emergencies, economic instability, and online harms.

Contents

Authors

The AI for science and government's 'Ecosystems of digital twins' working group (Hayley Bennett, Mark Birkin, Jennifer Ding, Andrew Duncan, Zeynep Engin)

DOI: 10.5281/zenodo.7840266

Executive summary

A digital twin is a virtual representation of a physical entity. By bringing together data, modelling and simulation, digital twins enable insights about the performance of structures, products and processes. They are powerful tools for understanding and optimising complex systems that are already beginning to demonstrate their potential across multiple sectors, from engineering and manufacturing to health and the environment.

In this white paper, we share learnings from The Alan Turing Institute's AI for science and government (ASG) programme, which has been conducting foundational research and applied work in digital twins under the 'Ecosystems of digital twins' (EDT) theme. ASG researchers are expanding the scale and complexity, and deepening the application and deployment of digital twins.

We organise digital twins into three groups according to their level of complexity, scaling up from a digital twin of a single asset or system, to a digital twin incorporating multiple assets or systems. At the third level, digital twins of differing levels of complexity are linked together in an 'ecosystem of digital twins'. We provide case studies demonstrating the EDT theme's multidisciplinary work on digital twins for each of these three levels.

Through digital twins, we can develop unprecedented capabilities for understanding, manipulating, and managing complex systems. Within the EDT theme, we have found that for the full promise of the technology to be realised, we must consider the development and deployment of digital twins at the ecosystem-level. By connecting together digital twins within a virtual environment, we can enable sharing of information between 'federated' assets – assets under different jurisdictions and restrictions – unlocking more advanced insights and capabilities than are possible at the lower levels. Within an ecosystem, connected digital twins can also promote more varied applications, for example, by combining representations of, and insights from, both physical and social systems to support decision-making.

There are, however, certain technical, social and institutional issues that need to be negotiated in developing ecosystems of digital twins. We therefore provide the following three recommendations in support of the progress which needs to be made:

1. Elevate cross-disciplinary activities and spaces for digital twins.

2. Invest in open infrastructure, with a focus on data and technical standards.

3. Prioritise tools for building trust in and understanding of digital twins.

By following these recommendations, with support from the right investments, we can build robust ecosystems of digital twins capable of helping us tackle some of the most important global challenges, from pandemics to adapting to climate change.

1. Introduction

"A digital twin is a digital representation of physical reality, coupled together through observations and simulations. It opens up new fundamental research questions, as well as new sociotechnical and socioeconomic possibilities."

Mark Girolami, Chief Scientist, The Alan Turing Institute

The concept of a digital twin originated in the engineering fields and is now well-established in academic research, business, government and industry as a digital representation of a real-world entity. This real-world entity might be as simple as a bridge or vehicle, or as complex as a farm or health system. Digital twins apply physical, mathematical and statistical models – often supported by sensor networks capable of supplying dense, fast-moving datasets – to gain insights into their real-world counterparts. Through virtual exploration of the impacts of complex decision-making, digital twins can effectively de-risk expensive decisions before they are implemented in the real-world. The technology is already improving data-driven decision-making across multiple sectors, leading to reduced risk and enhanced performance in the optimisation of engineering designs, business processes and government projects, as well as increased resilience and adaptability in rapidly changing situations, such as pandemics, supply chain shocks and financial crises. We are headed rapidly towards a future in which the world operates on insights derived from digital twins.

In 2017, the National Infrastructure Commission set out a roadmap towards a national digital twin; a digital representation of the UK's entire national infrastructure that would allow monitoring and simulation of assets such as transport networks and energy systems to enable better planning, prediction and understanding. The Alan Turing Institute was tasked with developing the foundations for the national digital twin and, in response, established the Ecosystems of digital twins (EDT) theme within its AI for science and government (ASG) programme. Under the EDT theme, the national digital twin is envisaged not as a single digital twin but as an integrated network of digital twins connecting key components of the UK's infrastructure through securely shared data: an 'ecosystem of digital twins'.

Meanwhile, digital twins have been gaining traction at national and international levels through other high-profile projects and programmes. The UK's Digital Twin Hub coordinates multi-sector partnerships through a base at the Connected Places Catapult, with the Turing playing a key role in the hub's governance. In Europe, the Gaia-X project aims to create a network of hubs focused on developing digital infrastructure. US initiatives include the Digital Twin Consortium and a National Academies of Sciences, Engineering, and Medicine-appointed committee on digital twins. The global digital twin market surpassed $8 billion (USD) in 2022 and is expected to reach $90 billion by 2032. This rapid pace of progress demands focused research investment now to maximise the potential of the technology and minimise harms.

The ambition of work under the EDT theme is to address policy design and decision-making questions for some of the most difficult challenges facing business, government and society. Since the start of ASG in 2018, our researchers have advanced the basic science, application and adoption of digital twins in domains from data-centric engineering and environmental science to agriculture and urban planning, focusing on building wider understanding and demonstrating the potential of digital twins through collaborative projects involving government and industry stakeholders. In this white paper, aimed at industry professionals, government departments and researchers with an interest in digital twins, we present an overview of the innovative and collaborative research, and its impacts, emerging from the EDT community.

Within the EDT community, we recognise that as digital twin technology continues to mature, its range of applications is expanding in both scale and complexity. Therefore, this white paper covers the development of new methods for the exploitation of digital twins at three levels. The system level (Section 2.1) focuses on digital twins for single assets, such as ships or wind turbines. The system-of-assets level (Section 2.2) focuses on digital twins for collections of assets that are similar in nature, such as truck fleets, or dissimilar, such as the components of a working farm. Most importantly, however, the focus is advanced to the third class, namely ecosystems of digital twins (Section 2.3), which combines and connects systems of different types and levels of complexity. Projects which advance the state-of-the-art at all three classes of digital twins are described, but it is to the emergence of this new class of ecosystems of digital twins that research under the EDT theme has been most strongly and distinctively addressed. The work of EDT researchers has generated new insights in pandemic monitoring, decarbonisation and the health impacts of climate change that can inform decision-making at all levels.

Based on the learnings from our work, we also provide three recommendations to guide industry stakeholders and policy makers globally in better harnessing the potential of digital twins to address challenges across business, government, and society (Section 3).

2. Scaling up the complexity of digital twins

In this section, we provide a unifying perspective on the outcomes and learnings from efforts under the EDT theme at the Turing. We illustrate the wide-ranging work of ASG researchers through projects of increasing complexity, from digital twins of a system and system of assets towards ecosystems of digital twins. Through this body of digital twins research, we explore the new considerations arising in pursuit of applications for more complex digital twins, particularly real-world applications, such as optimising production processes on a farm or developing data-informed policies to mitigate the impact of heat waves on local communities.

