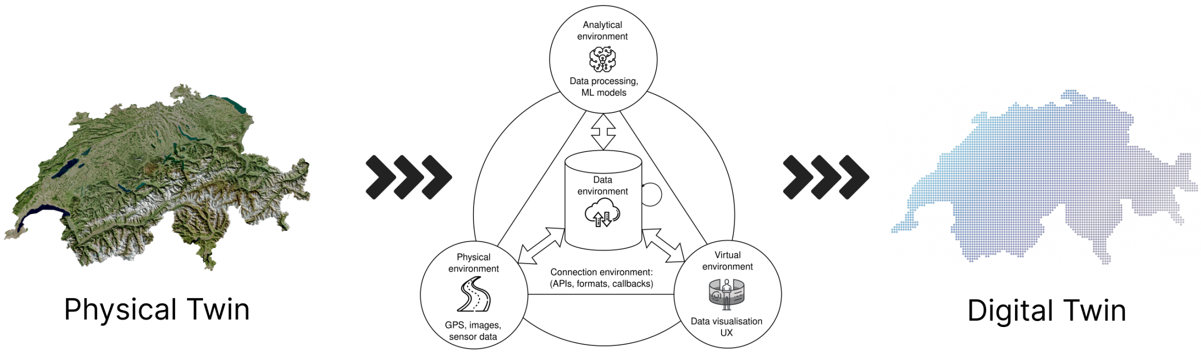


“CH on the move”: Introducing the Prototype Digital Twin of The Swiss Mobility System

Jascha Grübel

Carlos Vivar Rios

Milos Balac

Yanan Xin

Robin M. Franken

Sabrina Ossey

Martin Raubal

Kay W. Axhausen

Oksana Riba-Grognuz

STRC conference paper 2023 **May 1, 2023**

STRC | 23rd Swiss Transport Research Conference
Monte Verità / Ascona, May 10-12, 2023

"CH on the move": Introducing the Prototype Digital Twin of The Swiss Mobility System

Jascha Grübel
CSFM & Geoinformation Engineering
ETH Zurich
jgruebel@ethz.ch

Milos Balac
IVT
ETH Zurich
milos.balac@ivt.baug.ethz.ch

Robin M. Franken
ORDES
SDSC
robin.franken@epfl.ch

Martin Raubal
Geoinformation Engineering
ETH Zurich
mraubal@ethz.ch

Oksana Riba-Grognuz
ORDES
SDSC
oksana.riba@epfl.ch

Carlos Vivar Rios
ORDES
SDSC
carlos.vivarrios@epfl.ch

Yanan Xin
Geoinformation Engineering
ETH Zurich
yanxin@ethz.ch

Sabrina Ossey
ORDES
SDSC
sabrina.ossey@epfl.ch

Kay W. Axhausen
IVT
ETH Zurich
axhausen@ivt.baug.ethz.ch

May 1, 2023

Abstract

Mobility simulations are crucial for research, industry, and governance. However, simulating mobility systems with reasonable accuracy requires large-scale efforts of data collection, data transformation, data analysis, and data visualization. The paradigm of the digital twin offers a novel perspective on how to manage the data efficiently and make the simulations available more steadily at a lower cost. We introduce the first prototype of the "CH on the move" digital twin that is designed to be openly available to all interested parties to enable a common baseline for transport research in Switzerland. This digital twin is based on the extensible Open Digital Twin Platform (ODTP) that uses containerization, loose coupling, and micro-services to provide dynamically composable digital twins. In its first iteration, "CH on the move" provides an easy-to-use version of the eqasim pipeline for MATSim making it possible to initiate simulations of Switzerland with one click. Future versions of "CH on the move" aim to include a rich set of relevant data sources and analytical pipelines related to transport and mobility and make them easily accessible to research, industry, and governance.

