Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

Beyond Connected Digital Twins – from GIS to The World Avatar

Yong Ren Tan[1], Markus Hofmeister[2], Shin Zert Phua[1], George Brownbridge[3], Kushagar Rustagi[4], Jethro Akroyd[1,2], Sebastian Mosbach[1,2], Amit Bhave[3], Markus Kraft[1,2,3,4]

released: December 23, 2024

[1] CARES
Cambridge Centre for Advanced
Research and Education in Singapore
1 Create Way
CREATE Tower, #05-05
Singapore, 138602

[2] Department of Chemical Engineering
and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

[3] CMCL
No. 9, Journey Campus
Castle Park
Cambridge
CB3 0AX
United Kingdom

[4] CMPG
GRIPS – Gründerinnenzentrum Pirmasens
Delaware Avenue 1–3
66953
Pirmasens
Germany

Preprint No. 332

UNIVERSITY OF CAMBRIDGE

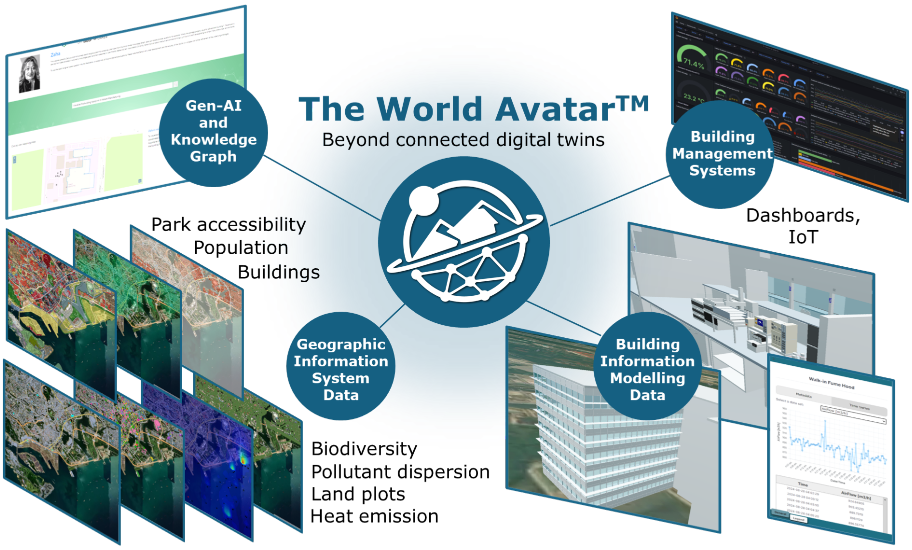
Keywords: Digital Twins, Knowledge Graphs, The World Avatar, Smart Cities, Decarbonisation

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

E-Mail: mk306@cam.ac.uk
World Wide Web: `https://como.ceb.cam.ac.uk/`

Abstract

This paper introduces The World Avatar (TWA), an open-source knowledge-world model developed to represent complex, dynamic systems by integrating spatial, temporal, and real-time data. TWA showed capability in addressing critical challenges in urban management, including resilience, infrastructure planning, and decarbonisation. We trace the origins and evolution of digital twins, culminating in the development of TWA as an ultimate form of knowledge-world model. The unique capability of TWA to facilitate seamless interaction between human stakeholders and diverse data sources is highlighted, showcasing its potential to solve urban challenges. This paper presents select use cases that demonstrate the abilities of TWA to integrate Geographic Information Systems (GIS) and Building Information Modelling (BIM) to address key exemplary issues related to climate change mitigation, energy optimisation, and strategic infrastructure placement. Our demonstrations revealed that TWA surpasses the limitations of traditional digital twins by enabling seamless integration of various data sources and supporting dynamic real-time decision making. Interfaces such as GIS, BIM, dashboards, mobile applications, and augmented reality enhance human-machine interactions within TWA. The presented use cases demonstrate the generation of impactful insights that would not have been possible without TWA's holistic integrated approach. As an open community project, TWA offers a scalable and adaptable path for cities to adopt sustainable and data-driven solutions, and to develop more resilient and intelligent urban infrastructures. We invite the community to collaborate and enhance infrastructure development using the innovative capabilities of TWA.



Highlights

- Integration of geospatial data and knowledge models with dynamic, real-time insights.
- TWA enables strategic decisions on climate resilience, infrastructure placement, and heat planning.
- Drives detailed analysis of decarbonisation policies, revealing socioeconomic impacts.

Contents

1 Introduction

Rapid advances in digitalisation are fundamentally transforming the field of urban survey data, moving away from traditional paper-based maps to highly sophisticated Geographic Information System (GIS) applications and cutting-edge multidimensional analytical tools [21, 52]. This transformation not only improves the accuracy and efficiency of urban planning, but also unlocks new opportunities for geospatial data to drive evidence-based decision making [50, 54]. Beyond conventional GIS applications, geospatial intelligence is increasingly being used in different areas such as decentralised energy systems powered by renewable sources, further advancing the goal of creating sustainable and resilient urban environments [7, 16, 28, 64, 87].

The achievement of such resilience and sustainability in urban planning requires the seamless integration of data from multiple domains (*e.g.* transportation, energy systems, public health, environmental monitoring, economic development, *etc.*) to allow holistic, consistent and actionable analyses for the development of integrated and comprehensive solutions [15, 42, 44, 78, 86]. Although GIS platforms play a critical role in managing and analysing geospatial data, they face challenges in supporting the full range of demands of modern urban ecosystems. For example, methods for reusing data and knowledge between projects remain limited [26, 27]. This often requires developers to recreate data models for each project, increasing costs and time while limiting the ability to comply with regulations or build on insights from previous GIS-based initiatives.

Organising GIS data to foster cross-disciplinary and cross-sector collaboration presents additional complexities [41]. Diverse file formats (*e.g.* KML, CSV, ACS, GeoJSON, TIFF, GML, *etc.*), and heterogeneous data structures, including structured and unstructured elements, often lead to inefficiencies in managing and analysing data. Unstructured data such as unclear tags, notes, and free-text remarks complicates interpretation and risks the loss of critical information [73]. In addition, implicit relationships between datasets frequently remain unrecognised, limiting the potential for deeper insights and understanding, a phenomenon often referred to as the "Semantic Gap" [53].

While efforts to integrate and share geospatial data, such as the Federal Geographic Data Committee, National Spatial Data Infrastructure, ISO 19115 standards, and the US Geological Survey's National Map, have advanced the field, existing platforms still face scalability and interoperability limitations [69, 73]. This can result in fragmented systems, with isolated components that hinder the effective use of urban data. These limitations underscore the growing need for scalable, interoperable solutions that can enhance data reuse, streamline collaboration between disciplines, and support the development of smart and sustainable cities aligned with the principles of environmental, social and governance, such as sustainable resource management, green infrastructure, and climate resilience, while ensuring equitable access to housing, transportation, and public services.

A promising approach to overcoming these limitations lies in the adoption of knowledge graphs (KGs), which offer the ability to connect disparate data and knowledge domains, fostering truly interoperable systems [9, 22]. KGs, prominently featured in Gartner's Artificial Intelligence (AI) hype cycle and recognised as essential enablers in their Emerging Tech Impact Radar, are becoming central to the development of infrastructures in the future “smart” world (see **Figure 1**) [56]. By resolving data silos and facilitating seamless integration, KGs pave the way for the efficient use of urban data across diverse applications, including digital twins, intelligent systems, and generative AI solutions.

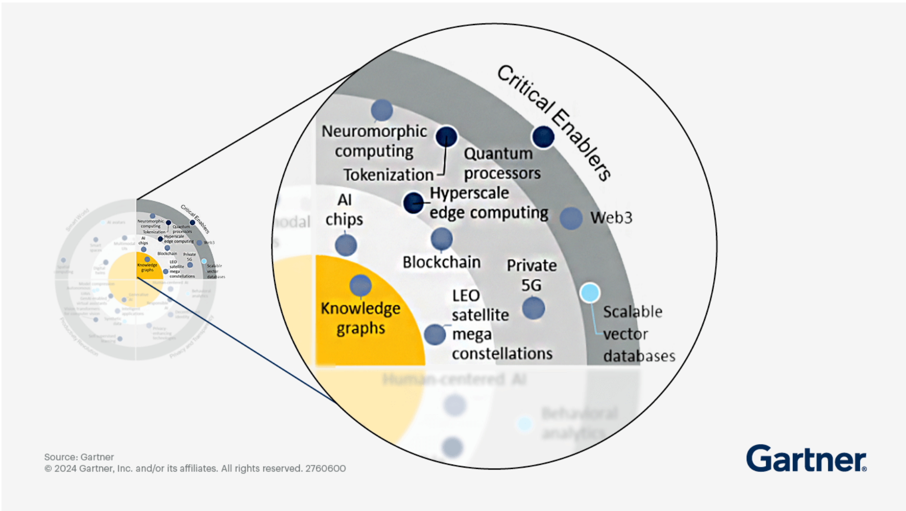


Figure 1: *The Gartner Emerging Tech Impact Radar 2024 recognises knowledge graphs as critical enablers to unlock productivity gains and create “smart” urban environments. Figure taken from [56].*

The dynamic KGs of The World Avatar (TWA) demonstrate how KGs can integrate and harmonise data from multiple dynamic sources to support advanced analyses in both spatial and temporal dimensions [48, 61]. When analysing the future use of energy systems in the sustainable city of the future, inter-sector collaboration is needed to maximise synergies between inter-connected systems. This includes power generation, urban infrastructure, and healthcare. These scenarios necessitates the use of KGs in the development of connected digital twins: distributed collaborative entities that share data and computational resources. These digital twins have the potential to address complex and multidisciplinary challenges by integrating various domains, providing a powerful framework to tackle the intricacies of modern and interconnected urban environments [4].

The **purpose of this paper** is to introduce and demonstrate the capabilities of TWA in addressing the key challenges faced when using GIS platforms for inter-disciplinary projects, particularly in relation to fragmentation, interoperability, and reuse of system or data. TWA offers a transformative approach by providing a unified framework to integrate GIS data, Building Information Modelling (BIM) data, and other domain-specific datasets and knowledge (*e.g.* laws, regulations, *etc.*). Through its use of dynamic KGs, TWA facilitates the seamless integration and harmonisation of diverse data sources, enabling the development of truly interoperable and scalable solutions for smart city planning. This paper presents the initial steps and key applications of TWA in creating smart, resilient urban environments, with a particular focus on climate change mitigation and adaptation. Using the advanced technology stack in TWA, we demonstrate how cities can progress toward more sustainable, efficient, and data-driven planning processes, ultimately contributing to the development of smart cities and resilient infrastructure in the face of global challenges.

2 The evolution of digital twins

The concept of digital twin originated in the 1960s at NASA (National Aeronautics and Space Administration) and gained prominence after the Apollo 13 incident [8, 29]. Initially, digital twin was developed to support the Apollo programme, where digital components were integrated with physical systems to facilitate real-time failure analysis. Today, NASA and the aerospace community continue to advance digital twin technology with increased fidelity, capable of simulating physical systems in extreme environments [29], which is vital for missions such as Artemis to the Moon and Mars, where limited connectivity and minimal human intervention are expected.

In the 2010s, digital twin technology was reintroduced to a wider range of sectors [30–32, 46]. Digital twin now provide detailed digital representations of assets, processes, and systems, capturing their current state and behaviour over time under varying conditions and constraints [40]. This capability has successfully addressed numerous real-world challenges in diverse industries [13, 19, 75, 81]. The global digital twin market is projected to grow from USD 10.1 billion in 2023 to USD 110.1 billion by 2028, driven primarily by the increasing demand for digital twin in the healthcare industry and the increased focus on predictive maintenance [65]. Additionally, 75 percent of large enterprises are currently investing in digital twin technology to scale AI-driven solutions [20].

Several organisations have played a pivotal role in the advancement of digital twin technologies and the promotion of their adoption in various industries. In particular, the Digital Twin Consortium (digitaltwinconsortium.org), Industrial Digital Twin Associations (industrialdigitaltwin.org) and the Digital Twin Hub (digitaltwinhub.co.uk) have been at the forefront of promoting digital twins applications in manufacturing, production and operations.

However, most digital twin solutions remain isolated and lack interoperability due to variations in configuration, hardware, and software, often driven by individual funding sources or proprietary interests [13, 63]. Interoperability, defined as the ability of tools, systems, and data to exchange and utilise each other's functionalities [24]—is essential for reusing data and software assets and addressing cross-domain challenges in a comprehensive way [80]. Hence, in this section, we aim to go through the evolution to shaping the concept of digital twin and beyond towards the concept of TWA. The illustration of the evolution is as shown in **Figure 2**.

Data & Models Traditionally, scientists, engineers, and quantitative professionals integrate 'data and models' in a largely *ad hoc* and manual way. Although the 'models' can range from simple, back-of-the-envelope estimates to complex computational models, the integration process to allow for the creation of digital representations of physical systems remains inefficient and labour-intensive. Manual and siloed data handling, model formulation, and analysis on spreadsheets (or even on paper) often lead to considerable time investment and a higher likelihood of human error. Furthermore, this inconsistent and fragmented approach hinders reproducibility [45], as critical integration steps are often undocumented or difficult to transfer to others, making it challenging to replicate findings and impeding progress and collaboration.