In putting forward this perspective, we recognise the need to first define what we mean by 'digital twin'.

This is challenging as the term is applied in different ways across different sectors. However, researchers and policy makers have narrowed in on certain shared characteristics and capabilities. Whilst the focus here is less on contributing to a digital twin definition – and rather on describing some of these shared characteristics and capabilities across domains – we adopt a broad definition that aligns with the EDT's cross-sector approach to digital twins and incorporates the following key components:

- A real-world asset or process, system of assets, or ecosystem.
- A digital representation of the former, through a mathematical, physical, or data-driven model.
- A platform for running simulations of the model.
- A data pipeline from the physical system to the digital twin.
- Tools for visualisation, analysis and verification of model outputs that inform decision-making and interventions in the real world.

With this understanding of what we mean by 'digital twin', in the following sections we ground this high-level concept in examples from ASG researchers' projects under the EDT theme, scaling up in complexity from a digital twin of a single asset (or process), to a digital twin of a system of assets, and finally to ecosystems of digital twins.

2.1 Digital twins of single assets

At the most basic level, a digital twin can represent a single structure, asset, system or process (Figure 1). In the engineering and manufacturing industries, digital twins of infrastructure assets such as ships, aeroplanes, bridges and their constituent components represent the cutting edge of innovation. Spurred on by interest and investment from industry, developments in digital twins technologies are contributing to de-risking large projects, optimising existing systems and distilling new insights that can help to ensure the safety and improve the lifetime efficiency of assets. For example, a Turing-built digital twin for a railway bridge in Staffordshire, UK, is enabling the rail company to monitor the structural health of its bridge in real-time, whilst some of the same techniques were employed in tests of the world's first 3D printed steel bridge prior to its installation in the city of Amsterdam.

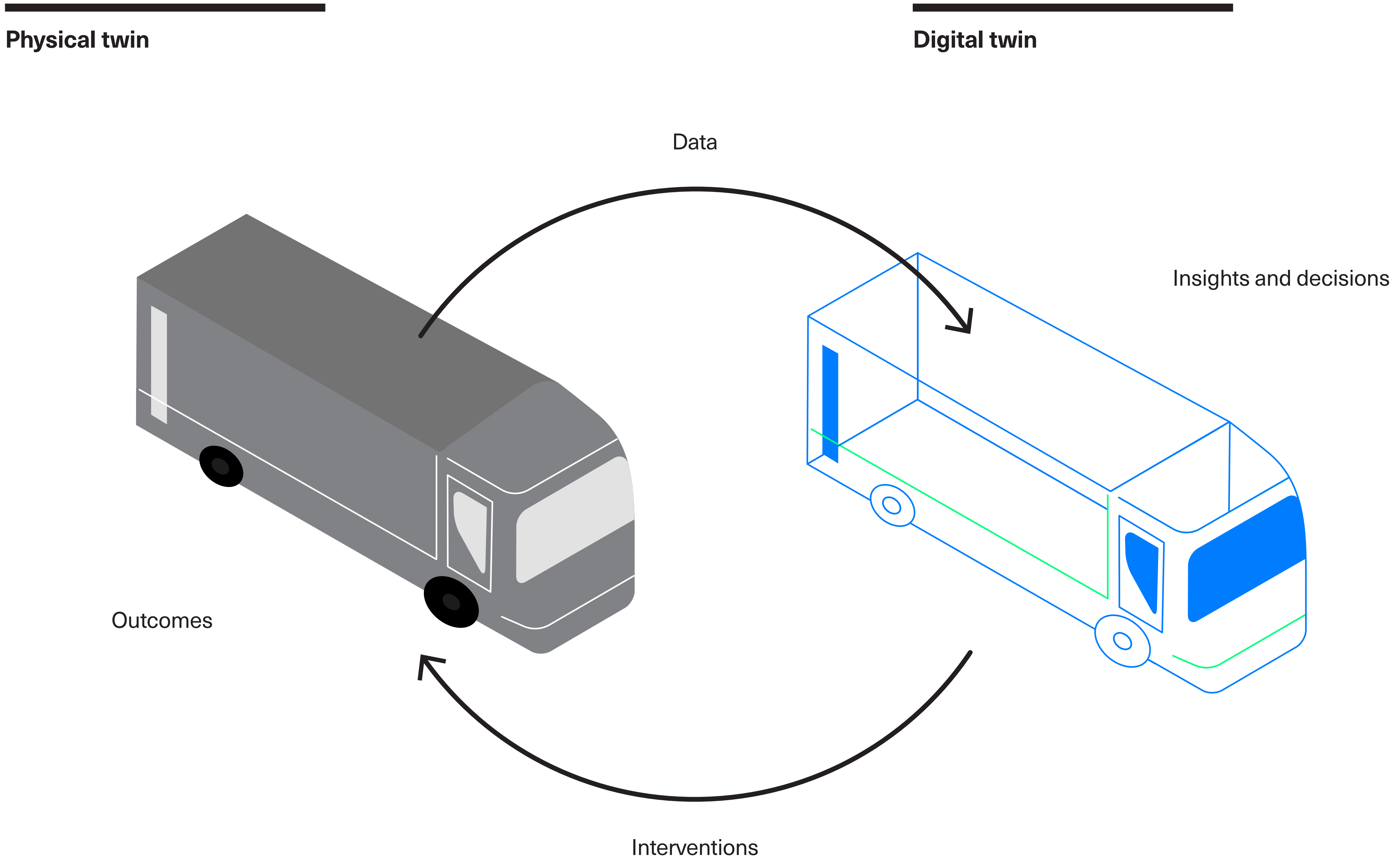


Figure 1. A digital twin for a single asset. In the example shown, data flows from the physical truck to its digital twin. Insights from the digital twin inform practical interventions in the life cycle of the physical twin, such as carrying out maintenance, and lead to improved performance.

This section introduces advancements in digital twin technology for infrastructure assets, through case studies focused on aerospace manufacturing and monitoring of maritime vessels. Both of these projects illustrate the value of cross-sector working and collaborations between academic researchers and industry for advancing the applications of digital twins. In the first case, one of the most important contributions of the ASG researchers involved is open-source software; utilising open tools and models allows the software to be adapted for diverse use cases across a broad spectrum of different sectors. 'Explainability' or interpretability of the models for users outside of the academic space is also key to ensuring their wider use.

