H. Roßnagel, C. H. Schunck, S. Mödersheim, D. Hühnlein (Hrsg): Open Identity Summit 2020,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2020 133

A Human Digital Twin as Building Block of Open Identity Management for the Internet of Things

Jan Zibuschka[1], Christopher Ruff[2], Andrea Horch[2], Heiko Roßnagel[2]

Abstract: In networked industry, digital twins aggregate product data along the entire life cycle, from design and production to deployment. This enables interoperability between different data sources and analysis functions and creates an integrated data environment. Human digital twins have the potential to create a similarly interoperable and integrated data environment for more user-centric use cases in the field of the Internet of Things. In this case, personal data is processed and transmitted; therefore, the underlying infrastructure is then not product data management but identity management. In this paper, we discuss general aspects of the human digital twin, its role in open identity management systems, and illustrate its application in the field of home, building and office automation. We identify advantages and limitations and suggest future research opportunities.

Keywords: digital twin; internet of things; interoperability; data protection; identity management

1 Introduction

Devices that we use every day are increasingly networked in the so-called Internet of Things (IoT). The communication between these devices poses both interoperability challenges, i.e. whether different devices are capable of exchanging data, and privacy challenges, such as how to control such data exchange [He16; ZHK19]. An identity management infrastructure, a key architectural component in many IoT ecosystem architectures [Ba19; ZHK19], can create an overarching, networked data space that allows both comprehensive analysis of the data and fine-grained control of their exchange.

There are corresponding developments in industrial IoT scenarios. Here, so-called digital twins [AE17], integrated artifacts that aggregate master data pertaining to a device type, observations of sensors at individual device instances, data processing functions, and derived data resulting from processing along the entire life cycle of a product or production machine. There are different approaches for specific implementation, but in general digital twins are intelligent, virtual images of physical devices [AE17]. They have the potential to enable new business models for companies in various verticals and serve as standardized units for cross-organizational data exchange, both in the interaction along the value chain and with regard to the resulting products [KHB18]. These changes affect the entire product life cycle [Ta18]. Besides industrial applications, digital twins

[1] Robert Bosch GmbH, Renningen, 70465 Stuttgart, Jan.Zibuschka@de.bosch.com
[2] Fraunhofer IAO, Nobelstr. 12, 70569 Stuttgart, firstname.lastname@iao.fraunhofer.de

have also found their way into building control systems, where they enable large-scale orchestration of devices [Kh19].

Like the digital twin of a device, the human digital twin has its origin in production technology [Ha20]. In addition to research in this field, where the focus is on capturing the behavior of workers in production environments, there are also applications in medical technology, where the patient's condition is monitored, and and treatment, such as dosage of medications, is controlled using a human digital twin [Ch19]. However, an application of the concept to end users is not covered by existing work.

In the following, we first characterize general characteristics of the human digital twin in a reference architecture, and then describe its application the consumer IoT fields of smart home and building, and how it enables identity management and privacy functions. We then discuss our findings.

2 Human Digital Twins

To characterize a human digital twin, we build on the more general C2PS[3] reference architecture for digital twins [AE17], removing components which are not applicable to humans. We identify the following subsystems, illustrated in Figure 1:

Virtual sensors represent sensors that collect information about the user for storage in the digital twin [AE17]. Unlike the digital twin of a device, sensors do not necessarily have a direct, physical connection to the human. Instead, any sensors that capture information concerning a subject can contribute data to the human digital twin [Ha20].

Observations of these sensors are stored in the digital twin [AE17]. The observations are usually available in a variety of formats and accuracies, and express various contextual data related to the user [Ja17].

Functional units process the information available in the digital twin [AE17]. In contrast to devices, where model-based approaches and physical simulation are central [Ta18] the human digital twin focuses on the empirical, statistical investigation of behavior [Ha20]. Simulations may be possible in cases such as medical applications, based on biological rather than physical processes [Ch19].

Derived knowledge results from this processing, is also stored in the digital twin and can in turn serve as input for subsequent processing steps. In particular, a derivation of user objectives and preferences from observations is possible [Ha20]. Derived events, as specified in C2PS [AE17], form a subset of the broader knowledge about users that a human digital twin can derive, which also includes

[3] Short for "cloud-based cyber-physical systems" [AE17], referring to a digital twin architecture reference model for such systems [AE17]

e.g. location types [KBB18] and user preferences [Ro14].

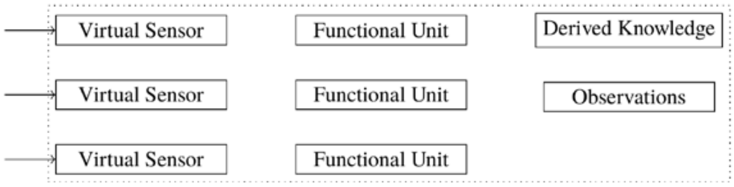

Fig. 1: Human digital twin reference architecture (adapted from [AE17])

In contrast to the device centric C2PS [AE17], a human digital twin does not contain virtual actuators or virtual power supply. The reasons for this are evident: The information system can observe but not control the human being, and the human being does not need to be powered by the electrical system.