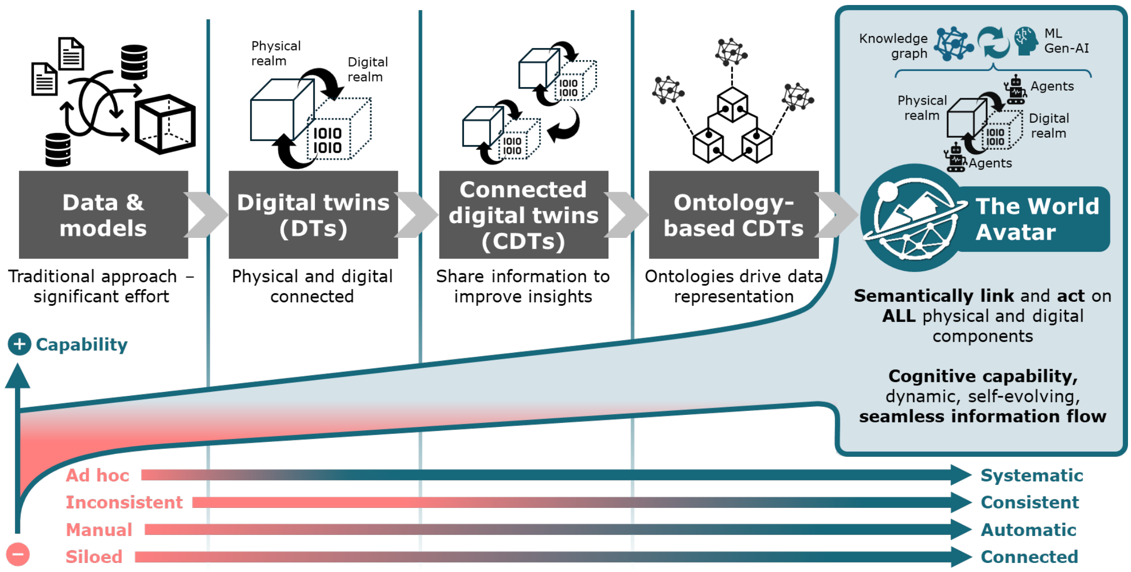


Figure 2: *The evolution of digital twins to The World Avatar.*

A significant limitation of the 'data and models' methodology is the heavy dependence on specific sets of observational data to estimate model parameters. As most models in the literature lack provenance information, inconsistencies often arise among models that attempt to describe the same physical phenomenon. The comparison of different models is further complicated due to differences in data sets, assumptions, and underlying methodologies in model development, even when they (aim to or are able to) achieve similar outcomes.

Furthermore, the integration of traditional 'data and models' methodology generally lacks a systematic feedback mechanism between digital models and the physical systems they represent. Typically, once a digital model is created, it remains static, without continuous updates to the digital model based on the change in the physical world, *i.e.* new experimental data or observations in the real world. Even if all the process conditions for an augmentation of the classical model can be obtained or updated, they are often performed manually *via* human input. This is prone to human error and lacks real-time fidelity. This absence of a dynamic feedback loop limits the adaptability of the digital model, reducing its accuracy over time when the physical conditions are altered.

Consequently, traditional approaches to data and the model are increasingly inadequate to meet the efficiency, automation, and real-time responsiveness required by contemporary scientific and engineering challenges [17]. While some models have begun to incorporate real-time data through Application Programming Interfaces (APIs), such as those used in classical control systems, these APIs often come with limitations, including restricted real-time data scope and potential vendor lock-ins. Such models that rely partially on real-time input are sometimes referred to as "digital shadows" [46]. Despite the presence of real-time input, they inherit many of the same limitations as static models, including restricted adaptability and limited feedback capabilities, which prevent them from fully leveraging real-time data for dynamic system updates.

Digital twin and connected digital twins As we increase complexity in the 'Data & Model' mode of working, we have the emergence of digital twins concept. In this paradigm, the physical and digital worlds are interconnected in a structured manner. The fundamental idea is to facilitate an exchange of information, allowing changes in the physical world to be reflected in the digital world and enabling insights from the digital model to inform and influence the physical counterpart. As such, in the industry, Model Predictive Controllers (MPCs) are often regarded as a foundational form of digital twins, as they analyse past data to generate and adjust control signals in real-time. In the literature, digital twin has received numerous definitions. The Digital Twin Consortium has defined digital twin as an integrated data-driven virtual representation of real-world entities and processes, with synchronised interaction at a specified frequency and fidelity [59].

Meanwhile, based on a review of the literature by Barricelli et al. [13], digital twin is often related to a virtual replica or description of a physical counterpart, which serves as an integrated construct that mirrors and simulates the real-world system. The digital twin ties together and links real-time information from the physical object with its digital clone, enabling prediction, testing, and analysis of its behaviour and performance under various conditions. Thelen et al. [75] proposed a five-dimensional framework for defining digital twin, which is based on the flow of data between physical and virtual systems. This includes physical system, digital system, update engine, prediction engine, and optimisation. This framework offers a comprehensive view of the interactions between physical and digital systems, which includes data exchange, modelling, and actions.

However, sole digital twin implementations are often isolated, resulting in siloed systems which have many challenges [25, 49, 70]. Naturally, there is a desire to develop connected digital twins to exchange information, thereby enhancing insights and interventions. Unfortunately, this integration is hindered by inconsistent data models. Potential solutions include adhering to a single vendor or introducing numerous APIs, but both approaches present the challenge of vendor lock-in and interoperability. Data and model feedback are essential for influencing physical systems, but disconnected implementations lead to the development of silos. Ensuring consistent data structures and concepts is crucial to avoid vendor lock-in and the complexity of managing multiple APIs.

Ontology-based connected digital twins Connected digital twins based on ontologies use ontologies to enhance data representation [18, 35, 58], The use of ontology-based connected digital twins is due to the lack of alignment in data structures, accuracy and interoperability in just the connected digital twins. With ontology-based digital twins, *i.e.* incorporating ontological descriptions into the models can overcome the interoperability issue. In connected digital twins based on ontology, the domain expert is key in the establishment and selection of ontologies and meta-models based on the input of domain and ontology knowledge [35]. This is so that all relevant state-of-the-art ontologies are taken into account. Then, a use case expert will further select relevant parts of the ontologies based on the requirements of a specific use-case [35]. For example, in the mechanical industry, ontology-based digital twins can handle heterogeneous data of multiple sources of mechanical parts with the room for extension due to the use of ontologies in connected digital twins, which can benefit the product manufacturing industry [12].

Cognitive digital twins, inspired by cognitive science, represent an evolution of ontology-based digital twins that leverages advanced technologies, *e.g.* semantic modelling, cognitive computing, and model-based systems engineering to enhance decision-making and system optimisation capabilities [23, 82, 90]. The concept first emerged in 2016 [3] as a digital representation encompassing intelligent capabilities that span all phases of the life cycle of its physical counterpart [70]. As the name implies, cognitive digital twins have human-like cognitive functions – selective attention, perception, memory, learning, and problem solving — which make them autonomous and capable of evolving with their physical counterparts in all phases of the life cycle [5, 55, 82, 90]. The vision of cognitive digital twins has a few characteristics: digital twin-based, has cognition capabilities, has full lifecycle management, has autonomy capabilities, and is able to continuously evolve with the real system throughout the lifecycle [90]. A formal definition of cognitive digital twins from Zheng et al. [90] is given as a digital representation of a physical system that is augmented with certain cognitive capabilities and support to execute autonomous activities; comprises a set of semantically interlinked digital models related to different lifecycle phases of the physical system including its subsystems and components; and evolves continuously with the physical system across the entire lifecycle.

Some notable cognitive digital twins have demonstrated practical benefits of cognitive digital twins in complex industrial systems, providing increased agility, resilience, and improved decision-making [57, 92], *e.g.* COGNITWIN (Cognitive plants through proactive self-learning hybrid digital twins) [2, 6], FACTLOG (Energy-aware Factory Analytics for Process Industries) [51, 68], and QU4LITY (Digital Reality in Zero Defect Manufacturing) [88, 89]. From these projects, cognitive digital twins enabled a unified framework to orchestrate interactions among complex production systems and processes that involve multiple subsystems and stakeholders from different domains or lifecycle phases, while also providing solutions for industrial systems that require higher levels of agility, resilience, reconfigurability, enhanced decision making, and autonomous reaction capabilities [90].

Despite their potential, cognitive digital twins face substantial challenges. Knowledge management remains a critical hurdle, as automated representation, acquisition, and continuous updating are essential for sustaining real-time dynamic systems [23, 82, 90]. Although ontologies are present in the cognitive digital twins framework, they are not fully autonomous; external agents and their adjustments are not inherently represented in the KG, limiting its ability to 'self-evolve' dynamically with real-world changes. Additional challenges arise in integrating heterogeneous cognitive digital twins due to disparate standards and protocols, which complicates cross-lifecycle orchestration [23, 82, 90]. Standardisation across industries is also a significant challenge, with different organisations adopting varying frameworks, further complicating the implementation of cognitive digital twins with such inconsistencies [23, 82, 90]. Furthermore, human expertise is indispensable in situations where data alone may be insufficient and the AI-driven decisions remained a 'black box', which poses challenge for existing cognitive digital twins frameworks to emphasise human-in-the-loop interactions [57, 92]. This is especially crucial in addressing 'corner cases', *i.e.* scenarios that occur outside of the 'trained', normal, typical or expected range of inputs, outputs, or behaviours of a data structure in dynamic systems [57, 92].

The World Avatar (TWA) TWA (www.theworldavatar.io) aspires to advance beyond the existing digital twins frameworks discussed, *i.e.* not only enabling semantic descriptions and integrations of physical and digital components but also by creating a fully autonomous, self-evolving system that facilitates continuous information exchange across these domains. TWA seeks to establish a dynamic, self-contained environment where semantically-represented computational agents interact with interconnected elements in the KG to autonomously ingest, process, and respond to incoming data. This will enable seamless data propagation throughout linked components, integrating advanced cognitive functions such as goal derivation and adaptive behaviour.

One of the core ambitions of TWA is to implement a hierarchical goal cascade that translates broad directives into specific actionable subgoals, maintaining a structured response capability similar to the principles of Life 3.0 [74]. With the intrinsic structural awareness, TWA can then autonomously reconfigure itself in response to environmental changes, preserving its operational relevance over time.

What sets TWA apart from typical digital twins is its emphasis on dynamism and automation. This can be achieved with the development of semantic computational agents which will act as autonomous knowledge components that manage and update the instantiated data. Any changes introduced by the semantically represented computational agents are required to be formally ontologised, ensuring consistency within the system, *i.e.* a closed system. The inputs and outputs of these agents can be semantically annotated to form entire chains of dependent information.

A native provenance framework of KG ensures that updates and changes to any individual data element propagate automatically throughout KG, including updates to all dependent information [10]. This automated information cascading, together with continuously running input agents that assimilate latest real-world data into the system, allows TWA to remain up-to-date with and responsive to new information and scenarios, *i.e.* TWA's derived information framework [10]. This self-consistent, platform-independent architecture of TWA will allow for logical setup across various platforms, supporting both distributed and local access. In the interest of accessibility, TWA will facilitate different levels of data access, connecting with distributed servers while accommodating a range of access permissions, depending on authorisation levels.

To achieve its potential, TWA will need to address some technical challenges such as federated queries and data privacy concerns, which are currently under development for optimised security and functionality. The ultimate aim is for TWA to become a universally adaptable framework with open access points for secure information exchange, establishing a path for advanced autonomous, sustainable, and intelligent systems.

3 The World Avatar

TWA is a general open source dynamic semantic model of world knowledge that aims to seamlessly integrate data between domains based on a unified knowledge repository [4]. Its key themes — Connect, Query, Imagine, and Control — reflect its comprehensive capabilities. TWA connects disparate data sources to provide real-time insights, supports complex what-if scenarios, and bridges the gap between digital and physical domains through robust control functions. These capabilities are underpinned by three foundational technologies: the Semantic Web, Linked Data, and Ontologies, which collectively enable data to be machine-readable, interoperable, and reusable across applications, domains, and community boundaries [38].

While a comprehensive discussion of these technologies is beyond the scope of this work, a brief overview is provided here. The Semantic Web is the extension of the World Wide Web, enabling data to be machine readable, shared, and reused [14]. Linked Data involves the structuring of metadata, enabling interlinking between datasets to create a connected network of information. Ontologies provide formal representations of domain-specific knowledge, defining relationships between data elements and ensuring coherence when integrating information from diverse sources [34].

These three components enable the creation, management, and integration of diverse data sources, forming a cohesive network of entities and relationships that allows for cross-domain analyses and consistent, versatile insights. **Figure 3** illustrates an example of a KG structure for flood contexts, where interconnected nodes represent concepts and data instances, and edges define relationships [62]. Unlike typical digital twins, TWA emphasises automation and dynamism, with semantic computational agents managing and updating instantiated data. These agents facilitate autonomous information cascades, where changes to individual data elements are propagated automatically throughout the KG, ensuring TWA remains current and responsive to evolving conditions.

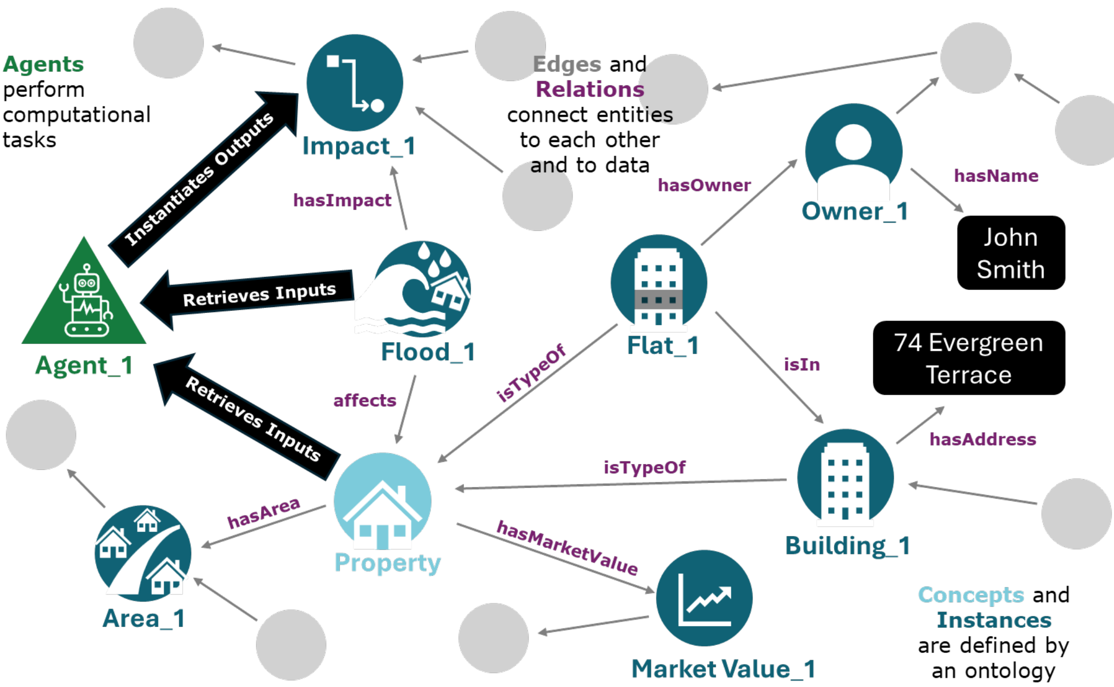


Figure 3: *KGs consist of nodes defining concepts and data instances, and edges denoting their respective relationships. The World Avatar dynamic KG, furthermore, includes ontologically represented computational agents as integral part of the graph, making it inherently dynamic [62].*

KGs are powerful tools that represent complex networks of entities and their interrelationships. In TWA, a dynamic KG serves as the foundation for integrating data from previously isolated sources, providing a unified view that can be queried and analysed. This interconnected structure allows for versatile yet consistent cross-domain analyses, as relationships and dependencies between different data elements are unambiguously defined.