Keywords

Digital Twin, Swiss Mobility System, Mobility Simulation, eqasim, MATSim

Suggested Citation

Contents

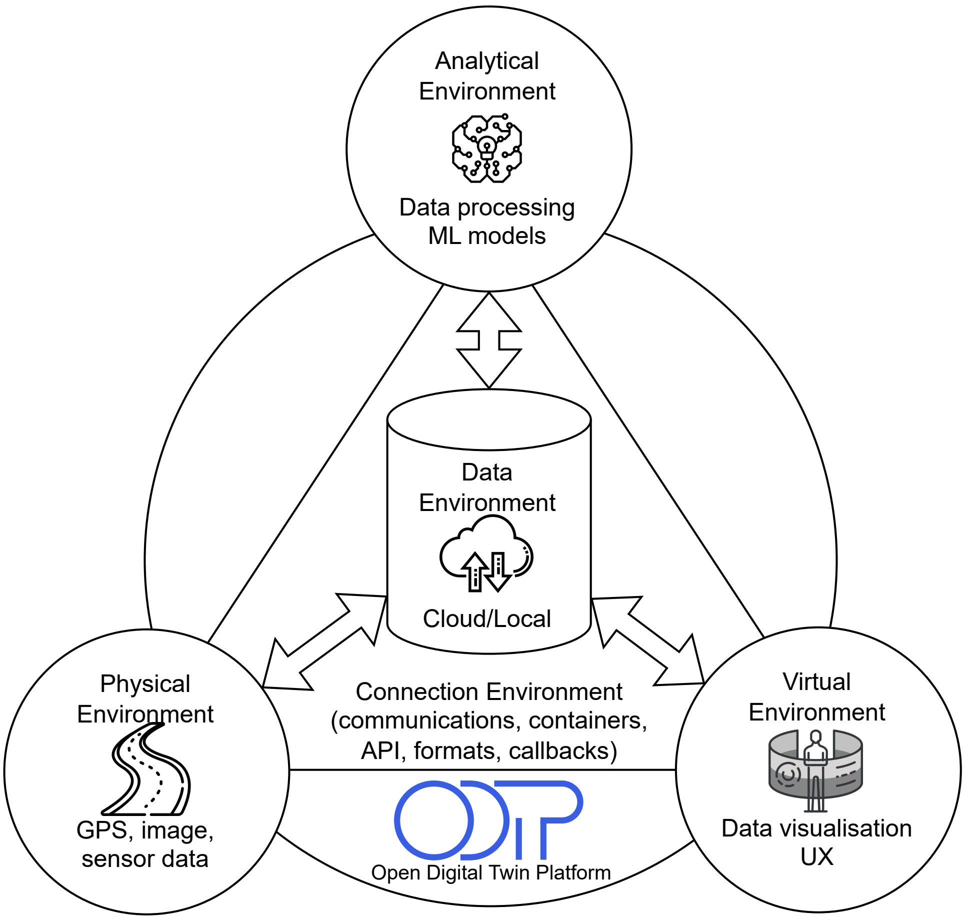
List of Tables

List of Figures

1 Introduction

We introduce a new project for a Digital Twin of the Swiss Mobility System. The project is currently in the late design phase and early implementation phase here we will provide a first preview of the Digital Twin. Our underlying model for Digital Twins is agnostic and scale-free and therefore applicable beyond mobility, see Fig. 1.

Figure 1: ODTP Overview: The platform follows a five component model for Digital Twins (Grübel *et al.*, 2022; Tao *et al.*, 2018). Each component represents a key environment for interaction with data. The physical environment captures variables of interest in the real world and provides raw data. The data environment attaches semantic meaning to the data and ensures compatibility. The analytical environment hosts data transformation applications from preprocessors to simulators and machine learning models. The virtual environment provides end-user representations from tables to dashboards to 2D and 3D visualisations. Lastly, the connection environment provides the framework that enables the hosting of different services within each environment and ensures their interoperability.