In the second case, a digital twin designed to enable remote monitoring of a ship's electronic systems demonstrates a safety-critical application. Here, ASG researchers worked closely with partners in the maritime industry to develop processes for assuring the digital twin for reliable use. This project highlights the value of artificially-generated data, or 'synthetic data', for software testing of digital twins prior to their implementation. The synthetic data provides a stand-in for real-world data – in this case simulating data from the vessel's electronic network – to test the digital twin's response. Synthetic data generation is also relevant to digital twins at the ecosystem level (see Section 2.3).

Case study: Streamlining jet engine design and manufacture

Effective Quadratures (equadratures) is an open-source software supporting the rapid development of digital twins. Developed and maintained by ASG researchers, it can be used to streamline the often complex and computationally expensive models used in digital twins by identifying the key variables controlling their outputs. The software has been adopted for diverse use cases. For example, United States Geological Survey researchers used equadratures to simplify their models of coastal erosion and it is now being adapted to improve the estimation of ozone measurements in a joint project involving the Turing and British Antarctic Survey researchers.

In a collaboration with industry partners at Rolls-Royce, the equadratures team applied its software to streamline the design and optimisation of components of the company's jet engines. Existing computer models are capable of simulating the effects of component design on jet engine performance – for instance, how differences in the shape and tilt of the fan blades in the engine's fan affect its efficiency. However, the large number of design variables for components like fan blades means these models can take days or even weeks to run, slowing down the search for new engine designs. Using equadratures, Rolls-Royce engineers were able to rapidly home in on the most important variables, revealing a more basic underlying structure for their models that makes the search for new designs more manageable. In this way, these simplified models can help to speed up the development of more efficient jet engines.

The ability of equadratures to produce fast, accurate simulations without sacrificing interpretability is of particular importance to industry stakeholders, who must be able to justify and explain why a particular design for an engineering structure or system may be advantageous. It may therefore be preferable to, for example, black-box approaches based on deep-learning, which typically offer little insight into how they make their predictions.

Due to its open-source nature, an international community of users and contributors is developing around equadratures, leading to continual improvement and introduction of new features, as well as the potential for new application areas. From modest beginnings, equadratures is evolving into a de-facto starting point for prototyping, testing and auditing digital twins in engineering systems.

Case study: Remote monitoring for maritime safety

Modern maritime vessels typically rely on many electronic devices, which must all be monitored for errors, faults and safety incidents. In this setting, digital twins can play an important role in remote monitoring of electronic systems, offering opportunities to detect and solve issues as they occur, and make improvements to safety and efficiency. Maritime technology company FurunoHellas developed its HermAce technology for just this purpose. HermAce is a digital twin that pairs with an onboard voyage data recorder (VDR), providing remote access to streaming data from a vessel's electronic systems. To ensure the reliability and accuracy of this new digital twin technology prior to its use within the maritime industry, ASG researchers partnered with maritime services organisation Lloyd's Register. Together, they worked to

independently verify and certify the HermAce VDR digital twin in line with the organisation's digital compliance framework for 'Digital Health Management Systems' – digital technologies that are used to gather data and generate insights on an asset's health.

In carrying out the verification, ASG researchers developed a software testing procedure that leveraged novel methods in synthetic data generation to generate vast numbers of test cases. These test cases simulated normal and abnormal operating conditions, and faults, allowing the digital twin to demonstrate its capabilities for monitoring the VDR and providing the correct information under many different scenarios. The software testing process, combined with a number of tests of the VDR hardware that sits on the ship itself, led in 2022 to the digital twin's certification by Lloyd's Register.

The cross-disciplinary partnership benefited from the data-centric engineering expertise of Turing researchers and maritime expertise of Lloyd's Register. Close collaboration was required in order for the researchers to translate industry requirements for assurance of the digital twin into a testing protocol and then communicate the results.

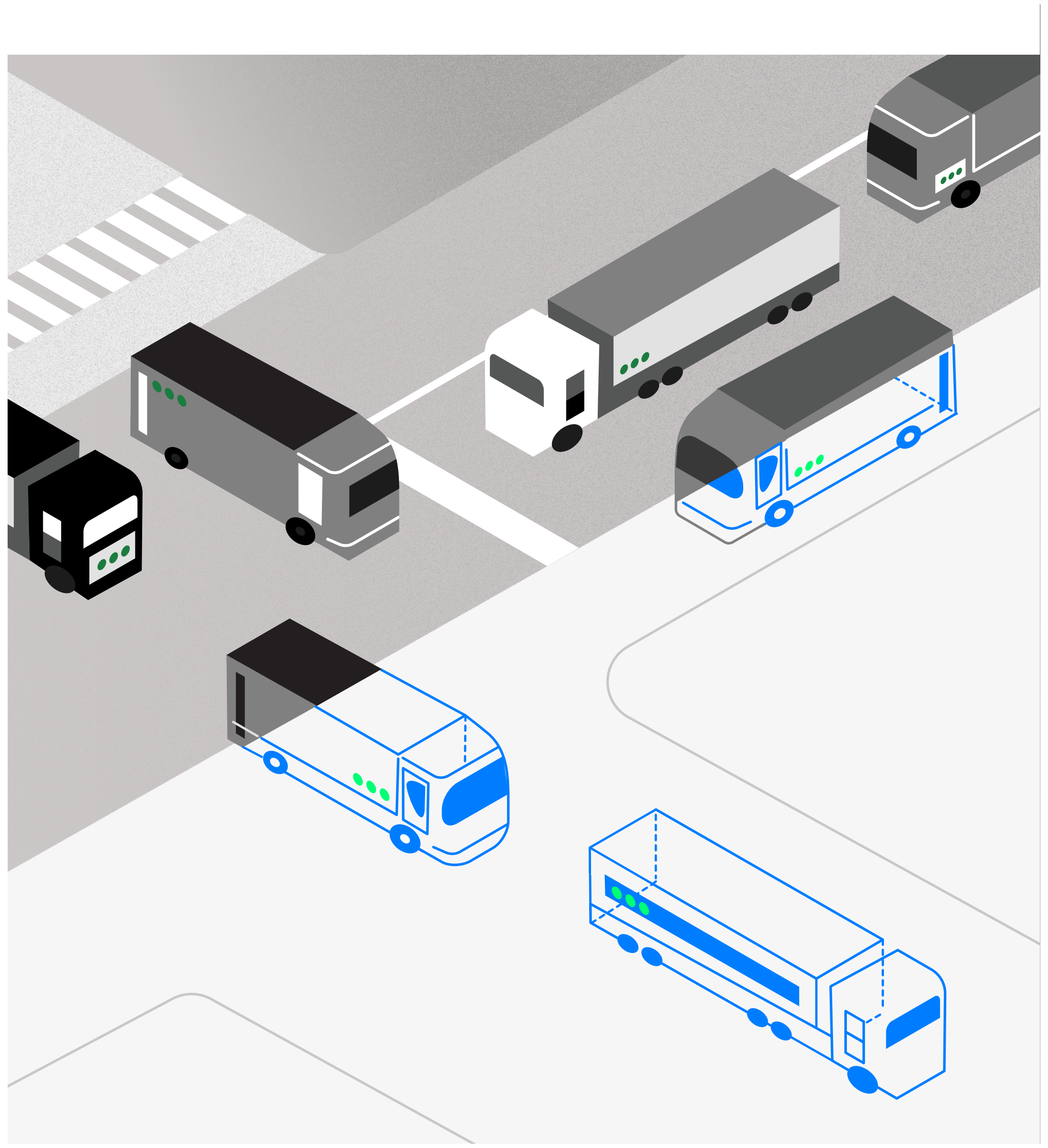
Through this collaboration, the researchers identified the need for more effective testing strategies for digital twins and are now developing new testing protocols that incorporate both software and more expensive hardware-based tests, in a semi-automated fashion, to balance reliability versus testing requirements. Such testing strategies could be of particular relevance for building trust in autonomous systems and could be tailored for specific applications, such as self-driving cars.