Designed to be platform-agnostic and open-source, TWA eliminates the risk of vendor lock-in and promotes a collaborative and transparent development environment. Its scalable, decentralised architecture supports secure data access, enabling local hosting and role-based access control for privacy. TWA's flexible agents integrate easily with existing software, facilitating seamless integration of future technologies. This ensures a holistic and evolving worldview accessible to all stakeholders, supporting coordinated and informed decision-making processes.

The use of ontologies within TWA enables interoperability and supports a variety of visualisation and analysis tools in the short term. In the long term, TWA preserves domain expertise by explicitly modelling previously implicit knowledge and experience, thereby preserving domain expertise. By separating data and knowledge representation from technical implementation, TWA maintains compatibility with multiple storage solutions, such as triple stores and relational databases, improving both flexibility and scalability. This makes TWA an invaluable tool for modern urban planning and management.

In urban planning, TWA integrates and connects data from numerous domains, including geospatial data. This capability enables public administrations to overcome the interoperability challenges inherent in current systems, significantly enhancing the efficiency and responsiveness of urban planning processes. **Figure 4**, for example, illustrates how TWA consolidates interconnected tasks across Geographic Information Systems (GIS), Building Information Modelling (BIM), and Building Management Systems (BMS) into a cohesive, integrated framework [66, 67].

The following sections will provide detailed explorations of TWA's data and knowledge capabilities, interaction methods, and various use cases.

4 Data and knowledge capabilities

Based on underlying ontologies and semantic agents, TWA can represent detailed geospatial information and stream live data input from the real world. The integration of geospatial geometry data, including both 2D and 3D representations, allows for comprehensive visualisations and analyses of the built environment. **Figure 5**, for example, combines the visualisation of 3D buildings (coloured according to their primary type of use or any other characteristics, *e.g.* the value of the property market) with areal polygons of expected flooding, as well as various readings of environmental sensors (*i.e.* water level, weather, rainfall, air quality) [39]. For instance, CityGML (www.ogc.org/standard/citygml/) is a standard for representing, storing, and exchanging 3D city models, enabling integration of urban geodata for various applications like urban planning, BIM, and simulations. By incorporating and representing CityGML within TWA framework, complex building ge-

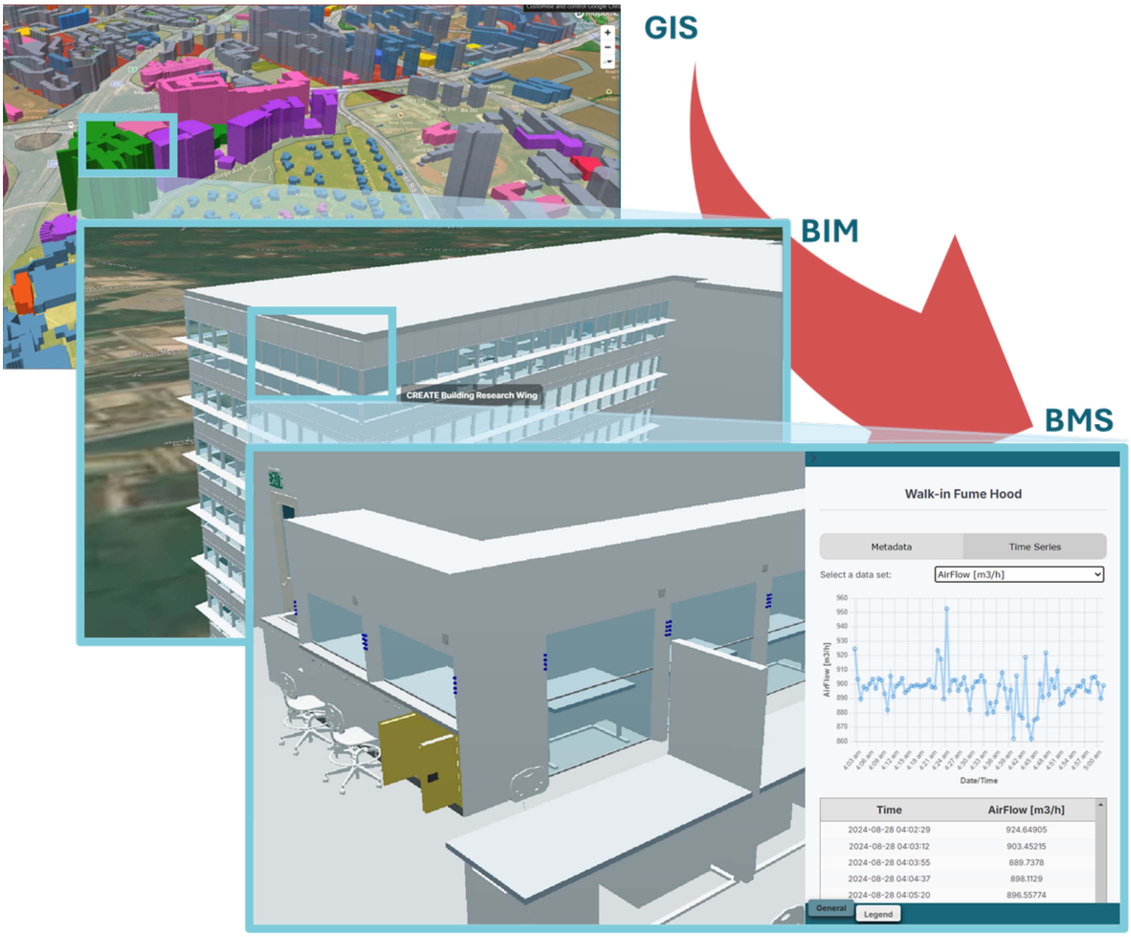


Figure 4: *Based on knowledge models, TWA allows to combine data of arbitrary spatial and temporal dimensions; shown for the combination of GIS and BIM data, further enriched with detailed information about individual devices in BIM [66, 67].*

ometries can be effectively represented through output agents of TWA. Beyond visualisation purposes, these geometries are used for further analyses and simulations, *e.g.* energy assessments and the accurate matching of buildings to their respective plots. While primarily focused on Mapbox, TWA also supports Cesium as mapping provider, offering flexibility in geospatial data visualisation and interaction.

TWA can ingest live-data streams through 'input agents' that continuously assimilate latest real-world data, ensuring the system remains current in time. Autonomous semantic TWA agents provide derived and calculated data, and can be invoked as necessary. Newly instantiated data can trigger automated updates to information with dependencies, allowing the system to model complex dynamic interdependencies between related information. This capability provides real-time insights into the current state of the world and the po-

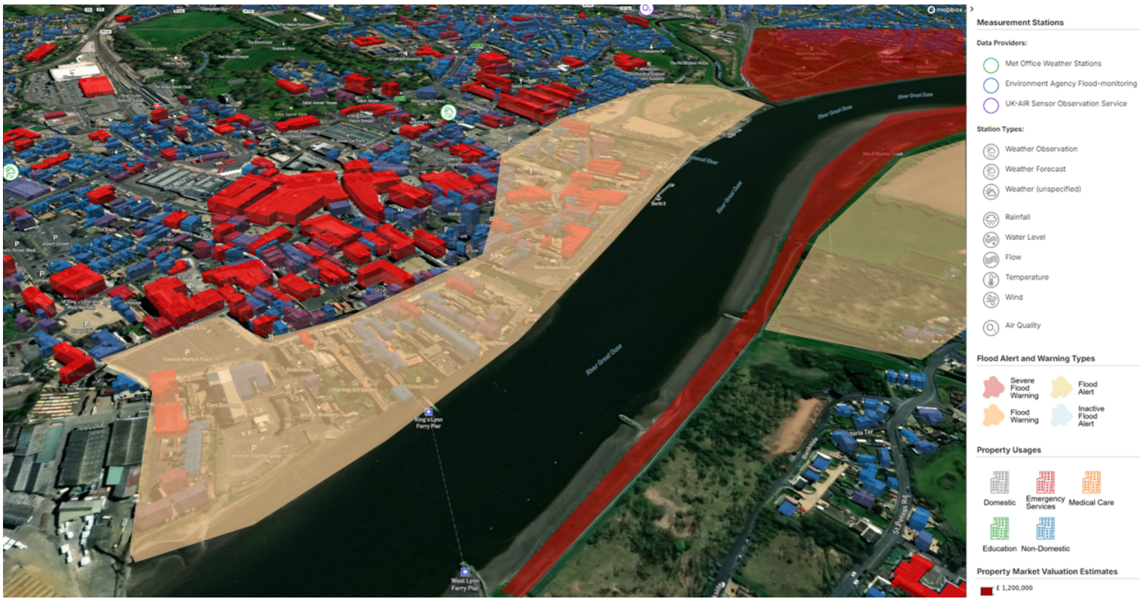


Figure 5: *The World Avatar's visualization capabilities extend to both 2D and 3D views, encompassing buildings, land plots, streets, and dynamic data [39].*

tential consequences of newly published data. This has been demonstrated through two case studies: one focusing on the dynamic assessment of flood hazards in the UK [39], and the other on the assessment of air pollution dispersion associated with district heat generation in Germany (see **Figure 6**) [40].

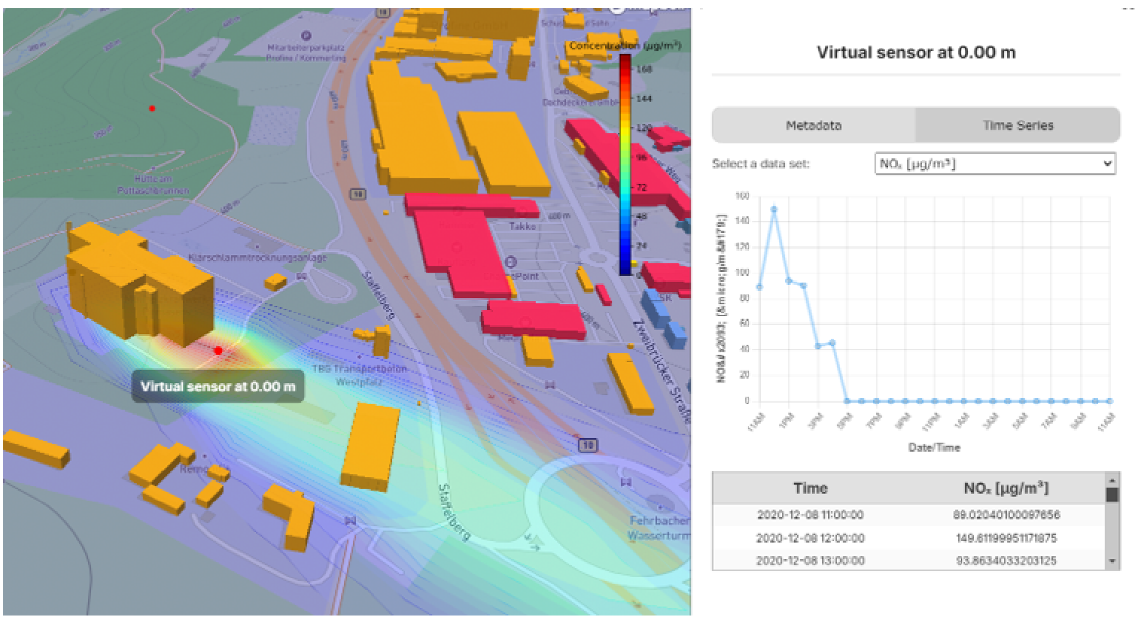


Figure 6: *Simulated air pollution dispersion associated with a certain district heat generation scenario. Based on a variety of available heat sources, TWA can derive the cost-optimal generator dispatch strategy and simulate the emission time series and their dispersion for associated air pollutants.*

TWA provides extensible computational capabilities through the seamless integration of versatile simulation tools. By incorporating around existing modelling suites, *e.g.* the City Energy Analyst (CEA, www.cityenergyanalyst.com) or AMS/EPA Regulatory Model (AERMOD, www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models), established models can be made available semantically (energy or emission dispersion, respectively). These semantic agents can directly interact with the underlying KG, allowing them to replace default assumptions built into top-down software tools with actual instantiated data, and thereby improving accuracy. For instance, as shown in **Figure 7**, the CEA agent utilises building-specific construction characteristics, along with local weather and terrain data, to assess a building's energy demand and renewable generation potential [77]. This approach contrasts with the default reliance on generic assumptions in the native CEA toolkit, leading to more meaningful results.