Source: Jascha Grübel

Indeed, ODTP can be applied on an arbitrary resolution for any research topic and hence in any application area from academica to industry and governance. Importantly, Digital Twins by design offer themselves as a platform for Open Research Data (ORD) (Nosek *et al.*, 2015) and the implementation of Findable, Accessible, Interoperable and Reusable (FAIR) data (Wilkinson *et al.*, 2016). We propose the Open Digital Twin Platform (ODTP) as a vehicle for ORD in research, industry and government.

1.1 Digital Twins

While there is no consensus on the meaning of Digital Twins (Grübel *et al.*, 2022), they usually consist of geo-referenced data to represent relevant elements of a real-world system for analysis and decision-making. This description is also appropriate for most Geo-Information Systems (GIS). However, the model simplicity—a Physical Twin and a Digital Twin (Grübel *et al.*, 2022)—that subsumes implementation complexity is alluring across industry, governance, and research. New Digital Twins are constantly developed in all fields and for all areas(Grübel *et al.*, 2022). Nonetheless, each of these Digital Twins is a unique ungeneralised artefact and there is still no open format for designing Digital Twins and exchanging data between Digital Twins (Roest, 2019). There are some initiatives for particular tasks like the Asset Administration Shell (Industrial Digital Twin Association, 2023) but in our model this would fall into the area of solutions for the Physical Environment rather than generalised Digital Twins.

In the mobility context, Digital Twins focus on describing individual, household, and aggregate spatial behavior with a temporal resolution anywhere from annual to real-time data. There is a wide array of ORD practices already in place such as NADIM (Federal Office of Transport, 2022), MATSim (Horni *et al.*, 2016), eqasim, (Hörl and Balac, 2021) and many more that could be combined into Digital Twins. Digital Twins would facilitate the access to information, the configuration of complex models and the communication of results. Embedding existing ORD practices in Digital Twins would allow researchers to only work on aspects that are fundamental to their work rather than having to also provide the surrounding framework themselves. Given the importance of reproducibility (Ioannidis, 2005), it becomes important that researchers don't only communicate their findings but also how to attain them (Stark, 2018). In this project, we aim to close this gap by joining existing open standards to coordinate them as a Digital Twin open standard. The Digital Twin format allows for practices of easier communication of ideas, results, and methods as the data is represented from collection over analysis to presentation. Thus, we aim to realize the potential of ORD practices in mobility research and beyond through the Digital Twin approach.

At the same time, industry is creating standards for Digital Twins that are proprietary (e.g., Microsoft, 2023) and could become the norm by the virtue of simply being there like early web browsers and operating systems. We identified the unique opportunity to create an open standard for Digital Twins right now because we are at the crossroads where the required technologies for Digital Twins are maturing, and we have no default yet. The ongoing debates have identified the Semantic Web (Berners-Lee *et al.*, 2001) as a core technology to make information exchangeable across systems through knowledge graphs upon which Digital Twins could be based (Akroyd *et al.*, 2021; Boje *et al.*, 2020; Microsoft, 2023). Whereas an ontology is a crucial component to interact with unknown Digital Twins based on a generic format (McGuinness and Van Harmelen, 2004)—a Digital Twin is more than just a knowledge exchange platform (Grübel *et al.*, 2022). The systematic nature of Digital Twins makes them a natural candidate to implement FAIR (findable, accessible, interoperable, and reusable) data standards (Schultes *et al.*, 2022) that guide ORD practices and enable reproducible research. In the best case, Digital Twins as an open standard can be a driver of development like the Internet. In the worst case, they could be bogged down in a battle of standards like HD-DVD vs BluRay that slow down adaptation because of uncertainty about which format will eventually become the default.

Industrial players from large multinational companies to consultants are keen to offer their Digital Twin solutions in the hope to fence off significant market shares from the imagined application fields of Digital Twins. The reasons why there is no open standard yet are multitude but include these wild-west assumptions of a large untapped opportunity that can be conquered and fenced off (Roest, 2019). Another issue is that the underlying technologies are only now maturing and are often still under development themselves. This makes early Digital Twins fickle and prone to breaking down. Most current actors hope to be a defining force for Digital Twins to come and therefore try to create facts on the ground. In this fast-paced environment, only few truly open initiatives have been formed. Mostly, they revolve around specific EU projects such as DUET (Raes *et al.*, 2021) or industry platforms such as Bosch's open IoT ecosystem DITTO. These approaches are often hardware-specific, task-specific, and platform-specific, and ultimately fail to provide a common standard that can be generalised across all Digital Twin applications from mobility over construction to industrial production and more.