2.2 Digital twins of systems of assets

The system-of-systems approach in engineering provides a framework for scaling up the complexity of digital twins. Thus, beyond digital twins of single assets and processes, the concept of a single-entity digital twin can be expanded to a system of assets (Figure 2). Whilst the individual assets within the system may be similar – as in a fleet of vehicles – or varied in nature, they will typically exist under common ownership, with full knowledge of the inter-relationships between the assets.

A digital twin at this next level of complexity is more than simply the sum of its parts. Researchers working under the EDT theme recognise the value of a system-of-assets level digital twin for extracting insights not just about the individual assets themselves but about how they relate to each other,

a) Digital twin of a truck fleet



b) Digital twin of an entire transportation network

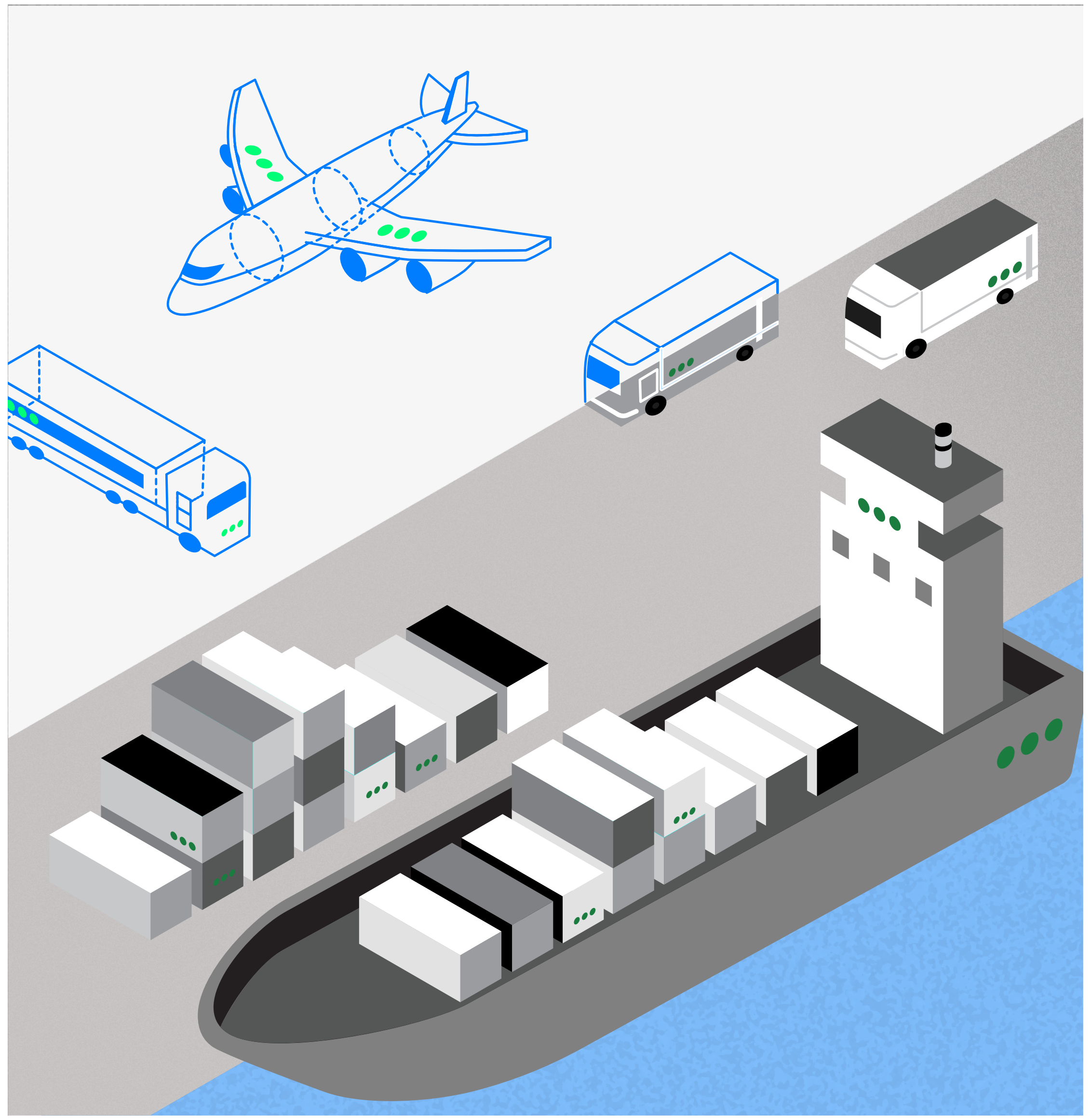


Figure 2. A digital twin of a system of assets may combine assets under common ownership that are a) similar in nature, as in a truck fleet, or b) varied in nature, as in a transportation network incorporating trucks, ships and aircraft. Although such a system of assets may be regarded as a single digital twin for understanding the system-level behaviour, it may comprise within it multiple asset-level digital twins linked together.

as well as about emergent properties of the entire system. In close partnership with industry and government stakeholders, ASG researchers are working across sectors to explore how digital twins representing multiple assets can be applied to tackle diverse, real-world challenges. This section introduces advancements in digital twins of systems of assets through projects focused on urban agriculture, commercial vehicles and the wind energy sector. These examples deepen the real-world applications of digital twins in specific sectors and lay the groundwork for more generalisable platforms for digital twins.