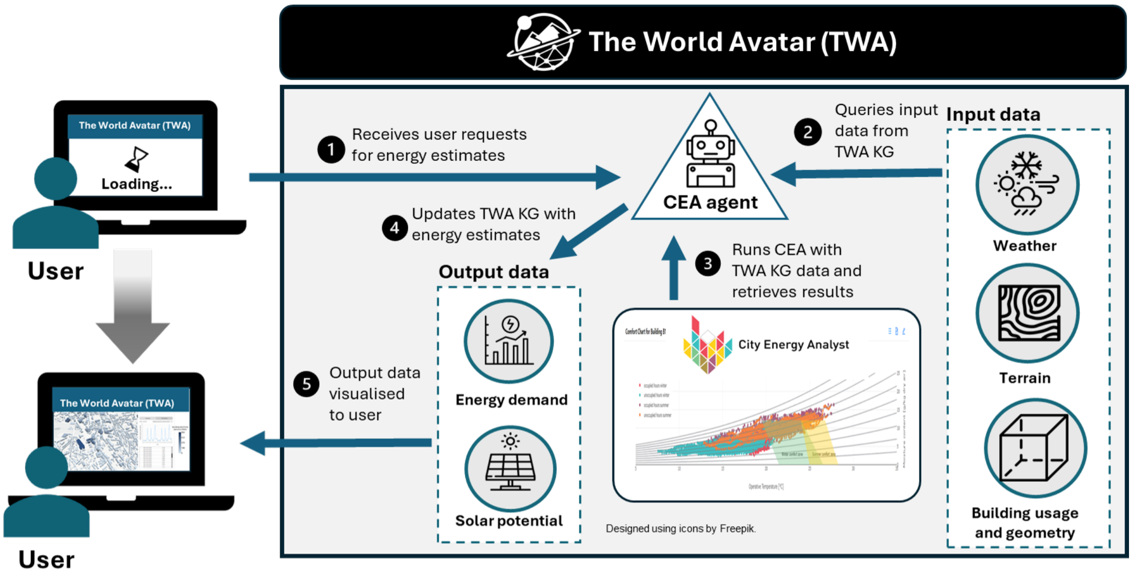


Figure 7: *Schematic depiction of the City Energy Analyst agent workflow, outlining key interactions with both users and the underlying KG [77].*

Agents can be used for "infilling" by simulating missing data required for various analyses across different domains of interest. TWA ensures aligned knowledge and data representation based on shared ontologies as well as comprehensive provenance tracking, enabling consistent analyses across all stakeholders. This is particularly beneficial for various departments within a city or state administration, as it supports role-based access rights, ensuring that only authorised personnel can access sensitive information.

5 Natural language interface

Traditionally, interacting with complex machine systems has required specialised knowledge of protocols and data query languages, creating a steep learning curve that often deters non-technical users from fully utilising available resources. To address this, TWA offers multiple user-friendly interaction channels, including dashboards for time-series and live data visualisation, mobile applications [66, 67], augmented reality goggles [66, 67], and a chatbot interface [60, 76, 91] powered by large language models (LLMs). These interfaces are supported by TWA output agents, which bridge the gap between human users and the system, ensuring accessibility for a wide range of audiences with varying preferences and needs.

Among these interaction methods, the natural language interface stands out as a particularly attractive option. By combining the capabilities of LLMs [47] with real-time data from TWA's extensive KG, the interface provides users with reliable, fact-oriented responses. This integration enables users to interact with TWA using everyday language, democratising access to TWA and making its extensive and information-rich KG accessible to a broader audience.

The core strength of this approach lies in its ability to understand and interpret user intents accurately by tapping on the capabilities of LLMs. When a user poses a question or makes a request, the natural language interface comprehends this input and translates it into precise machine-readable queries. These queries are then processed by TWA, which retrieves relevant data or computes derived information and presents it back to the user in a clear and concise manner. This seamless interaction enhances user experience and ensures that responses are based on the most current and accurate data available [60, 76, 91].

One of the notable advantages of our system is the mitigation of hallucinations, a common issue observed in stand-alone LLMs, *e.g.* ChatGPT (www.chatgpt.com). Hallucination refers to the generation of incorrect or fabricated information, often due to the lack of real-time data access or contextual understanding. By leveraging TWA's real-time data streams and verified KG, our system significantly reduces the likelihood of hallucination. Users can trust that the responses they receive are grounded in fact and reflect the latest available information.

In addition to text-based outputs, TWA caters to the diverse needs of GIS and BIM professionals by offering multi-modal responses. Depending on the nature of the query, TWA can generate tables, charts, annotated maps, and other visual aids that enhance comprehension and decision-making (see **Figure 8** from www.theworldavatar.io/demos/zaha). This versatility ensures that information is not only accurate but also tailored to the specific requirements of surveyors, geo-informatics professionals, and other stakeholders.

6 Use cases

This section highlights several use cases relevant to urban planning, geoinformatics, and surveying. While these examples are tailored to the scope of this journal, they represent only a subset of the broader applications enabled by TWA. Beyond urban contexts, the TWA framework is designed to support a fully connected and automated augmented reality that is generic and adaptable across diverse domains, including chemistry, materials science, and laboratory automation [11, 43, 60, 66, 67]. Collectively, these use cases serve as proof-of-concept for TWA's flexibility and its capability to address complex, interdisciplinary challenges. Additionally, they represent intermediary steps toward realising a comprehensive augmented reality framework for designing sustainable and resilient cities.

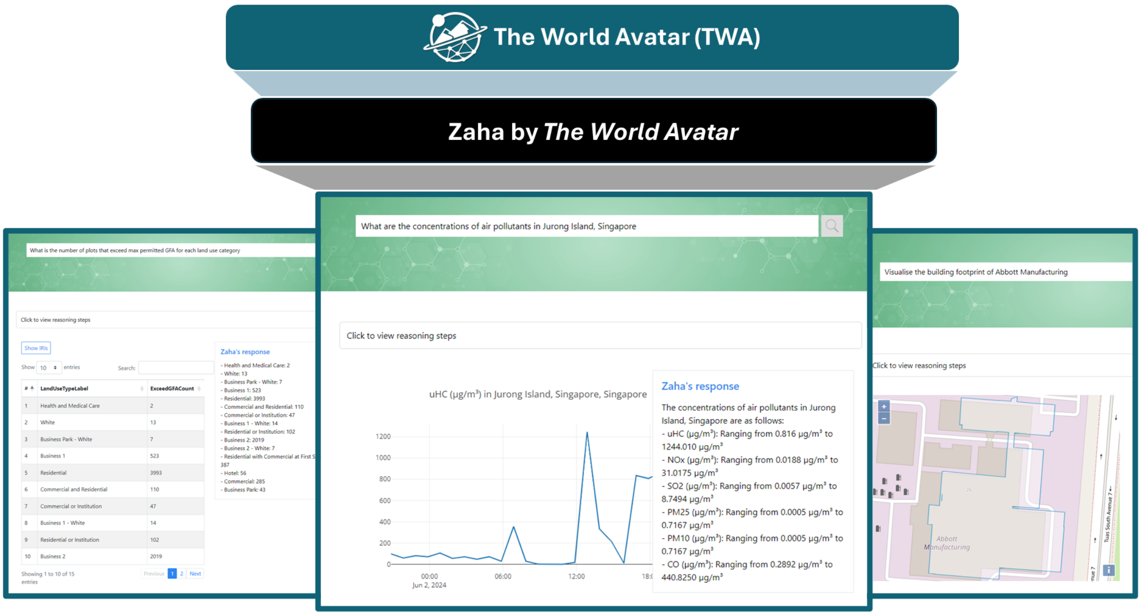


Figure 8: *The World Avatar's chatbot interface supports a wide range of questions and response types, including text, tables, graphs, and map views – all directly queried from the underlying dynamic KG from Zaha by TWA from www.theworldavatar.io/demos/zaha.*

6.1 Resolving the boundary between GIS and BIM

With its underlying knowledge models, TWA can connect GIS and BIM representations and resolve potential ambiguity between both domains. This capability facilitates "borderless" navigation from the city scale down to individual buildings, as well as installed devices, and beyond. Additionally, it allows for the seamless integration of BMS data, ensuring a cohesive and comprehensive approach to managing and analysing spatial and building information enabling fault prediction and predictive maintenance. Merging these overlapping, yet currently isolated, domains, *i.e.* GIS (spatial analysis), BIM (building life cycle management), and BMS (optimise building operations), TWA enables a more comprehensive perspective for holistic analyses. For instance, as depicted in **Figure 9**, combining GIS's spatial analysis with BIM's detailed building models and BMS's operational data can facilitate more accurate simulations of energy usage and urban planning scenarios [40].

6.2 Programmatic plot finder

City administrations face significant challenges in managing land use functions such as planning, infrastructure development, permit issuance, and compliance monitoring. These processes often require civil servants to manually search for target land plots by navigating through various PDFs, online documents, and GIS layers to verify regulatory compliance. Additionally, the need to coordinate with other departments for approvals and access to relevant documents adds further delays. The inherent complexity of land use regulation,

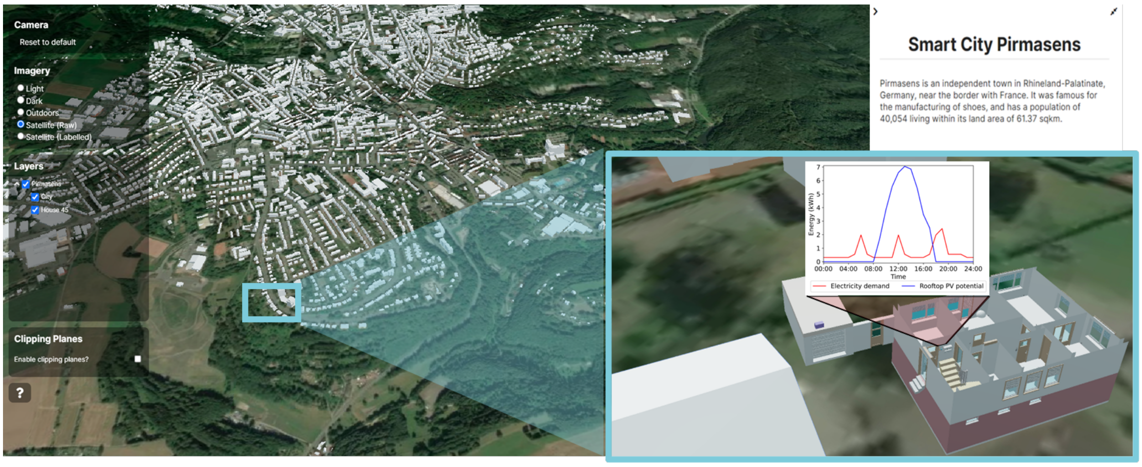
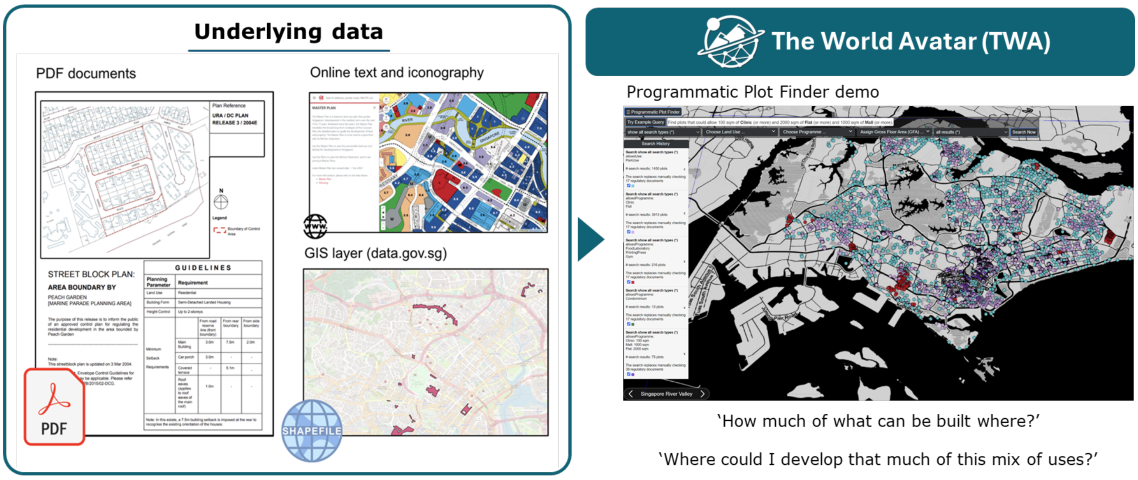


Figure 9: *Bridging the boundaries between legacy GIS, BIM, and even BMS applications. Connecting various representations within one single system supports "borderless" and unambiguous navigation from city scale down to individual device characteristics [40].*

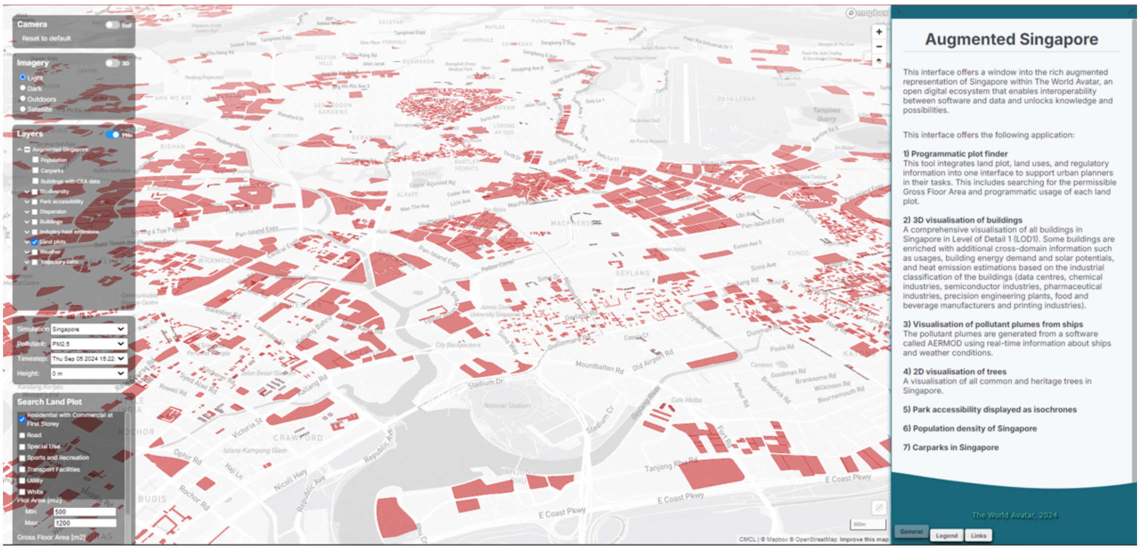
driven by diverse designations for recreational, residential, economic, and infrastructure purposes, compounds these issues. Each designation entails unique regulations tailored to the priorities of stakeholders, such as enhancing quality of life for residential areas or ensuring safety for utility plots. Moreover, regulations governing adjacent land uses can influence those of the target plot, further increasing complexity. This cumbersome, opaque, and manpower-intensive workflow creates frustration among stakeholders and places a heavy administrative burden on city governments.