Nonetheless, Digital Twins are attractive for many communities because they inherently couple data collection, data analysis and data presentation and provide users ranging from policy makers and researchers to citizens with an integrated view on information. As a point of departure for this paper we will focus on mobility. This includes the generation of synthetic populations (SynPops (Müller and Axhausen, 2011)) which are a key component for mobility modelling based on the "New Census" strategy of Switzerland. A large number of synthetic population models is available, but their comparison is currently not possible due to the differences in assumptions. Embedding these models in a Digital Twin will allow us to make them comparable and advance the discourse on "SynPops" through validation as well as giving mobility modelling in Switzerland a stronger and open foundation. However, at the same time, we engage other research fields such as energy and land use to come up with a holistic standard that may be deployable across disciplinary boundaries.

1.2 Open Research Data

Research data is often inaccessible, incomprehensible and consequently irreproducible (Ioannidis, 2005; Stark, 2018). To address these shortcomings, Open Research Data (ORD) (Nosek *et al.*, 2015) aims to promote the easy use and access to research data according to the Findable, Accessible, Interoperable and Reusable (FAIR) data paradigm (Wilkinson *et al.*, 2016). However, issues of data usage go hand in hand with issues of research design, research execution and research reproduction (Aguilar *et al.*, 2022). Digital Twins offer a platform to conduct research, in particular experiments (Grübel, 2023b). Combining Digital Twins and ORD offers new possibilities on improving reproducibility and preproducibility in research and beyond (Grübel, 2023a). Digital Twins provide a framework to document, understand, and reproduce data flows and enable the precise re-enactment of experiments.

On the one hand, the Digital Twin-based ORD practices that were hoped for so far are still lacking (Roest, 2019) as standards diverge and too many actors pursue too many different goals. On the other hand, we are close to building Open Digital Twins as the constituent components of Digital Twins (see Fig. 1 for a detailed explanation of our model (Grübel *et al.*, 2022)) can be implemented with existing ORD solutions. The physical and data environment can be provided by projects that implement the Federal Open Data Strategy in Switzerland such as NADIM (Federal Office of Transport, 2022), opendata.swiss (Confederation, Cantons, & Communes, 2022) and municipalities and cantons who openly integrate and aggregate mobility knowledge across Switzerland from all available open sources and some access-limited sources as well. The analytical environment of a Digital Twin could be provided by an open-source knowledge infrastructure for collaborative and reproducible data science such as Renku (Swiss Data Science Center, 2022) implementing, for example synthetic population models (Beckman *et al.*, 1996). The virtual environment to access, interact with, and visualise the data can be based on classical dashboards in Cesium (Cesium GS, 2019) environments. Eventually, it might be possible to implement them with modern visualisation approaches based on extended reality (XR) applications (Grübel *et al.*, 2021). At the core of this paper will be to provide ORD solutions for the connection environment such as developed in the Experiments as Code paradigm (Aguilar *et al.*, 2022) that allow us to interlink these ongoing efforts. Thereby, we enable a simple interaction with mobility data and enable reproducible workflows that push ORD practices in mobility research to the next level.