By adopting a systems perspective of multiple, interconnected assets, a system-level digital twin can exploit the pooled data streams of individual assets within the network to extract more valuable insights. In the case of a fleet of vehicles, for example, the foundational research outlined in the first case study hints at how pooling data from across the fleet could allow better predictions to be made about the performance of individual vehicles. Meanwhile, the second case study, based on an underground, hydroponic farm, demonstrates the benefits of a digital twin that brings together data from a collection of different components all within a common setting. In this case, integrating data from a variety of sources allows growers to optimise conditions and resource use within a defined space.

More efficient system-level digital twins and faster simulations will increasingly allow more complex questions and scenarios to be addressed. As work under the EDT theme demonstrates, it is possible to build these system-level digital twins without resorting to black-box methods that obscure their internal workings. The data requirements and computational burden of operationalising these more complex digital twins may also be decreased by building generalisable platforms – based on open-source infrastructure – that can be tailored to individual projects, for example, in unique settings like an underground farm.

Case study: Modelling machine fleets

Models built by ASG researchers to analyse the performance of truck fleets and wind turbines are laying the foundations for digital twins for whole populations of machines. This population-level approach demonstrates how digital twins for systems of multiple similar assets can facilitate deeper insights than from individual assets.

The researchers developed their population-level models for machines based on models used for human populations in political science and health research. Their models enable sharing of data across a machine fleet, borrowing insights from trends within the fleet when data is scarce for any specific machine. To verify their approach, the researchers collaborated with commercial vehicle manufacturer Scania to develop digital twins for predictive health monitoring of their heavy-duty trucks – their models use data from truck fleets to predict the time-to-failure more accurately for components on individual trucks. The aim is to help Scania reduce down time and maintenance costs for its trucks by providing more accurate predictions to inform maintenance schedules. By prioritising explainability of their models, the researchers are also enabling users to understand the specific drivers of engine health degradation.

The value of this population-level approach is in the pooling of data streams from a large number of machines to improve the ability of digital twins to predict any given machine's future health. This can help to address situations where data streams from individual machines may be spotty and inconsistent – perhaps due to the monitoring of a rare occurrence like mechanical failure. In the case of new machines, for example, insights gained from the wider fleet may be the only way to make predictions about time-to-failure as no prior data is available for the individual machine.

Demonstrating the value of a cross-sector approach, ASG researchers have applied similar models in the wind energy space, working with industry stakeholders and the University of Sheffield to identify opportunities for supporting more reliable performance monitoring of wind turbines as well as modelling of power output. They are now applying their population-level approach to develop other fleet-level digital twins in the energy and civil infrastructure sectors.

Case study: Data-driven decisions at an underground farm

Urban agriculture company Zero Carbon Farms uses a digital twin called 'CROP' to guide management of its underground, hydroponic farm. The farm, which produces 100–150 tonnes of salad each year, is based in tunnels designed for use as an air-raid shelter near Clapham tube station in London. Its digital twin is the product of a long-term, multidisciplinary collaboration between growers from Zero Carbon Farms, researchers from the University of Cambridge and ASG programme, and software engineers from the Turing's Research Engineering team.

CROP (the Crop Research Observation Platform) realises the end-to-end vision of digital twins, from sensor data all the way to simulation capabilities. Sensors allow growers to monitor lighting, humidity, water quality and temperature in the tunnels, and analyse the effects of past conditions on yields. The streaming data is also used to calibrate a physics-based model of the farm, grounding it in reality. By incorporating local weather forecasts into this physics-based model, CROP can predict future conditions in the tunnels up to three days in advance, guiding proactive responses, such as changing lighting settings or ventilation to protect crops in a heatwave.

The hybrid data-physics approach used in CROP incorporates physics to reduce reliance on data-driven methods and increase explainability, enabling decision makers to understand and verify its results. From the grower's perspective, the platform provides the most advanced tools possible for optimising crop growth conditions and yields, and efficiently managing resources. With the digital twin fully integrated into the farm's daily operations, the Zero Carbon Farms team can monitor the farm from the propagation of crops to harvest. Meanwhile, the platform can still be expanded to integrate new technologies such as computer vision for visual monitoring of plant growth.

Whilst the current digital twin is bespoke to Zero Carbon Farms' London-based project, the research team hopes to develop the open infrastructure created by the research engineers at the Turing as a widely available digital twin framework to support other indoor farming projects. In this sense, the broader value of the platform lies in its reproducibility, which provides a means to lower the barrier to entry to digital twin technology within agriculture and thereby realise its potential power for informing urban food systems around the world.

2.3 Ecosystems of digital twins

Having described digital twins of increasing levels of complexity, we now advance the concept of a system that can connect multiple digital twins of different types and complexity levels: an ecosystem of digital twins (Figure 3). This multi-level perspective allows us to organise digital twins into a hierarchy as follows: asset (or process), system of assets and ecosystem.

Ecosystems of digital twins are capable of representing interconnected systems as complex as an entire city or national health service. Reflecting the intricate and interconnected reality of ecosystems in nature, ecosystems of digital twins can operate within different ranges of spatial resolution – from building to country-level – and time, potentially spanning decades of data. These capabilities position ecosystems of connected digital twins for use in decision and policy-making contexts, where it is important to capture a nuanced state of reality across spatiotemporal scales.

The added value of an ecosystem of digital twins lies in its potential to provide insight and capabilities far beyond that of the individual digital twins it is comprised of. However, extracting this value requires a means for the connected digital twins to collate information and resources – for instance, to combine data curated within healthcare systems with movement data controlled by mobile phone networks without violating privacy constraints. Thus, to enable the generation of unique insights that are only possible at the ecosystem scale, we require new data infrastructure and technical standards for linking together heterogeneous digital twins. These key components will provide the means for sharing of information across an ecosystem of federated assets – assets that are under different jurisdictions, with differing privacy and security constraints – in a mutually understandable and secure manner. With this shared information, it will be possible to achieve more together than apart.

The EDT theme has built upon its years of research in digital twins to create models, methods, and standards for assembling connected, interoperable digital twins within ecosystems. This section introduces advancements in ecosystems of digital twins through projects focused on digital twins for analysing and improving energy efficiency, and for supporting health planning and data-driven policy intervention for climate mitigation. These examples demonstrate how work under the EDT theme is building towards federated ecosystems of digital twins by combining data on demographics, buildings, energy, healthcare, environment and movement patterns. They are also instructive in drawing out the social and behavioural dimensions of digital twins. In contrast to digital twins of physical systems described at the asset and system levels, in these examples, complex human behaviours and social interactions are represented. By taking an ecosystems approach, it will be possible to benefit from complementary insights gained from these different types of digital twins.