To address these challenges, TWA introduces two key innovations. First, TWA ingests and integrates regulatory documents and GIS data into a dynamic land plot search engine, enabling users to search for plots based on specific criteria. Unlike conventional plot search engines, which are typically limited to residential and commercial uses, TWA is designed to meet the broader needs of city administrations, *e.g.* regulatory compliance, infrastructure management, real-time data integration, public engagement, disaster preparedness, future scalability, *etc.* This implementation significantly improves efficiency, potentially replacing over one million manual verifications of regulatory documents annually [72]. **Figure 10(a)** depicts the integration of data and knowledge into a target network of connected constraints, and Figure 10(b) showcasing an example search result through the search interface by TWA for residential plots with commercial use on the first storey, spanning areas of 500 to 1200 square meters.

The second innovation involves the representation of institutional knowledge in a comprehensive, extensible knowledge model that integrates geospatial and temporal dimensions. Using formal ontologies, TWA semantically represents and connects land use regulations with other domains, creating a robust compliance-checking framework. This approach supports dynamic and automated procedures that go beyond simple plot searching. For example, by linking 3D building data with land plot and regulatory information, the semantic agents of TWA can autonomously verify whether existing buildings comply with permissible gross floor area requirements for specific plots. The dynamic nature of these



(a) Schematic depiction of connecting underlying data.



(b) Search results for residential with commercial at first storey plots of 500 to 1200 sqm.

Figure 10: *The World Avatar's land plot search engine.*

agents ensures that updates to regulations, building specifications, or plot data automatically trigger workflows and propagate changes to web visualisations. This capability allows compliance procedures to remain accurate and current, reflecting regulatory changes in real time.

These advancements position TWA in a crucial role in the transformation of urban planning and land use management. Its novel methods align with similar approaches found in the literature [33, 71, 79], further validating its effectiveness in addressing the complexities of modern land use regulation. By reducing administrative burdens and improving transparency, TWA supports more efficient and integrated decision-making processes for city administrations.

6.3 Strategic placement of infrastructure

City planners strive to ensure that their cities provide sufficient and equitable access to relevant infrastructure and amenities for all residents. TWA can support these tasks by assessing the coverage by amenities using isochrones of various time frames and transport modes (*i.e.* walking, cycling, driving) [61]. By comparing various placement alternatives for new infrastructure regarding population demographics distribution, existing amenities coverage, and land use regulations, TWA can streamline the planning process of urban planners. This involves identifying areas with poor accessibility and validating improved coverage due to new infrastructure placements, as illustrated in **Figure 11**, in which the accessibility to pharmacies is used as an example. The ability to measure and optimise the coverage of amenities is invaluable for applications such as planning emergency services. By determining the optimal locations for critical infrastructure (*e.g.* fire stations, police stations, emergency medical services), planners can minimise response times and enhance public safety.

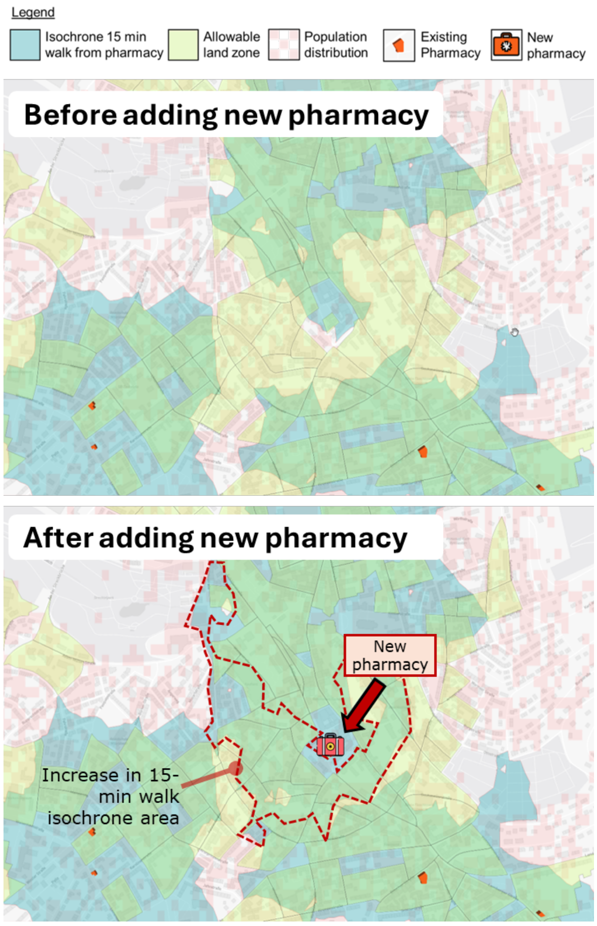


Figure 11: *Infrastructure planning through the multi-factor considerations where red grids represent population distribution, yellow polygons represent allowable zones for pharmacies to be built, blue represents 15-minute walk isochrone from pharmacies [61].*

Moreover, integrating population demographic attributes (*e.g.* age, gender) enables planners to assess the effectiveness of emergency services in reaching specific groups, such as the elderly and children. This is particularly useful for applications like placing mobile vaccination units or calculating the number of elderly individuals served by each pharmacy. In addition to healthcare services and emergency services, TWA also assesses the accessibility coverage of other essential amenities, including educational facilities, retail shops, recreational facilities, banks, *etc.* This can be used to improve the development of a "15-minute city", where residents can meet their daily needs (*e.g.* work, home, food, health, education, culture, sports, and leisure) within 15 minutes of walking or cycling from their residence.

6.4 Climate Resilience Demonstrator

The role of TWA in enabling interconnected, data-driven analyses opens new possibilities for addressing complex, interdisciplinary challenges. The Climate Resilience Demonstrator (CReDo) serves as a prime example of how TWA can be leveraged to enhance the resilience of critical network infrastructure systems. By utilising a system-of-systems approach in TWA, vulnerabilities in the systems can be identified through the prediction of potential cascading failures across interconnected networks, such as energy, water, and telecommunications. Extensive ontologies are the key to ensure data consistency across the diverse domains and integrating data from multiple stakeholders in TWA. With high-quality input data, and high-fidelity real-time data maintained in TWA, this holistic perspective enables stakeholders to optimise system resilience at minimal cost, moving beyond unaffordable and often inadequate infrastructure hardening strategies [1, 37].

CReDo integrates data on asset types, modelled operational states, locations, and their logical connections (see **Figure 12**). The data is ontologised to enable interoperability between the infrastructure details with flood simulation results under various climate change scenarios. Using simulated flood maps based on historical and projected data, CReDo provides insights into network interdependencies and failure propagation. For example, the platform models how flood-induced failures can propagate both within and across networks, offering both tactical and strategic planning dimensions. This dual perspective enables the simultaneous management of immediate risks and long-term preparedness.

While the implementation in CReDo focuses on the UK, the extensibility of TWA allows for broader applications. For instance, TWA could support sea-level rise vulnerability assessments in regions like Singapore, incorporating diverse factors such as infrastructural, economic, cultural, and planning constraints. By combining geospatial and non-geospatial data into a unified representation, TWA facilitates integrated analyses that account for land use regulations, population distributions, and existing infrastructure.

With TWA, urban planners can evaluate multiple conflicting constraints, avoiding high-value developments in vulnerable areas while adhering to regulatory and land-use policies. **Figure 13** illustrates how combining domains such as population distribution, designated land use, building types, and sea-level rise vulnerability zones enhances multi-perspective decision-making. This approach with TWA helps minimise risks while maximising the effectiveness of urban development plans.

Figure 12: *The World Avatar enables strategic planning and scenario analyses to understand potential failure propagation of network infrastructure due to flooding, i.e., whether and how a potential failure could cascade (1) across networks and (2) out of the geographic scope of the flooding (in red). Screenshot taken from CReDO.*

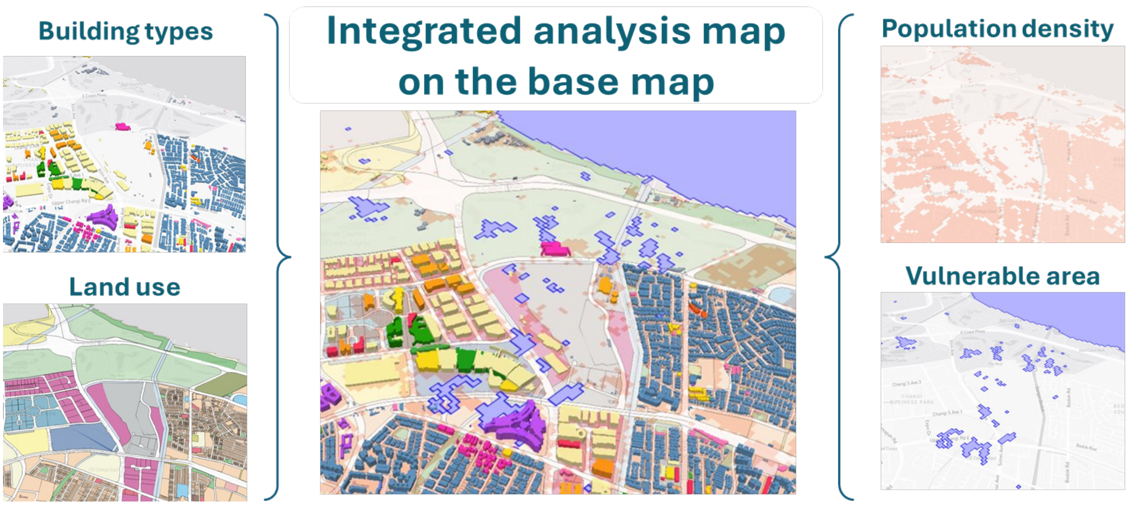


Figure 13: *The combination of population distribution, designated land use, building types, vulnerable area from sea level rise enables a multi-perspective visualisation, enhancing integrated analysis.*

6.5 Optimise evacuation routing

In the event of disaster, such as a flood, specialised vehicles such as helicopters, boats, ambulances, fire trucks, and high-water trucks are limited resources. However, rapid, efficient, and specialised assistance is crucial, especially for aging and vulnerable demographics. Loss of local medical services combined with restricted accessibility to people at risk are factors that will drastically increase the mortality rate. Hence, advanced preparedness and anticipatory actions are required to provide a quicker and more effective response during a sudden disaster.

One way to achieve this is by integrating flood depth level dynamically into routing calculations to account for distinct operational speeds, allowable water wading depth, and deployment times. TWA supports this approach by semantically connecting all necessary data, enabling the effective coordination and allocation of vehicles based on rescue locations and current flooding conditions. **Figure 14** illustrates this approach, showcasing the selection of various routes circumventing flooded areas for vehicles with three different wading depths, leading to a swift and efficient emergency response [61].

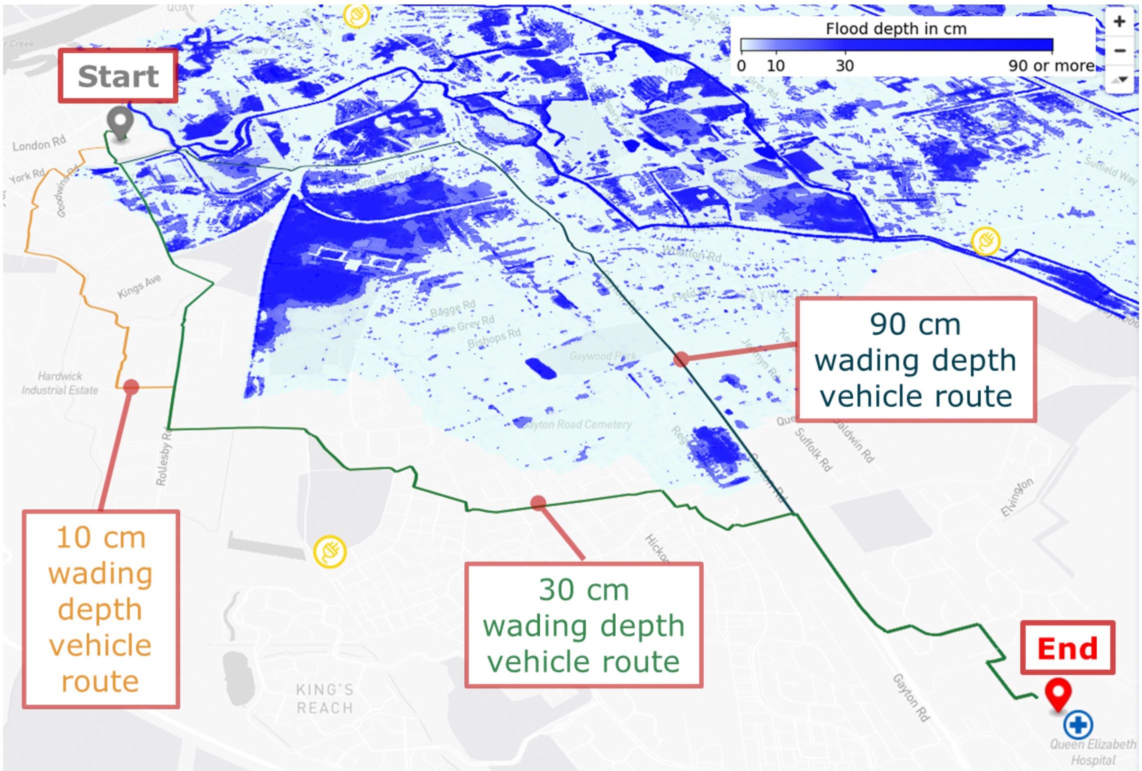


Figure 14: *Optimal route selection under flooded conditions where orange, green, and purple lines represent the fastest paths taken by vehicles with 10cm, 30cm, and 90cm wading depth capabilities, respectively [61].*

6.6 Support municipal heat planning

The City Energy Analyst (CEA) agent has been developed within TWA to offer versatile energy simulation capabilities, enabling precise assessment of current energy demands and estimations of renewable energy potential required to developed future energy planning scenarios. By integrating the CEA simulation engine into TWA, previous generic built-in assumptions can be replaced with actual and up-to-date building properties as well as environmental conditions to enhance simulation accuracy [77].