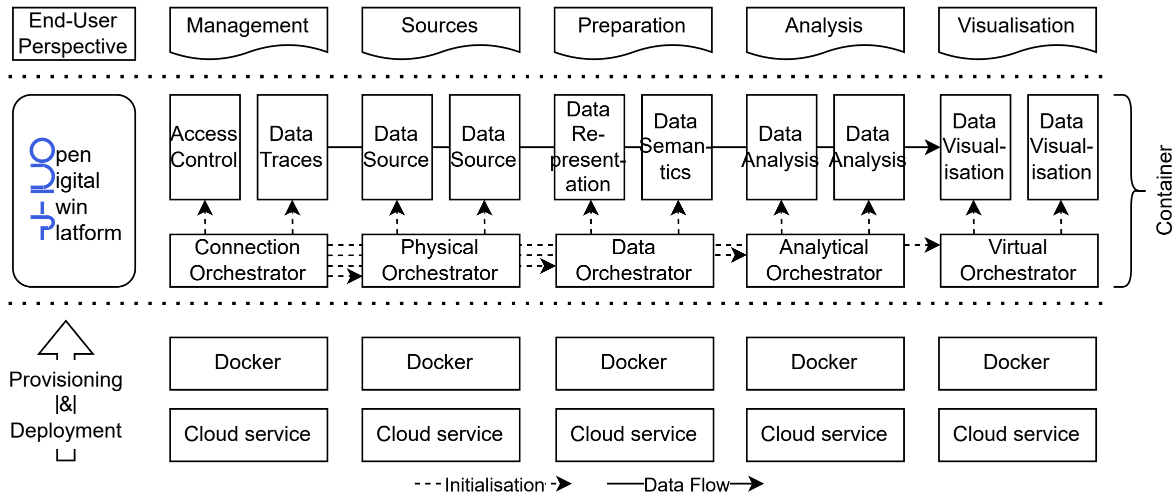
2 System and Implementation

The overall architecture of ODTP, see Fig. 1, is based on an analysis of existing digital twin solutions (Grübel *et al.*, 2022) that finds common denominators across models and abstractions. The five environments of the ODTP architecture seek to conceptually split tasks into meaningful and modular steps. We make a clear separation between data acquisition, data representation, data analysis and data visualisation. Whereas the abstract figure suggests each environment to be monolithic, reality is more nuances with a large number of micro-services (De Lauretis, 2019) that build up the required features of each environment, see Fig. 2. Micro-services allow for a task-specific assembly of services for each Digital Twin enabling performant yet general designs.

In general, micro-services provide several benefits such maintainability, modularity, scalability, implementation-independence, and blackboxing (Dmitry and Manfred, 2014). Micro-services are facilitated by container-based virtualisation of operating system features (Amaral *et al.*, 2015). In particular, the docker environment has become the go-to provider for containerisation. The key advantages are that each containerised service perceives itself to be the only in the system and all its dependencies are cleanly separated. Interactions between the dependencies in codebase of different services become impossible.

In ODTP, the micro-services are orchestrated by the platform to ensure that they satisfy

Figure 2: ODTP Implementation (Grübel *et al.*, 2023): The platform is implemented on top of docker services running in the cloud or on a local cluster. ODTP initialises first orchestrators for each environment of the Digital Twin which in turn initialise all containers used for a particular Digital Twin. From the end-user perspective, each environment takes on a particular task with regards to their data needs.