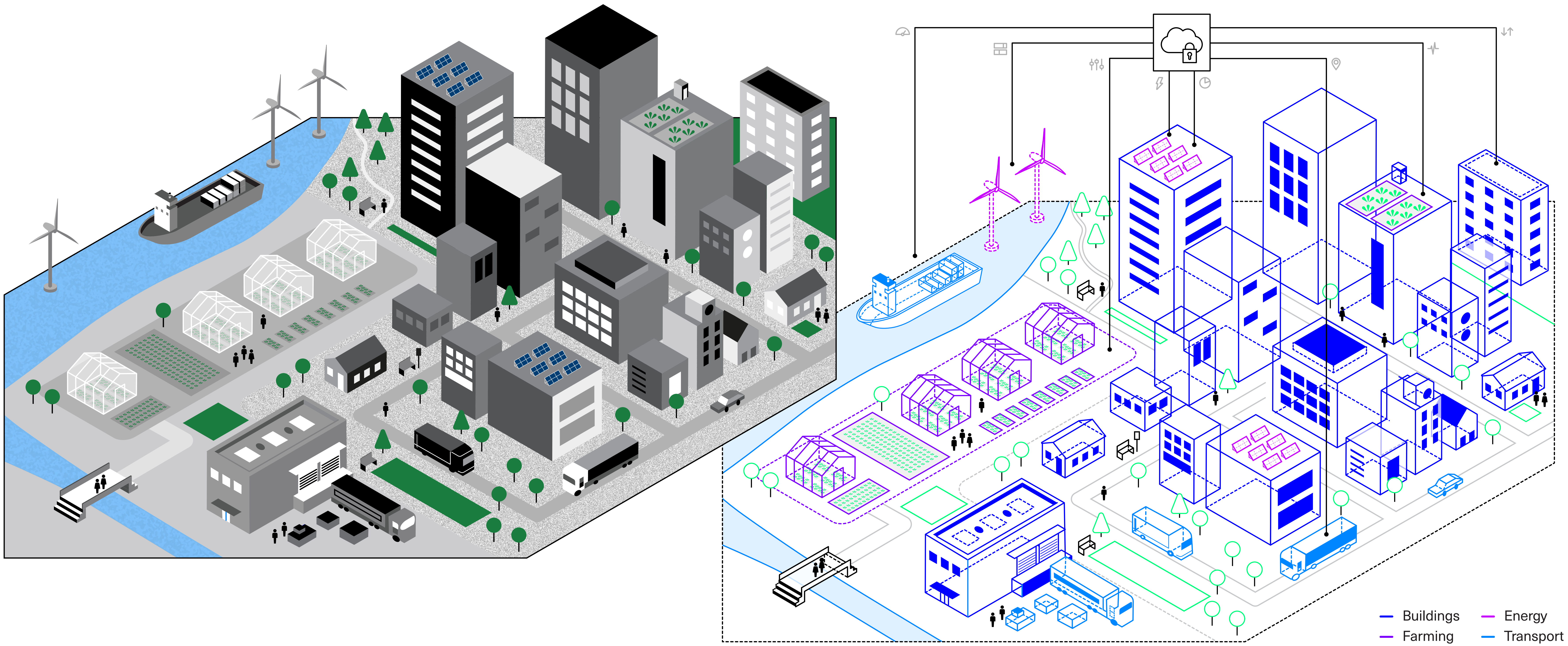


Figure 3. An ecosystem of digital twins. Digital twins for different types of assets and systems of assets, under different jurisdictions, are connected within in a virtual, ecosystem-like landscape, through the secure sharing of data.

Importantly, at all levels, the successful functioning of digital twins in the real world is dependent on how society chooses to utilise them in decision-making.

Both projects described in this section take advantage of the computational modelling technique known as 'microsimulation' to combine information from different domains. Microsimulations are large-scale simulations that account for the differences and interactions between individuals. Through microsimulation, it is possible to provide a granular view of a simulated (synthetic) population that combines information about housing, health, mobility, economic status and other factors in a way that mirrors important structural aspects of the real population. The artificially-generated data attached to each simulated individual in the population could include everything from the number of bedrooms in their house to their occupation and risk of heart disease. This powerful technique, when exploited in combination with information shared securely through the federated system that we describe above, can provide the basis for a connected, ecosystems approach to digital twinning.

Applied in the context of planning or epidemics, we can now envisage how this ecosystems approach can be of benefit. In the example of monitoring an infectious disease, combining foundational microsimulations with data and models from across a federated system of digital twins – to take in information such as health records, working patterns and travel data – could be invaluable for understanding transmission and designing interventions to minimise infections or hospitalisations. Whilst we do not yet have all the components we need to extract the maximum value from this ecosystems approach, in Section 3 we propose the next steps towards making our vision of an ecosystems approach to digital twinning a reality.

Case study: Connecting digital twins to support policy-making

The DyME modelling framework developed by ASG researchers provides the building blocks for a whole series of connected simulation applications that can inform policy-making. Bringing together datasets, tools and methods that can be used for large, granular simulations spanning diverse domains – from epidemics to the environment – it represents an ecosystem-like approach to integrating digital twins.

Created through partnerships nurtured in the Royal Society's Rapid Assistance in Modelling the Pandemic (RAMP) programme and bringing together expertise from fields as diverse as epidemiology, computer modelling, traffic planning and financial markets, DyME was originally developed by ASG researchers as the Dynamic Microsimulation Model for Epidemics. This model functions as a digital twin for simulating epidemics that can quantify the impact of lockdown timings on COVID-19 infections. Its microsimulation approach, which takes account of differences in the living situations, health status and behaviour of individuals in a population, is capable of producing localised insights and ‘what if?’ scenario assessments. When combined with appropriate data and models from across multiple domains, these types of insights could enable more targeted policy interventions in future epidemics.

Inspired by the possibilities of DyME for supporting targeted public health interventions, another ASG team led the creation of DyME-CHH (for Climate, Heat and Health). This model can form the basis of digital twins for analysing the effects of climate-related heat exposure on health. Data from the model will feed the University of Exeter’s Local Climate Adaptation Tool (LCAT), designed in collaboration with local authorities to support evidence-based policy-making at the local level. The microsimulation approach used by DyME-CHH will enable decision makers to understand the impacts of heat exposure at all levels, from national and regional levels right down to the individual level, for future years and under different climate scenarios.

DyME-CHH uses an ASG-funded, open-source tool called the Synthetic Population Catalyst (SPC) to generate its synthetic population data, which is integrated with climate projections and data on mobility, buildings and public spaces, to feed the model. From initial applications in disease modelling with DyME for epidemics, SPC now represents a shared resource and easy-to-use tool for quickly producing synthetic population data that can power complex demographic models across multiple domains.