In 2023, the German Bundestag passed the Heat Planning Act that makes municipal heat planning mandatory and aims at climate-neutral heat supply by 2045 [36]. There are four phases to the municipal heat planning: inventory analysis, potential analysis, development of heating plans, and development of an implementation strategy. Granular heat density and renewable generation potential maps are essential for the analysis of municipal heat planning; however, publicly available datasets tend to be rather coarse or might even lack relevant information. The CEA agent can provide both granular, building-level heating demand and solar potential estimates to complement potentially missing actual data for required analyses.

For a mid-sized town in Germany, the CEA agent has been demonstrated to be applicable in supporting municipal heating planning, especially from inventory analysis to the development of heating plans [77]. For both the inventory and potential analyses, the CEA agent provides granular heat and potential maps for the quick identification of areas of high heating demand or high solar energy potential (see **Figure 15**). The data provided by the CEA agent on the building heating demand and the heat generation from solar collectors are also available as time series, and can be aggregated across different spatial and time scales. The aggregation capability allows policy makers to better understand the variation of heating demand or solar potential across different areas such as districts or postcode zones.

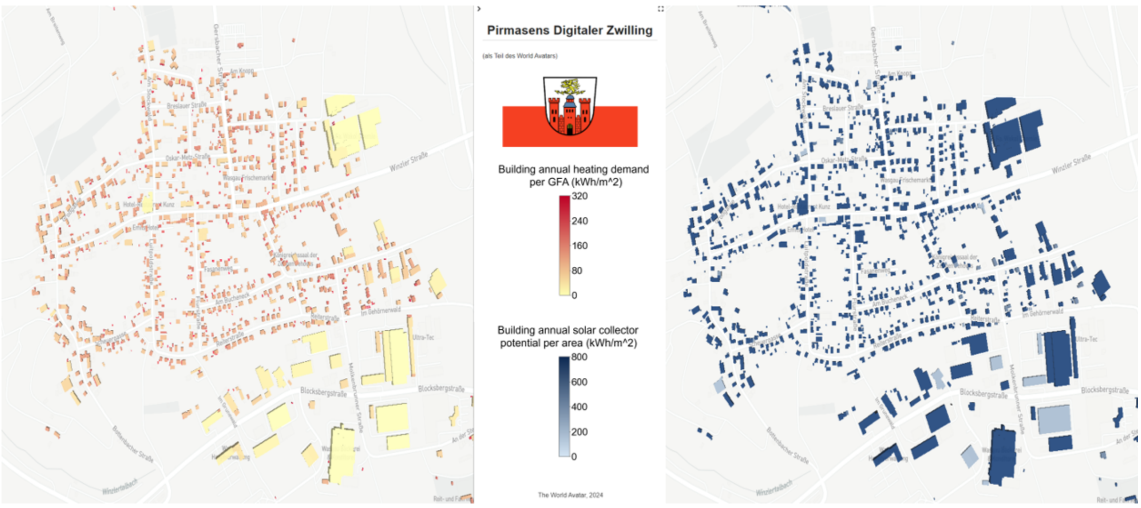


Figure 15: *Building-resolved heating demand and solar generation potential simulated by the City Energy Analyst agent, leverage detailed building geometries as well as local weather and terrain data [77].*

For the development of heating plans, the data provided by the CEA agent enables the planning of installing on-site solar collectors to offset heating demand. We analysed the installation of solar collectors with and without thermal heat storage and derived the net present value (NPV) of the two scenarios. We found that around 69 percent of the buildings investigated in the city has positive NPV after 25 years (typical lifetime of the investigated solar systems). Based on the NPV results, and the heat density map, we were able to recommend the central area of the city for the prioritised installation of solar collectors due to the high density of buildings with high NPV and high heating demand in that area.

6.7 Analysis of social inequality from heat pumps

In scenarios such as evaluating the socioeconomic impact of large-scale adoption of heat pumps in the domestic heating sector of the UK, TWA has also been useful in linking data from different agencies to gain insight [85]. As illustrated in **Figure 16**, with TWA enabled integrated geospatial data analysis, detailed analysis of how household fuel consumption, fuel poverty, climate data, and future energy prices affect regional inequalities can be performed. This is crucial to understanding how different regions respond to decarbonisation efforts.

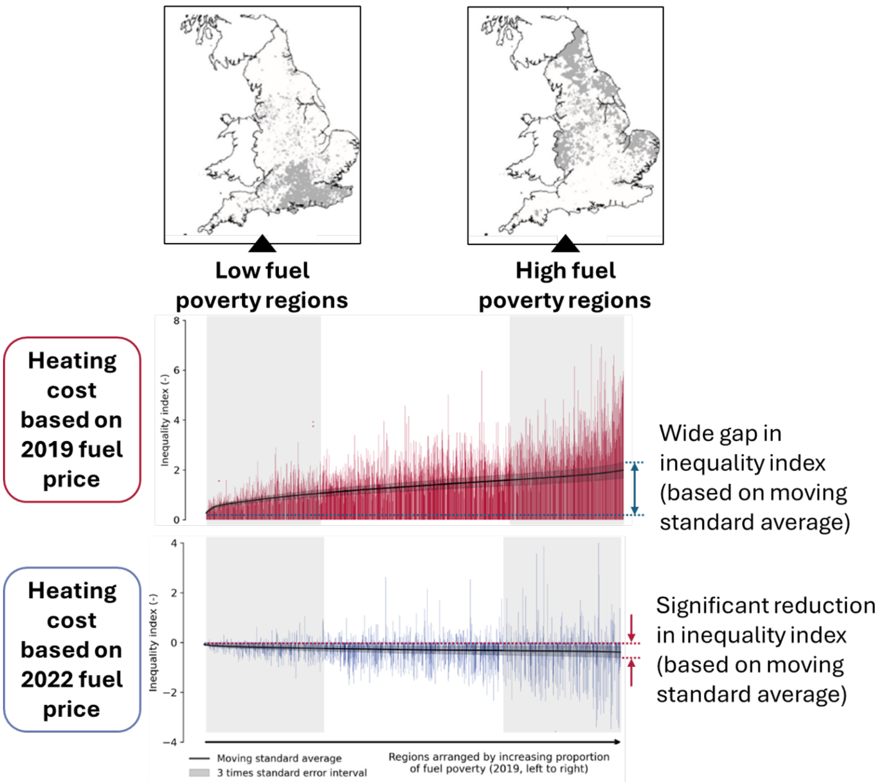


Figure 16: *Inequality index for electricity and gas at 2019 and 2022 prices. Each bar represents a single Lower Layer Super Output Areas region. Regions are arranged in the order of increasing proportion of fuel poverty (2019, left to right). The maps (top) show the 10,000 regions with the lowest (left) and highest (right) proportions of fuel poverty (2019). Figure adapted from [85].*

The adoption of heat pumps could have varied results depending on the electricity-to-gas price ratio. For example, under the 2019 fuel prices, regions with high fuel poverty, particularly in northern England, show increased inequality, while the 2022 prices reduce this inequality. The analysis emphasises the sensitivity of inequality to the price ratio, illustrating how TWA can provide critical insights into the unintended social consequences of decarbonisation policies, which would be difficult to capture with traditional models. This comprehensive data integration enables policymakers to predict future trends and adjust strategies accordingly.

Further inspection of the results from Figure 16 revealed that the inequality index responded to the socioeconomic vulnerability of regions to price changes, with northern regions, where fuel poverty is more prevalent, being disproportionately affected. The dynamic modelling capabilities of TWA allow for scenario analyses that highlight potential risks, such as how a return to the 2019 price levels could exacerbate inequality. Conversely, the 2022 price scenario indicates that under the right conditions, heat pump adoption can help reduce both carbon emissions and inequality. Policymakers can leverage these insights to formulate financial support mechanisms, such as subsidies targeted at regions with high fuel poverty, to mitigate the adverse effects of fluctuating fuel prices. The data-driven decision-making enabled by TWA ensures that decarbonisation efforts are aligned with social equity goals, addressing both environmental sustainability and regional disparities.

6.8 Power system decarbonisation with small modular reactor deployment

TWA's dynamic and interoperable structure is designed to be reusable across various sectors, enabling more informed decision-making in clean energy transitions. In the current case study, TWA facilitated the integration and modelling of large, complex datasets relevant to energy systems (*e.g.* infrastructure data, power system models, and socio-demographic attributes), especially in efforts to assess clean energy trajectories and enable the smooth transition towards decarbonised power system.

One notable application of this approach is outlined in a study on decarbonising the UK power system [84]. TWA populated with power plant data and administrative region details, allows computational agents to automate tasks like data processing, scenario analysis, and simulation. This has enabled the exploration of how different deployment strategies for Small Modular Reactors (SMRs) could facilitate the UK's energy transition. The computational agents not only simulated the effects of SMRs but also performed geospatial queries to assess their proximity to population centers and overall network efficiency [83, 84].

Another key study explored the role of SMRs in the UK's energy transition under different carbon tax scenarios, focusing on minimising transmission losses and population risk [83]. The use of SMRs presents an opportunity to complement renewable energy sources like wind and solar, offering consistent base-load power while reducing greenhouse gas emissions. Analysis using TWA highlights the importance of optimising SMR placement to strike a balance between cost, risk, and energy efficiency, enabling the design of effective clean energy policies that align with national decarbonisation goals, as shown in **Figure 17** [83, 84].

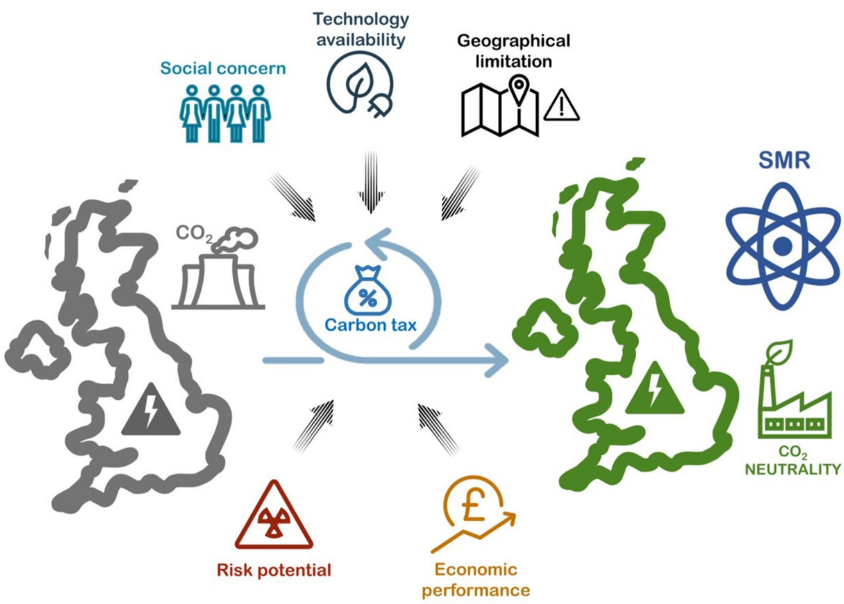


Figure 17: *Balancing between cost, risk, and energy efficiency, enabling the design of effective clean energy policies. Figure reprinted from Xie et al. [83].*

This integration of TWA's real-time data processing and dynamic modelling capabilities marks a significant step forward in achieving a sustainable, decarbonised power grid. The use of KG facilitates a multi-dimensional understanding of complex power systems, ensuring that solutions like SMRs are deployed optimally within future low-carbon energy networks.

7 Discussion and future work

TWA represents a significant advancement in integrating dynamic, multi-domain data to address complex challenges in urban systems. Through the demonstration of typical GIS use cases, this paper highlights the capabilities of TWA to enable fine-grained, real-time analyses of urban phenomena. However, it is important to emphasise that these use cases are not isolated applications; rather, they are integral components of a larger distributed TWA system. TWA functions as a unified and interconnected entity, but its distributed nature ensures scalability and adaptability across various contexts and domains.

The focus on GIS use cases in this paper reflects the target readership, which is predominantly within the GIS community. These examples demonstrate the capacity of TWA to address critical challenges in urban planning and management, such as integrating data from BMS and other fine-grained city-level representations. However, the potential applications of TWA extend far beyond GIS, with aggregated values and connections becoming increasingly significant as the framework evolves. Establishing these connections will rely on a detailed, fine-grained representation of cities, ensuring that diverse domains, such as infrastructure, energy systems, and social data, are seamlessly integrated.

At its core, TWA operates at the logical level, abstracting complexities related to hardware and software dependencies. Although this paper demonstrates specific technical implementations using standard technologies, such as Docker stacks, these are merely tools to support the logical framework. The ultimate aim is to achieve hardware and software independence, allowing users to interact with TWA without having to consider underlying technical requirements. This abstraction is essential for accessibility and scalability, enabling broader adoption across industries and disciplines. The ontologisation of hardware and software dependencies further strengthens this vision, ensuring that technical implementation details are modular and adaptable.

Do note that this paper focusses on demonstrating the foundational capabilities of TWA, the ongoing development of TWA is geared toward simplification the engagement and usage of TWA for different domains. Future work will focus on expanding the use cases, improving interoperability, and addressing challenges related to data security, federated queries, and dynamic updates. Ultimately, TWA seeks to enable a comprehensive, robust, and user-friendly means of modelling, analysing, and optimising interconnected systems, with far-reaching implications for public health, urban sustainability, resilience, and innovation.

8 Conclusion

The World Avatar (TWA) represents a significant step forward in the building of a world model, offering a dynamic and scalable framework that integrates data across multiple domains. Using knowledge graphs to unify GIS, BIM, and other environmental models, TWA enables more comprehensive urban planning and management. Its ability to seamlessly connect disparate data sources and provide real-time analysis empowers stakeholders to make informed decisions on critical issues such as climate resilience, energy optimisation, and infrastructure placement.

One of TWA's core strengths lies in its open-source, platform-agnostic design, which ensures accessibility for a wide range of users, from local authorities to research institutions. The modularity and flexibility of the system allow the integration of various data sources, fostering collaboration between sectors, and enabling cities of all sizes to adopt cutting-edge planning tools without prohibitive costs. The use cases provided in the current paper such as municipal heat planning and small modular reactor deployment, demonstrate the capacity of TWA in addressing complex challenges in energy, sustainability, and social equity.