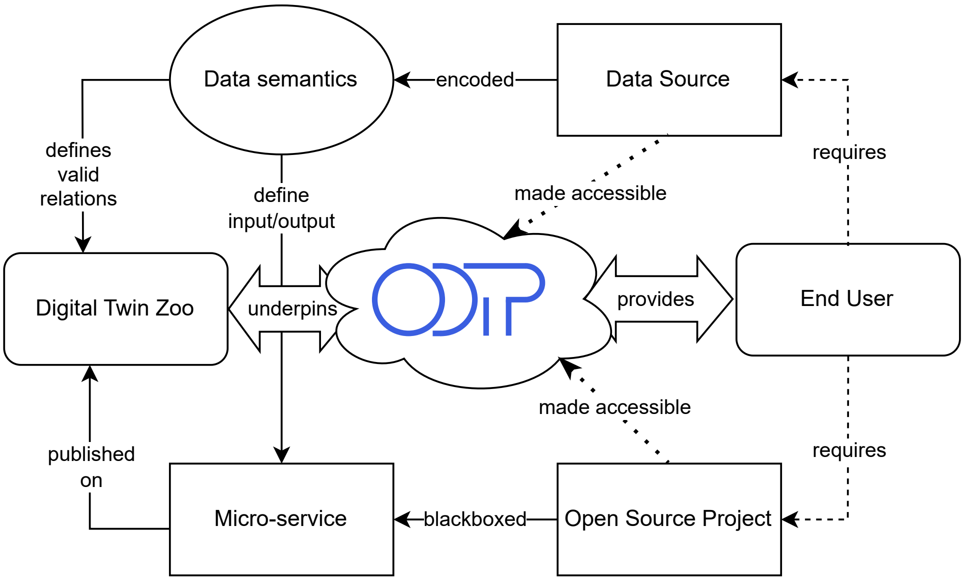


Source: Jascha Grübel

all criteria of respective Digital Twin environments. The connection orchestrator starts up a generic version of ODTP and uses existing data traces to determine the required micro-services. The environment orchestrators each load micro-service containers from trusted sources. End-users perceive the Digital Twin through the kind of interactions that they are pursuing such as managing the Twin, sourcing the data, preparing the data, analysing the data and visualising the data. Docker container provide the run time for these micro-services located either in the cloud or on local clusters.

Micro-services for ODTP are packaged versions of existing open-source projects that are made available as a containerised version, see Fig. 3. The major advantages are that the end user is not required to correctly install all dependencies and different components will not interact unless specified to do so. Instead, once a user specifies a requirement to ODTP, it can load the corresponding data and micro-service from the Digital Twin Zoo. The zoo terminology is loan word from machine learning approaches of model zoos—a large collection of prepared models for immediate use to reduce workload. The Digital Twin Zoo will host all micro-services containers that are publicly available and enable ODTP to instantiate a particular Digital Twin that an end user requested.

Figure 3: ODTP Micro-services: Users require software that is containerised into a micro-services and made available through the Digital Twin Zoo to ODTP. The data semantics ensure that a chain of micro-services is valid.

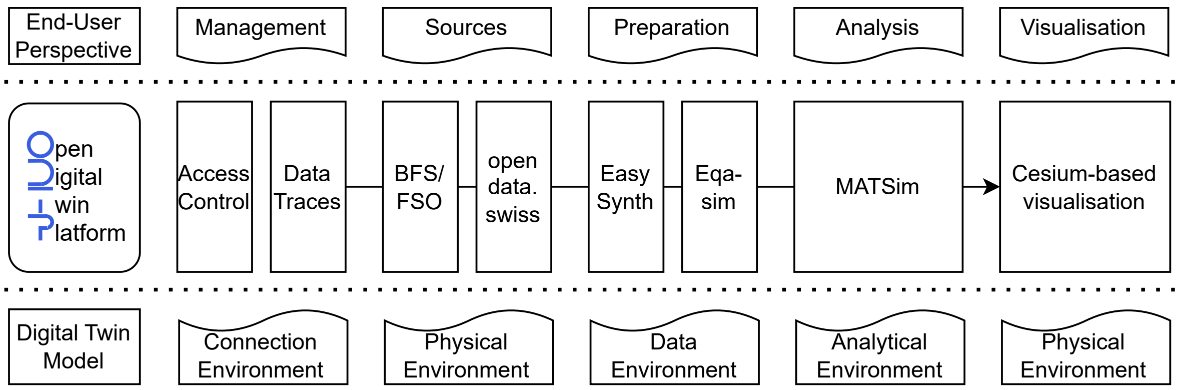


Source: Jascha Grübel

For the "CH on the move" Digital Twin, we implement a first prototype that only covers a particular aspect of mobility. We provide equilibrium-based simulations of the whole of Switzerland. MATSim has been applied to Switzerland on several occasions with different goals. However, the usability is still low and the barrier to entry is large. The prototype changes this by providing a "one-click" version of MATSim.