Case study: Analysing home energy use and efficiency

Home energy use accounts for more than 10% of global carbon emissions and is a major contributor to the rising cost of living in the UK. However, differences in types of buildings, and how they are heated and insulated, can lead to very different emissions and costs for the same-sized homes. The EnergyFlex project was created by ASG researchers to identify building-level inequalities in energy efficiencies across the UK and support targeting of local policies to reduce household energy use.

The EnergyFlex team created a digital twin of the UK housing population using publicly available data and microsimulation techniques – drawing on the Synthetic population estimation and scenario projection (SPENSER) model developed by ASG researchers to generate a synthetic housing population that preserves important characteristics of the real population. They then applied an energy model to this

synthetic population, allowing simulation of energy usage and performance at the household level. Data on the synthetic population includes socioeconomic data about the populations living in the buildings, as well as the energy use profile of each building based on features like building age, location, and material composition. The synthetic approach allows housing data to be used at a finer scale than is possible for real public data because it does not compromise residents' privacy.

Using their synthetic housing population, the team created a series of microsimulation and visualisation tools that can be used to help identify strategies for more effective residential decarbonisation policies. Instead of a one-size-fits-all approach towards retrofitting action, such as installing smart meters in every home, EnergyFlex allows for targeted policy creation and evaluation, by quantifying the impact of different interventions at local, regional, and national scales.

By expanding knowledge and 'what-if' scenario-testing capabilities for decision makers, EnergyFlex can help public and private sector stakeholders identify the types of buildings and interventions that should be prioritised. In order to guide their research, and improve the accessibility and usability of the tools, EnergyFlex engaged with local government and private sector organisations, including Haringey Council and Parity Projects, from the start.

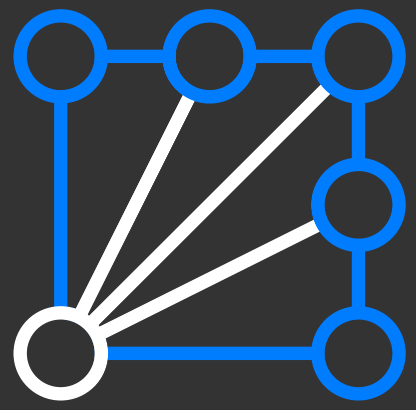
EnergyFlex is available on a cloud-based platform provided by the Data & Analytics Facility for National Infrastructure (DAFNI), allowing users access to the deep insights the digital twin provides without requiring specialist knowledge.

3. Recommendations

Based on learnings from the ASG programme under the EDT theme, our community of research supports the following three recommendations for future directions of innovation in the digital twins space. These recommendations will help to inform the new Turing Research and Innovation Cluster in Digital Twins, which builds on the wider ASG portfolio of digital twins research and innovation.

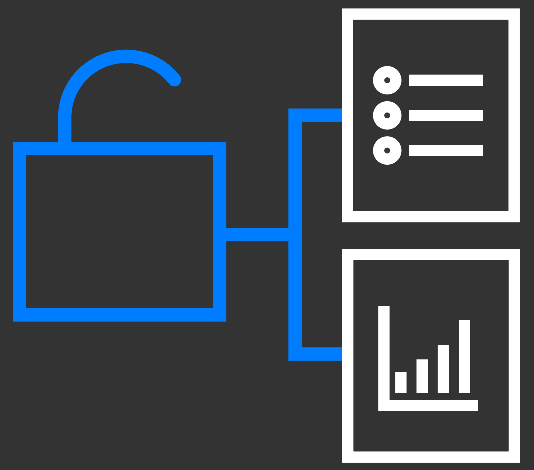
1

Elevate cross-disciplinary activities and spaces for digital twins



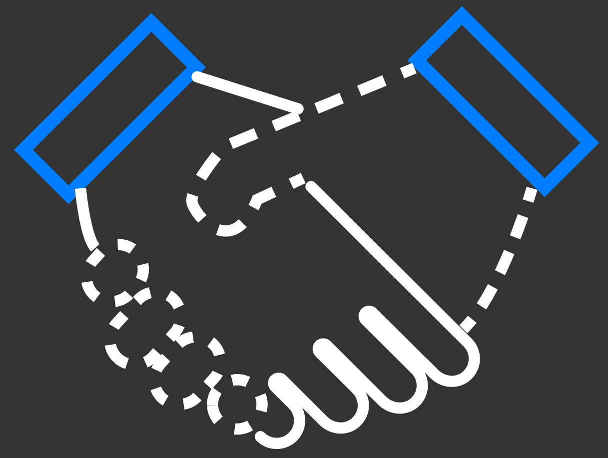
2

Invest in open infrastructure, with a focus on data and technical standards



3

Prioritise tools for building trust in and understanding of digital twins



1 Elevate cross-disciplinary activities and spaces for digital twins

The inherently multidisciplinary nature of digital twins exposes skills gaps across the sectors and individual organisations where they are being deployed. For example, engineers may have expertise in modelling and simulation, but have less working knowledge of big data techniques or state-of-the-art machine learning methods. Specific sectors may be experts in how to work with and interpret their domain's datasets, such as in urban analytics and their work with geospatially-oriented datasets, but may not be up to date in the latest advancements in model calibration, uncertainty quantification and machine learning-based methods.

Multidisciplinary environments help to address these skills gaps by convening researchers from a broad range of roles and fields. Under the EDT theme, experts from civil engineering, urban geography, economics, mathematics, public health, computer science and more are assembled to work on common problems. Such interdisciplinary spaces create opportunities for sparking new ideas and connections, sharing tools and methods, and catalysing collaboration. Moreover, greater involvement for the social and economic sciences extends the range and quality of applications of digital twins from purely physical systems towards digital twins representing complex human behaviours, increasing their relevance for policy-making and civic design. Prioritising cross-disciplinary ways of thinking and working will lead to outcomes that are useful and used by more people in different sectors and organisations, paving the way for the creation of standards and infrastructure that make possible the proliferation and operationalisation of digital twins at scale.

Cross-disciplinary collaborations can also help initiate the development of data pipelines between connected digital twins within larger ecosystems of digital twins. Accomplishing this will require engaging non-traditional roles and expertise, such as Data Stewards, Research Software Engineers and Research Application Managers – a role created at the Turing to help drive forward real-world applications of research by facilitating collaborations beyond academia. These roles can help shape research outputs to meet necessary levels of system safety, data interoperability, and software sustainability so that they can be integrated into industry workstreams and inform government policy.