As development continues, challenges remain in refining user interfaces, enhancing data security, and managing the complexity of integrating large-scale data systems. However, the potential of TWA to transform urban management is clear, with ongoing collaborations offering valuable insights to refine the platform. By continuing to engage stakeholders and expand its capabilities, TWA is well-positioned to become an essential tool for building smarter, more resilient cities in the face of emerging global challenges.

Acknowledgements

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. M. Hofmeister acknowledges financial support provided by the Cambridge Trust and CMCL. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Nomenclature

AERMOD AMS/EPA Regulatory Model

AI Artificial Intelligence

API Application Programming Interface

BIM Building Information Modelling

BMS Building Management System

CEA City Energy Analyst

CReDo The Climate Resilience Demonstrator

GIS Geographic Information System

KG Knowledge Graph

LLM Large Language Model

NLP Natural Language Processing

TWA The World Avatar

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to enhance the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Conflicts of interest

There are no conflicts to declare.

References

[1] Digital Twin Hub, Climate Resilience Demonstrator, 2022. URL `digitaltwinhub.co.uk/credo/credo/`. Accessed October 1, 2024.

[2] S. Abburu, A. J. Berre, M. Jacoby, D. Roman, L. Stojanovic, and N. Stojanovic. COGNITWIN – Hybrid and Cognitive Digital Twins for the Process Industry. In *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 1–8, 2020. doi:10.1109/ICE/ITMC49519.2020.9198403.

[3] A. E. Adl. The Cognitive Digital Twin, 2016. URL `www.slideshare.net/aeladl/the-cognitive-digital-twin`. Accessed October 1, 2024.

[4] J. Akroyd, S. Mosbach, A. Bhave, and M. Kraft. Universal Digital Twin - A Dynamic Knowledge Graph. *Data-Centric Eng.*, 2:e14, 2021. doi:10.1017/dce.2021.10.

[5] M. A. Al Faruque, D. Muthirayan, S.-Y. Yu, and P. P. Khargonekar. Cognitive Digital Twin for Manufacturing Systems. In *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 440–445, 2021. doi:10.23919/DATE51398.2021.9474166.

[6] Ö. Albayrak and P. Ünal. Smart Steel Pipe Production Plant via Cognitive Digital Twins: A Case Study on Digitalization of Spiral Welded Pipe Machinery. In V. Colla and C. Pietrosanti, editors, *Impact and Opportunities of Artificial Intelligence Techniques in the Steel Industry*, pages 132–143, Cham, 2021. Springer International Publishing.

[7] A. Alhamwi, W. Medjroubi, T. Vogt, and C. Agert. FlexiGIS: an Open Source GIS-based Platform for the Optimisation of Flexibility Options in Urban Energy Systems. *Energy Procedia*, 152:941–946, 2018. doi:10.1016/j.egypro.2018.09.097.

[8] B. D. Allen. Digital Twins and Living Models at NASA, 2001. URL `https://ntrs.nasa.gov/citations/20210023699`. Accessed October 1, 2024.

[9] A. Aryani, M. Fenner, P. Manghi, A. Mannocci, and M. Stocker. Open Science Graphs Must Interoperate! In L. Bellatreche, M. Bieliková, O. Boussaïd, B. Catania, J. Darmont, E. Demidova, F. Duchateau, M. Hall, T. Merčun, B. Novikov, C. Papatheodorou, T. Risse, O. Romero, L. Sautot, G. Talens, R. Wrembel, and M. Žumer, editors, *ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium*, pages 195–206, Cham, 2020. Springer International Publishing. ISBN 978-3-030-55814-7.

[10] J. Bai, K. F. Lee, M. Hofmeister, S. Mosbach, J. Akroyd, and M. Kraft. A Derived Information Framework for a Dynamic Knowledge Graph and its Application to Smart Cities. *Future Gener. Comput. Syst.*, 152:112–126, 2024. doi:10.1016/j.future.2023.10.008.

[11] J. Bai, S. Mosbach, C. J. Taylor, D. Karan, K. F. Lee, S. D. Rihm, J. Akroyd, A. A. Lapkin, and M. Kraft. A Dynamic Knowledge Graph Approach to Distributed Self-Driving Laboratories. *Nat. Commun.*, 15(462):1, 2024. doi:10.1038/s41467-023-44599-9.

[12] Q. Bao, G. Zhao, Y. Yu, S. Dai, and W. Wang. Ontology-Based Modeling of Part Digital Twin Oriented to Assembly. *Proc. Inst. Mech. Eng., Part B: J. Eng. Manufact.*, 236(1–2):16–28, 2022. doi:10.1177/0954405420941160.

[13] B. R. Barricelli, E. Casiraghi, and D. Fogli. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access*, 7:167653–167671, 2019. doi:10.1109/ACCESS.2019.2953499.

[14] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):28–37, 2001.

[15] S. Bibri. The Anatomy of the Data-Driven Smart Sustainable City: Instrumentation, Datafication, Computerization and Related Applications. *J. Big Data*, 6, 2020. doi:10.1186/s40537-019-0221-4.

[16] G. Bowman, T. Huber, and V. Burg. Linking Solar and Biomass Resources to Generate Renewable Energy: Can We Find Local Complementarities in the Agricultural Setting? *Energies*, 16(3), 2023. doi:10.3390/en16031486.

[17] S. Camarella, M. P. Conway, K. Goering, and M. Huntington. Digital Twins: The Next Frontier of Factory Optimization, 2024. URL `www.mckinsey.com/capabilities/operations/our-insights/digital-twins-the-next-frontier-of-factory-optimization`. Accessed October 1, 2024.

[18] H. M. Carlin, P. A. Goodall, R. I. M. Young, and A. A. West. An Interactive Framework to Support Decision-Making for Digital Twin Design. *J. Ind. Inf. Integr.*, 41: 100639, 2024. doi:10.1016/j.jii.2024.100639.

[19] A. Castro, G. Cruz, V. Bough, and S. Tiwari. Enhancing the Customer Journey with Gen AI–Powered Digital Twins, 2024. URL `www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/enhancing-the-customer-journey-with-gen-ai-powered-digital-twins`. Accessed October 1, 2024.

[20] A. Cosmas, G. Cruz, S. Cubela, M. Huntington, S. Rahimi, and S. Tiwari. Digital Twins and Generative AI: A Powerful Pairing, 2024. URL `www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/digital-twins-and-generative-ai-a-powerful-pairing`. Accessed October 1, 2024.

[21] D. G. Costa, J. C. N. Bittencourt, F. Oliveira, J. P. J. Peixoto, and T. C. Jesus. Achieving Sustainable Smart Cities through Geospatial Data-Driven Approaches. *Sustainability*, 16(2), 2024. doi:10.3390/su16020640.

[22] R. Das and M. Soylu. A Key Review on Graph Data Science: The Power of Graphs in Scientific Studies. *Chemom. Intell. Lab. Syst.*, 240:104896, 2023. doi:10.1016/j.chemolab.2023.104896.

[23] R. D. D'Amico, J. A. Erkoyuncu, S. Addepalli, and S. Penver. Cognitive Digital Twin: An Approach to Improve the Maintenance Management. *CIRP J. Manuf. Sci. Technol.*, 38:613–630, 2022. doi:10.1016/j.cirpj.2022.06.004.

[24] A. Eibeck, M. Q. Lim, and M. Kraft. J-Park Simulator: An Ontology-Based Platform for Cross-Domain Scenarios in Process Industry. *Comput. Chem. Eng.*, 131:106586, 2019. doi:10.1016/j.compchemeng.2019.106586.

[25] P. Eirinakis, K. Kalaboukas, S. Lounis, I. Mourtos, J. M. Rožanec, N. Stojanovic, and G. Zois. Enhancing Cognition for Digital Twins. In *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 1–7, 2020. doi:10.1109/ICE/ITMC49519.2020.9198492.

[26] F. T. Fonseca, M. J. Egenhofer, C. A. Davis, and K. A. V. Borges. Ontologies and Knowledge Sharing in Urban GIS. *Comput. Environ. Urban Syst.*, 24(3):251–272, 2000. doi:10.1016/S0198-9715(00)00004-1.

[27] F. T. Fonseca, M. J. Egenhofer, C. A. Davis, and G. Câmara. Semantic Granularity in Ontology-Driven Geographic Information Systems. *Ann. Math. Artif. Intell.*, 36: 121–151, 2002. doi:10.1023/A:1015808104769.

[28] A. A. A. Gassar and S. H. Cha. Review of Geographic Information Systems-Based Rooftop Solar Photovoltaic Potential Estimation Approaches at Urban Scales. *Appl. Energy*, 291:116817, 2021. doi:10.1016/j.apenergy.2021.116817.

[29] E. Glaessgen and D. Stargel. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, 2012. doi:10.2514/6.2012-1818.

[30] M. Grieves. *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management*. Space Coast Press, 11 2011. ISBN 0982138008.

[31] M. Grieves. *Working Paper: Origins of the Digital Twin Concept*. Florida Institute of Technology: Melbourne, FL, USA, 08 2016. doi:10.13140/RG.2.2.26367.61609.

[32] M. W. Grieves. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *Dassault Systèmes*, pages 1–7, 2014. URL `https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf`.

[33] A. Grisiute, H. Silvennoinen, S. Li, A. Chadzynski, M. Raubal, M. Kraft, A. von Richthofen, and P. Herthogs. A Semantic Spatial Policy Model to Automatically Calculate Allowable Gross Floor Areas in Singapore. In M. Turrin, C. Andriotis, and A. Rafiee, editors, *Computer-Aided Architectural Design. INTERCONNECTIONS: Co-computing Beyond Boundaries*, pages 455–469, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-37189-9.

[34] T. R. Gruber. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199–220, 1993. doi:10.1006/knac.1993.1008.

[35] A. Göppert, L. Grahn, J. Rachner, D. Grunert, S. Hort, and R. H. Schmitt. Pipeline for Ontology-Based Modeling and Automated Deployment of Digital Twins for Planning and Control of Manufacturing Systems. *J. Intell. Manuf.*, 34:2133–2152, 2021. doi:10.1007/s10845-021-01860-6.

[36] R. Habeck. Weg frei für eine klimafreundliche und bezahlbare Wärmeversorgung, 2023. URL www.bmwk.de/Redaktion/DE/Pressemitteilungen/2023/11/20231117-klimafreundliche-und-bezahlbare-waermeversorgung.html. Accessed October 1, 2024.

[37] A. H. Hay, B. Karney, and N. Martyn. Reconstructing Infrastructure for Resilient Essential Services During and Following Protracted Conflict: A Conceptual Framework. *Int. Rev. Red Cross*, pages 1001–1029, 2019. doi:10.1017/S1816383120000053.

[38] P. Hitzler, M. Krötzsch, and S. Rudolph. *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, 2009.

[39] M. Hofmeister, J. Bai, G. Brownbridge, S. Mosbach, K. F. Lee, F. Farazi, M. Hillman, M. Agarwal, S. Ganguly, J. Akroyd, and M. Kraft. Semantic Agent Framework for Automated Flood Assessment using Dynamic Knowledge Graphs. *Data-Centric Eng.*, 5:e14, 2024. doi:10.1017/dce.2024.11.

[40] M. Hofmeister, K. F. Lee, Y.-K. Tsai, M. Müller, K. Nagarajan, S. Mosbach, J. Akroyd, and M. Kraft. Dynamic Control of District Heating Networks with Integrated Emission Modelling: A Dynamic Knowledge Graph Approach. *Energy AI*, 17:100376, 2024. doi:10.1016/j.egyai.2024.100376.

[41] K. Janowicz, P. Hitzler, W. Li, D. Rehberger, M. Schildhauer, R. Zhu, C. Shimizu, C. K. Fisher, L. Cai, G. Mai, J. Zalewski, L. Zhou, S. Stephen, S. Gonzalez, B. Mecum, A. Lopez-Carr, A. Schroeder, D. Smith, D. Wright, S. Wang, Y. Tian, Z. Liu, M. Shi, A. D'Onofrio, Z. Gu, and K. Currier. Know, Know Where, KnowWhereGraph: A Densely Connected, Cross-Domain Knowledge Graph and Geo-Enrichment Service Stack for Applications in Environmental Intelligence. *AI Magazine*, 43:30–39, 2022. doi:10.1002/aaai.12043.

[42] D. Kohli, R. Sliuzas, N. Kerle, and A. Stein. An Ontology of Slums for Image-Based Classification. *Comput. Environ. Urban Syst.*, 36(2):154–163, 2012. doi:10.1016/j.compenvurbsys.2011.11.001.

[43] A. Kondinski, P. Rutkevych, L. Pascazio, D. N. Tran, F. Farazi, S. Ganguly, and M. Kraft. Knowledge Graph Representation of Zeolitic Crystalline Materials. *Digit. Discov.*, 3:2070–2084, 2024. doi:10.1039/D4DD00166D.

[44] A. Kovacs-Györi, A. Ristea, C. Havas, M. Mehaffy, H. H. Hochmair, B. Resch, L. Juhasz, A. Lehner, L. Ramasubramanian, and T. Blaschke. Opportunities and Challenges of Geospatial Analysis for Promoting Urban Livability in the Era of Big Data and Machine Learning. *ISPRS Int. J. Geo-Inf.*, 9(12), 2020. doi:10.3390/ijgi9120752.

[45] M. Kraft and S. Mosbach. The Future of Computational Modelling in Reaction Engineering. *Phil. Trans. R. Soc. A.*, 368(1924):3633—-3644, 2010. doi:10.1098/rsta.2010.0124.

[46] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn. Digital Twin in Manufacturing: A Categorical Literature Review and Classification. *IFAC-PapersOnLine*, 51(11):1016–1022, 2018. doi:10.1016/j.ifacol.2018.08.474.

[47] Y. LeCun, Y. Bengio, and G. Hinton. Deep Learning. *Nature*, 521:436–444, 2015. doi:10.1038/nature14539.