To that end, the "CH on the move" Digital Twin uses the ODTP infrastructure and micro-services for MATSim, eqasim, EasySynth and Cesium to generate the first prototype, see Fig. 4. EasySynth is providing access to public data by automatically accessing `opendata.swiss`. FSO data is loaded from a secured private server as its access is restricted. Eqasim is processing population data into functional synthetic populations for MATSim. Lastly, MATSim is providing mobility simulations to understand mobility in Switzerland.

Figure 4: ODTP Prototype of "CH on the move": Our prototype implements multi-modal mobility simulations based on MATSim (Horni *et al.*, 2016) and eqasim (Hörl and Balac, 2021) with automated data acquistion (Micallef *et al.*, 2023). Lastly, data visualisation will be conducted in Cesium (Cesium GS, 2019).



Source: Jascha Grübel

3 Discussion and conclusion

We introduce the Open Digital Twin Platform (ODTP) as a vehicle to generate a Digital Twin of the Swiss Mobility System called "CH on the move". Both ODTP and "CH on the move" are in their early development stages. This paper outlines the technology underpinning the system and our overall goals for future developments.

As it stands, "CH on the move" solely provides MATSim mobility simulations but the underlying ODTP allows us to expand our capabilities in the near future. Among others, we plan to include further mobility simulators to allow users to compare outputs. We will integrate real-time data sources of mobility such as the National Data Infrastructure for Mobility (NaDIM) developed by the Swiss Federal Office of Transport. To make data use more versatile, we plan to integrate data wrangling such as trackintel (Martin *et al.*, 2023). We are also working on including tools for causal inference on traffic forecasting.

The presented prototype is only the starting point for a comprehensive Digital Twin for the Swiss Mobility system that aims to provide researchers across all areas of transport with a reproducible baseline for data and models. "CH on the move" is also meant to be a community project and the open source approach will allow any interested researcher to contribute micro-services that can be re-used by everybody.

Acknowledgements

This research is funded by swissuniversities as part of the National Strategy and Action Plan for Open Research Data under the grant "An Open Digital Twin Platform for Research on the Swiss Mobility System" (ODTPR-SMS).

4 References

Aguilar, L., M. Gath-Morad, J. Grübel, J. Ermatinger, H. Zhao, S. Wehrli, R. W. Sumner, C. Zhang, D. Helbing and C. Hölscher (2022) Experiments as Code, available on arXiv.

Akroyd, J., S. Mosbach, A. Bhave and M. Kraft (2021) Universal digital twin-a dynamic knowledge graph, *Data-Centric Engineering*, **2**.

Amaral, M., J. Polo, D. Carrera, I. Mohomed, M. Unuvar and M. Steinder (2015) Performance evaluation of microservices architectures using containers, paper presented at the *2015 ieee 14th international symposium on network computing and applications*, 27–34.

Beckman, R. J., K. A. Baggerly and M. D. McKay (1996) Creating synthetic baseline populations, *Transportation Research Part A: Policy and Practice*, **30** (6) 415–429.

Berners-Lee, T., J. Hendler and O. Lassila (2001) The semantic web, *Scientific american*, **284** (5) 34–43.

Boje, C., A. Guerriero, S. Kubicki and Y. Rezgui (2020) Towards a semantic construction digital twin: Directions for future research, *Automation in Construction*, **114**, 103179.

Cesium GS (2019) About Cesium, `https://cesium.com/about/`. Accessed on 2023-04-28.

Confederation, Cantons, & Communes (2022) Swiss open government data, `https://opendata.swiss/en`. Accessed 28 June 2022.

De Lauretis, L. (2019) From monolithic architecture to microservices architecture, paper presented at the *2019 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*, 93–96.

Dmitry, N. and S.-S. Manfred (2014) On micro-services architecture, *International Journal of Open Information Technologies*, **2** (9) 24–27.

Federal Office of Transport (2022) Nationalen dateninfrastruktur mobilität, `https://www.bav.admin.ch/bav/en/home/general-topics/mmm.html`. Accessed 28 June 2022.