The translation and extension of digital twins research into forms that are relevant to real-world applications should be cultivated within traditional academic institutions as well as national institutes, like the Turing in the UK, that bring together different domains, roles and organisation types. The new methods of collaboration developed in these interdisciplinary spaces should be shared with science and innovation networks worldwide.

2 Invest in open infrastructure, with a focus on data and technical standards

As we have suggested, the full potential of digital twin technology for addressing global challenges through real-world applications like planning, policy interventions and safety-critical decision-making can only be realised by connecting together digital twins for multiple different assets, processes, systems and systems of assets within ecosystems of digital twins. Building these types of interconnected systems at scale will require new data infrastructure and new technical standards. Under the EDT theme, we have identified the fundamental components required for meaningful deployment and use of ecosystems of digital twins based on this new infrastructure:

- Access to open data resources that can benefit different domains and organisations (e.g. synthetic population datasets).
- Technical standards for linking together heterogeneous digital twins acting across different temporal and spatial scales.
- Tools for extracting insights from interconnected digital twins that are meaningful and understandable for diverse users and use cases.
- Tools for enabling ecosystems of digital twins in federated settings.

ASG researchers have developed open datasets and data infrastructure that can serve as a foundation to address all four of these fundamental challenges. Our teams have created these open resources as part of our research work and have prioritised sharing them openly so that others can make use of these datasets and tools. They include resources like the Synthetic Population Catalyst, which can be used to generate synthetic data to feed population models across multiple domains. Like other Turing tools, they are freely available, well-documented, and created with a broad spectrum of users in mind, including those with less technical experience.

Building open infrastructures to connect digital twins within ecosystems can contribute to the legacy of publicly funded research work, helping to ensure that whatever is developed is available for use by others working on similar problems in academia, government, or industry. Prioritising investment in open resources can also help reduce the duplication of effort and instead advance innovation cycles more quickly. In this way, investments in open infrastructure can benefit researchers and stakeholders beyond a project team, in the UK and around the world.

3 Prioritise tools for building trust in and understanding of digital twins

Even if a technology is capable of impressive feats, its real-world application may be limited by fundamental concerns around how it was developed and deployed. In the field of AI, for instance, concerns have often surrounded biased or privacy-infringing training datasets, black-box algorithms that obscure their internal workings and discriminatory outcomes. Thus, an important lesson to learn from the AI space is that how a technology is created and how it interfaces with people matters as much as how it performs.

To address these pressing concerns, it is important to prioritise building tools for trust and understanding of digital twins alongside investments in advancing the technology itself. Within safety-critical domains such as aerospace, maritime and critical infrastructure engineering, tools are needed to provide verification and assurance of digital twins in order to assess risk and confidence levels. Within public-facing domains such as health or policy-making, tools that can facilitate data-driven explanations are required to expand the confidence and scope of knowledge a user has beyond their domain-specific expertise for making difficult decisions.

Another dimension of trust relates to the ability of an ecosystem of digital twins to operate under constraints of privacy and federation, where data pertaining to participating individuals and institutions must remain private, while still contributing value to the wider ecosystem. The adoption of privacy-enhancing technologies such as synthetic data generation and methods for assimilation of data under federated settings can reinforce the importance of and capabilities for building trust in the emerging technology of digital twins.

Trust in an emerging technology is built on collaboration and understanding between those involved in its development and its users. Thus, for a more sustainable and trustworthy direction of innovation, we should co-create tools and standards in partnership with industry and policy makers.

4. Conclusions

In this white paper, we have shared the unique perspective gained by ASG researchers through their work on digital twins under the EDT theme. This perspective comes from the Turing's foundational and applied research in digital twins across multiple disciplines, industries and sectors.

We have outlined a hierarchy for classifying digital twins according to their level of complexity and shown how, under the EDT theme, ASG researchers are developing the means to share information between digital twins from different levels of this hierarchy to build ecosystems of connected digital twins.

It is our belief that the deep insights provided by digital twins and ecosystems of digital twins can help researchers, industry and policy makers globally to tackle some of the most pressing challenges facing business, society and government – from decarbonisation to infectious disease. However, the full potential of the technology cannot be realised unless certain technical, social and institutional issues are addressed. Therefore, we have made three recommendations that can advance the real-world applications and adoption of digital twins, and help make ecosystems of digital twins a reality.

In the first place, building ecosystems of digital twins demands a multidisciplinary approach to research and we therefore recommend a higher level of priority to establish and nurture such environments. Second, the need for diverse and complementary expertise dictates an enhanced capability for the sharing of tools and methods, leading to our recommendation for greater focus on open research and infrastructure, with associated standards for co-production and sharing of resources, including data. Third, the need to operate across conventional boundaries between systems, and their associated communities of developers and users, dictates the need to use methods that can ensure privacy, and enhance trust and understanding of the tools and their application.

By following these recommendations, we can build robust ecosystems of digital twins that empower researchers, engineers, and decision makers to tackle the challenges of tomorrow. With the right investments and a commitment to collaboration and transparency, we can unlock new insights and opportunities that will benefit society as a whole.

Acknowledgements

The authors would like to thank the following people for their input during the development of this paper and for their work on the profiled ASG projects:

Dani Arribas-Bel, Michael Batty, David Beaven, Fernando Benítez-Paez, Ruth Bowyer, Lawrence Bull, Dustin Carlino, Ruchi Choudhary, Fehmi Cirak, Mafalda Cunha, Theo Damoulas, Eric Daub, Domenic Di Francesco, Alistair Ford, Mark Girolami, Jim Hall, Nick Holliman, Polly Hudson, Zaynab Ismail, James Lloyd, Nik Lomax, Ben MacArthur, Omar Matar, Aida Mehonic, Richard Milton, Grigorios Mingas, Mark Mon-Williams, Neave O’Clery, Martin O’Reilly, Indranil Pan, Achintya Rao, Jonathan Rowe, Hadrien Salat, Gavin Shaddick, Hannah Sweeney, Lukasz Szpruch, Kirstie Whitaker, Anna Zanchetta.

We also thank the creative communications company Soapbox for designing the paper, the digital twins illustrations and the recommendations icons.

This work was supported by Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1 and EPSRC Grant EP/W006022/1, particularly the 'Complex systems engineering', 'Urban analytics' and 'Ecosystems of digital twins' themes within those grants and The Alan Turing Institute.

Image credits

Front and back cover: Novikov Aleksey
Page 4: Yiucheung
Page 11: Aun Photographer
Page 15: Betsy McCue
Page 20: K303

The Alan Turing Institute

turing.ac.uk
@turinginst