[48] M. Q. Lim, X. Wang, O. Inderwildi, and M. Kraft. *The World Avatar - A World Model for Facilitating Interoperability*, pages 39–53. Springer International Publishing, 2022. ISBN 978-3-030-86215-2. doi:10.1007/978-3-030-86215-2_4.

[49] M. Liu, S. Fang, H. Dong, and C. Xu. Review of Digital Twin about Concepts, Technologies, and Industrial Applications. *J. Manuf. Syst.*, 58:346–361, 2021. doi:10.1016/j.jmsy.2020.06.017. Digital Twin towards Smart Manufacturing and Industry 4.0.

[50] Y. Loukili, Y. Lakhrissi, and S. Ali. Geospatial Big Data Platforms: A Comprehensive Review. *KN J. Cartogr. Geogr. Inf.*, 72:293–308, 2022. doi:10.1007/s42489-022-00121-7.

[51] J. Lu, X. Zheng, L. Schweiger, and D. Kiritsis. A Cognitive Approach to Manage the Complexity of Digital Twin Systems. In S. West, J. Meierhofer, and C. Ganz, editors, *Smart Services Summit*, pages 105–115, Cham, 2021. Springer International Publishing.

[52] T. Malaker and Q. Meng. Urban Disparity Analytics Using GIS: A Systematic Review. *Sustainability*, 16(14), 2024. doi:10.3390/su16145956.

[53] J. L. Mennis. Derivation and Implementation of a Semantic GIS Data Model Informed by Principles of Cognition. *Comput. Environ. Urban Syst.*, 27(5):455–479, 2003. doi:10.1016/S0198-9715(02)00019-4.

[54] A. Mobasheri, F. Pirotti, and Agugiaro. Open-source Geospatial Tools and Technologies for Urban and Environmental Studies. *Open Geospatial Data, Softw. Stand.*, 5, 2020. doi:10.1186/s40965-020-00078-2.

[55] K. E. Mokhtari, I. Panushev, and J. J. McArthur. Development of a Cognitive Digital Twin for Building Management and Operations. *Front. Built Environ.*, 8:1–18, 2022. doi:10.3389/fbuil.2022.856873.

[56] T. H. Nguyuen. 30 Emerging Technologies That Will Guide Your Business Decisions, 2024. URL `www.gartner.com/en/articles/30-emerging-technologies-that-will-guide-your-business-decisions`. Accessed October 1, 2024.

[57] P. Niloofar, S. Lazarova-Molnar, F. Omitaomu, H. Xu, and X. Li. A General Framework for Human-in-the-Loop Cognitive Digital Twins. In *2023 Winter Simulation Conference (WSC)*, pages 3202–3213, 2023. doi:10.1109/WSC60868.2023.10407598.

[58] H. Nordahl, M. Rindarøy, S. Skjong, L. T. Kyllingstad, D. Walther, and T. Brekke. An Ontology-Based Approach for Simplified FMU Variable Connections With Automatic Verification of Semantically Correct Configuration. *International Conference on Offshore Mechanics and Arctic Engineering*, Volume 6A: Ocean Engineering:V06AT06A022, 08 2020. doi:10.1115/OMAE2020-18135.

[59] S. Olcott and C. Mullen. Digital Twin Consortium Defines Digital Twin, 2020. URL www.digitaltwinconsortium.org/2020/12/digital-twin-consortium-defines-digital-twin/. Accessed October 1, 2024.

[60] L. Pascazio, D. Tran, S. D. Rihm, J. Bai, S. Mosbach, J. Akroyd, and M. Kraft. Question-Answering System for Combustion Kinetics. *Proc. Combust. Ins.*, 40(1): 105428, 2024. doi:10.1016/j.proci.2024.105428.

[61] S. Z. Phua, M. Hofmeister, Y.-K. Tsai, O. Peppard, K. F. Lee, S. Courtney, S. Mosbach, J. Akroyd, and M. Kraft. Fostering Urban Resilience and Accessibility in Cities: A Dynamic Knowledge Graph Approach. *Sustain. Cities Soc.*, 113:105708, 2024. doi:10.1016/j.scs.2024.105708.

[62] S. Z. Phua, K. F. Lee, Y.-K. Tsai, S. Ganguly, J. Yan, S. Mosbach, T. Ng, A. Moise, B. P. Horton, and M. Kraft. Urban Vulnerability Assessment of Sea Level Rise in Singapore through The World Avatar. *Appl. Sci.*, 14(17), 2024. doi:10.3390/app14177815.

[63] H. Y. Quek, F. Sielker, J. Akroyd, A. N. Bhave, A. von Richthofen, P. Herthogs, C. v. d. L. Yamu, L. Wan, T. Nochta, G. Burgess, M. Q. Lim, S. Mosbach, and M. Kraft. The Conundrum in Smart City Governance: Interoperability and Compatibility in an Ever-Growing Ecosystem of Digital Twins. *Data & Policy*, 5:e6, 2023. doi:10.1017/dap.2023.1.

[64] B. Resch, G. Sagl, T. Törnros, A. Bachmaier, J.-B. Eggers, S. Herkel, S. Narmsara, and H. Gündra. GIS-Based Planning and Modeling for Renewable Energy: Challenges and Future Research Avenues. *ISPRS Int. J. Geo-Inf.*, 3(2):662–692, 2014. doi:10.3390/ijgi3020662.

[65] Research and Markets. Global Digital Twin Market by Application (Predictive Maintenance, Business Optimization, Performance Monitoring, Inventory Management), Industry (Automotive & Transportation, Healthcare, Energy & Utilities), Enterprise and Geography - Forecast to 2028, 2023. URL https://www.researchandmarkets.com/reports/5146336/global-digital-twin-market-by-application?srsltid=AfmBOoq3MLjzpKUVJMsSmfuqaXjJSxZghfuGPbrYUnvTiY9UBnRL5ptZ. Accessed October 1, 2024.

[66] S. D. Rihm, Y. R. Tan, W. Ang, M. Hofmeister, X. Deng, M. T. Laksana, H. Y. Quek, J. Bai, L. Pascazio, S. C. Siong, J. Akroyd, S. Mosbach, and M. Kraft. The Digital Lab Manager: Automating Research Support. *SLAS Technol.*, 29(3):100135, 2024. doi:10.1016/j.slast.2024.100135.

[67] S. D. Rihm, Y. R. Tan, W. Ang, H. Y. Quek, X. Deng, M. T. Laksana, J. Bai, S. Mosbach, J. Akroyd, and M. Kraft. The Digital Lab Facility Manager: Automating operations of research laboratories through "The World Avatar". *Nexus*, 1(3):100031, 2024. doi:10.1016/j.ynexs.2024.100031.

[68] J. M. Rožanec, J. Lu, J. Rupnik, M. Škrjanc, D. Mladenić, B. Fortuna, X. Zheng, and D. Kiritsis. Actionable Cognitive Twins for Decision Making in Manufacturing. *Int. J. Prod. Res.*, 60(2):452–478, 2022. doi:10.1080/00207543.2021.2002967.

[69] R. Sacks, I. Brilakis, E. Pikas, H. S. Xie, and M. Girolami. Construction with Digital Twin Information Systems. *Data-Centric Eng.*, 1:e14, 2020. doi:10.1017/dce.2020.16.

[70] R. Sharma and H. Gupta. Leveraging Cognitive Digital Twins in Industry 5.0 for Achieving Sustainable Development Goal 9: An Exploration of Inclusive and Sustainable Industrialization Strategies. *J. Clean. Prod.*, 448:141364, 2024. doi:10.1016/j.jclepro.2024.141364.

[71] Z. Shi, H. Silvennoinen, A. Chadzynski, A. von Richthofen, M. Kraft, S. Cairns, and P. Herthogs. Defining Archetypes of Mixed-Use Developments using Google Maps API Data. *Environ. Plan. B: Urban Anal. City Sci.*, 50(6):1607–1623, 2023. doi:10.1177/23998083221141428.

[72] H. Silvennoinen, A. Chadzynski, F. Farazi, A. Grišiūtė, Z. Shi, A. von Richthofen, S. Cairns, M. Kraft, M. Raubal, and P. Herthogs. A Semantic Web Approach to Land Use Regulations in Urban Planning: The OntoZoning Ontology of Zones, Land Uses and Programmes for Singapore. *J. Urban Manag.*, 12(2):151–167, 2023. doi:10.1016/j.jum.2023.02.002.

[73] K. Sun, Y. Zhu, P. Pan, Z. Hou, D. Wang, W. Li, and J. Song. Geospatial Data Ontology: The Semantic Foundation of Geospatial Data Integration and Sharing. *Big Earth Data*, 3(3):269–296, 2019. doi:10.1080/20964471.2019.1661662.

[74] M. Tegmark. *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York : Alfred A. Knopf, 2017.

[75] A. Thelen, X. Zhang, O. Fink, Y. Lu, S. Ghosh, B. D. Youn, M. D. Todd, S. Mahadevan, C. Hu, and Z. Hu. A Comprehensive Review of Digital Twin — Part 1: Modeling and Twinning Enabling Technologies. *Struct. Multidiscip. Optim.*, 65 (354):1–55, 2022. doi:10.1007/s00158-022-03425-4.

[76] D. Tran, L. Pascazio, J. Akroyd, S. Mosbach, and M. Kraft. Leveraging Text-to-Text Pretrained Language Models for Question Answering in Chemistry. *ACS Omega*, 9: 13883–13896, 2024. doi:10.1021/acsomega.3c08842.

[77] Y.-K. Tsai, M. Hofmeister, S. Ganguly, K. Rustagi, Y. R. Tan, S. Mosbach, J. Akroyd, and M. Kraft. Municipal Heat Planning within The World Avatar. Submitted for publication; preprint available online, 2024. URL `https://como.ceb.cam.ac.uk/preprints/326/`.

[78] P. H. Verburg, K. Neumann, and L. Nol. Challenges in Using Land Use and Land Cover Data for Global Change Studies. *Glob. Change Biol.*, 17(2):974–989, 2011. doi:10.1111/j.1365-2486.2010.02307.x.

[79] A. von Richthofen, P. Herthogs, M. Kraft, and S. Cairns. Semantic City Planning Systems (SCPS): A Literature Review. *Journal of Planning Literature*, 37(3):415–432, 2022. doi:10.1177/08854122211068526.

[80] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S. A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. Van Der Lei, E. Van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data*, 3:1–9, 2016. doi:10.1038/sdata.2016.18.

[81] N. Wu, F. Zhang, J. Wang, X. Wang, J. Wu, J. Huang, J. Tan, R. Jing, J. Lin, S. Xie, and Y. Zhao. A Review on Modelling Methods, Tools and Service of Integrated Energy Systems in China. *Prog. Energy*, 5(3):032003, 2023. doi:10.1088/2516-1083/acef9e.

[82] Z. Xiaochen, L. Jinzhi, A. Rebeca, R. Jože Martin, L. Stavros, K. Kostas, and K. Dimitris. *Cognitive Digital Twins*, pages 51–65. CRC Press Taylor & Francis Group, 2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431 U.S.A., 2024. ISBN 9781003425724. doi:10.1201/9781003425724-6.

[83] W. Xie, J. Atherton, J. Bai, F. Farazi, S. Mosbach, J. Akroyd, and M. Kraft. A Nuclear Future? Small Modular Reactors in a Carbon Tax-Driven Transition to Clean Energy. *Appl. Energy*, 364:123128, 2024. doi:10.1016/j.apenergy.2024.123128.

[84] W. Xie, F. Farazi, J. Atherton, J. Bai, S. Mosbach, J. Akroyd, and M. Kraft. Dynamic Knowledge Graph Approach for Modelling the Decarbonisation of Power Systems. *Energy AI*, 17:100359, 2024. doi:10.1016/j.egyai.2024.100359.

[85] J. Xu, S. Mosbach, J. Akroyd, and M. Kraft. Impact of Heat Pumps and Future Energy Prices on Regional Inequalities. Submitted for publication; preprint available online, 2024. URL `https://como.ceb.cam.ac.uk/preprints/323/`.

[86] L. Zhang, B. Zhang, T. Guo, F. Chen, P. Runcie, B. Cameron, and R. Rooney. *Linking Complex Urban Systems: Insights from Cross-Domain Urban Data Analysis*, pages 221–239. Springer Nature Singapore, Singapore, 2020. ISBN 978-981-13-6605-5. doi:10.1007/978-981-13-6605-5.

[87] X. A. Zhang, P. Kinder, M. Strager, S. Taylor, and G. Schwartzman. Distributed Energy Infrastructure Development: Geospatial and Economic Feasibility in Rural

West Virginia. *Environ. Dev. Sustain.*, 72:293–308, 2022. doi:10.1007/s10668-024-05363-1.

[88] X. Zheng, F. Psarommatis, P. Petrali, C. Turrin, J. Lu, and D. Kiritsis. A Quality-Oriented Digital Twin Modelling Method for Manufacturing Processes Based on A Multi-Agent Architecture. *Procedia Manuf.*, 51:309–315, 2020. doi:10.1016/j.promfg.2020.10.044. 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021).

[89] X. Zheng, J. Lu, R. Arista, X. Hu, J. Lentes, F. Ubis, J. Sorvari, and D. Kiritsis. Development of an Application Ontology for Knowledge Management to Support Aircraft Assembly System Design. *CEUR Workshop Proc.*, 2021. doi:10.24406/publica-fhg-412664.

[90] X. Zheng, J. Lu, and D. Kiritsis. The Emergence of Cognitive Digital Twin: Vision, Challenges and Opportunities. *Int. J. Prod. Res.*, 60(24):7610–7632, 2022. doi:10.1080/00207543.2021.2014591.

[91] X. Zhou, S. Zhang, M. Agarwal, J. Akroyd, S. Mosbach, and M. Kraft. Marie and BERT-A Knowledge Graph Embedding Based Question Answering System for Chemistry. *ACS Omega*, 8:33039–33057, 2023. doi:10.1021/acsomega.3c05114.

[92] J. Šturm, P. Zajec, M. Škrjanc, D. Mladenić, and M. Grobelnik. Enhancing Cognitive Digital Twin Interaction using an LLM Agent. In *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*, pages 103–107, 2024. doi:10.1109/MIPRO60963.2024.10569919.