Grübel, J. (2023a) The design, experiment, analyse, and reproduce principle for experimentation in virtual reality, *Frontiers in Virtual Reality*, **4**.

Grübel, J. (2023b) *Handbook of Digital Twins*, chap. Experiments as Digital Twins, 20, CRC Press.

Grübel, J., M. Gath-Morad, L. Aguilar, T. Thrash, R. W. Sumner, C. Hölscher and V. R. Schinazi (2021) Fused Twins: A Cognitive Approach to Augmented Reality Media Architecture, paper presented at the *MAB '20: Proceedings of the 5th Media Architecture Biennale Conference.*

Grübel, J., T. Thrash, L. Aguilar, M. Gath-Morad, J. Chatain, R. W. Sumner, C. Hölscher and V. R. Schinazi (2022) The hitchhiker's guide to fused twins: A review of access to digital twins in situ in smart cities, *Remote Sensing*, **14** (13) 3095.

Grübel, J., C. Vivar Rios, M. Balać, Y. Xin, R. M. Franken, S. Ossey, M. Raubal, K. W. Axhausen and O. Riba-Grofnuz (2023) Open Digital Twin Platform, `https://csfm.ethz.ch/en/research/integrative-projects/digital-twin/odtp.html`.

Hörl, S. and M. Balac (2021) Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data, *Transportation Research Part C: Emerging Technologies*, **130**, 103291.

Horni, A., K. Nagel and K. W. Axhausen (2016) *The multi-agent transport simulation MATSim*, Ubiquity Press, London.

Industrial Digital Twin Association (2023) Asset administration shell, `https://industrialdigitaltwin.org/en/technology`. Accessed on 2023-04-28.

Ioannidis, J. P. (2005) Why most published research findings are false, *PLoS medicine*, **2** (8) e124.

Martin, H., Y. Hong, N. Wiedemann, D. Bucher and M. Raubal (2023) Trackintel: An open-source python library for human mobility analysis, *Computers, Environment and Urban Systems*, **101**, 101938.

McGuinness, D. L. and F. Van Harmelen (2004) Owl web ontology language overview, *W3C recommendation*, **10** (10) 10.

Micallef, D., M. Balać, S. Ossey, O. Riba-Grofnuz and J. Grübel (2023) Towards an automated, open, and reproducible syn- thetic population of Switzerland, paper presented at the *Swiss Transport Research Conference 2023*.

Microsoft (2023) Assure digital twins, `https://azure.microsoft.com/en-au/products/digital-twins/`. Accessed on 2023-04-28.

Müller, K. and K. W. Axhausen (2011) Hierarchical ipf: Generating a synthetic population for switzerland, *Arbeitsberichte Verkehrs-und Raumplanung*, **718**.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen *et al.* (2015) Promoting an open research culture, *Science*, **348** (6242) 1422–1425.

Raes, L., P. Michiels, T. Adolphi, C. Tampere, T. Dalianis, S. Mcaleer and P. Kogut (2021) Duet: a framework for building secure and trusted digital twins of smart cities, *IEEE Internet Computing*, 1–9.

Roest, M. (2019) An open source platform for digital twins?, `https://www.linkedin.com/pulse/open-source-platform-digital-twins-mark-roest/`. Accessed 30 March 2022.

Schultes, E., M. Roos, L. O. B. da Silva Santos, G. Guizzardi, J. Bouwman, T. Hankemeier, A. Baak and B. Mons (2022) Fair digital twins for data-intensive research, *Frontiers in Big Data*, **5**.

Stark, P. B. (2018) Before reproducibility must come preproducibility, *Nature*, **557** (7706) 613–614.

Swiss Data Science Center (2022) Renku, `https://renkulab.io`. Accessed 28 June 2022.

Tao, F., H. Zhang, A. Liu and A. Nee (2018) Digital twin in industry: State-of-the-art, *IEEE Trans. Ind. Informat.*, **15**, 2405–2415.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne *et al.* (2016) The fair guiding principles for scientific data management and stewardship, *Scientific data*, **3** (1) 1–9.