Similarly, also in contrast to devices, direct communication between people on the physical level is outside the boundaries of the technical system. Therefore, we assume that the implementation of human digital twins leverages the transmission of observations of user attributes and derived knowledge via virtual connections in the sense of C2PS, which corresponds to federated identity management [Ro14].

3 Fields of Application

3.1 Smart Home

In Smart Home applications, a human digital twin helps to observe and simulate the actions and behavior of residents, guests and other visitors like craftsmen, living or working in the Smart Home environment for a finite timeframe. For this purpose, various sensors, like cameras, motion or pressure sensors may capture relevant data of the people in the home. Further sensors like thermostats, humidity or sound sensors, can measure additional environmental data of the Smart Home. Sensors and electrical switches on lights and the other Smart Home devices record the current state of the devices as well as modifications made by the residents or visitors. The data of the people in the Smart Home environment and the data of the environment itself is necessary in order to gain knowledge about the preferences and behavior of the people in the environment as well as to be able to derive further knowledge about the context of specific situations. Additionally, residents can provide additional context information like data from a personnel calendar or personal preferences like color preferences of the lighting conditions.

Using the captured data in the digital twin, the Smart Home can derive context information, i.e. whether a specific person entering the home is identified and subsequently classified as a resident, guest or an unauthorized person. Additionally, it can use the data to provide general and personalized services or to adjust parameters like room lighting, temperature of a room or the sound volume of devices within the environment to the preferences of one person or even several people in a specific room. Electrically adjustable furniture or appliances can also be adjusted to the required settings or physical needs of an individual or a group of people.

The Smart Home system can use the data about the environment or individuals to learn detecting special contexts or even (medical) emergencies in order to provide special services or to call for assistance. Furthermore, the Smart Home could implement preventive measures, e.g. recommending and supporting a diet for individual residents when detecting weight problems and support individuals in forming healthier (sleeping) habits by adjusting the living environment accordingly

To provide these services and adjust the environment to the needs and preferences of the residents and visitors or to detect the context, the Smart Home system needs to collect a variety of general and personal data from the people within the environment and the environment itself. The collected data has a high risk of abuse if access is not carefully managed and secured. Identification of individuals at the location or data related to medical condition are prominent examples of this risk. Evidently, the personal data, but also the environmental data is very sensitive and should be carefully managed and safeguarded.

The Smart Home system also needs to implement a concept to manage the data transfer between residents and visitors. Disclosure of private data has to be avoided. This also includes actions allowing individuals in the environment to derive personal information about other individuals in the environment, e.g. showing the weight in the smart mirror in the bathroom to a visitor or displaying private appointments to other residents.

3.2 Smart Building

Similar to the application in a Smart Home environment, the concept of a digital twin can be expanded on and subsequently applied to operate smart buildings. However, the focus and central point of the data collection is the individual residing in the building. A smart building constantly produces valuable data, which in combination with the data captured in the human digital twin, can act as decision support for human or machine-operated systems using rule engines or artificial intelligence. The derived knowledge can then be used to adjust various parts of the building’s infrastructure automatically. Several useful applications and benefits of the combination of human digital twins in smart buildings are listed below:

Energy Efficiency

By locating and identifying residents inside the building and thereby identifying rooms that are currently occupied and other facilities in use, the controlling system could dynamically adjust the heating, ventilation or lightning conditions, so that energy resources are efficiently used and managed according to the current requirements.

Individual preferences as well as environmental context data could also be taken into account, to adjust the building conditions to an optimal level.

Automation

The digital twin can enable highly individual but at the same time and transparent forms of building automation by using real time and historic information about occupants as well as the infrastructure and resources. Repetitive or context sensitive tasks like i.e. turning on sprinklers or moving shades according to certain rules or lightning conditions could be easily monitored and configured.

Access Control

Using and combining the captured and derived knowledge about building infrastructure, occupants' location information as well as their respective security clearance levels stored in the digital twin, a control system could automatically grant or deny access to certain resources, facilities, appliances or parts of the buildings in a secure, transparent and convenient manner.

Predictive Maintenance

Similarly, to smart production environments, the context information about building infrastructure and various appliances or machinery in combination with a human digital twin can be used to quickly detect and subsequently repair malfunctions by informing responsible personnel. Furthermore, by using historical data and machine learning, the data can be used to predict were failures or malfunctions are most likely to happen in the future.

3.3 Smart Office

Smart Office systems implement applications from Smart Buildings for energy efficiency as well as security and cost improvements. Additionally, Smart Office solutions contain similar applications like Smart Homes in order to create a pleasant, healthy and motivating working environment. Using a human digital twin in this context can benefit a Smart Office concept by providing additional historic and real time context information to be used to create smart and adaptive office environments.

A Smart Building system, which integrates the Smart Office environment, enables applications like the regulation of heating, lights or air conditioning depending on the number of people in a room. Airing or heating of (meeting) rooms can be prepared in advance, based on contextual data like calendar entries for a room in the booking system of the Smart Office.

Waiting times for elevators can be optimized by analyzing the data of different sensors and predicting the demands on the different floors of the building.

Smart Office environments, which are usually integrated into Smart Buildings, also provide access control systems to control and monitor the access to the building itself or to single departments. Additionally, Smart Offices provide services like the optimization of the utilization of technical infrastructure like servers and other (network) components.

Similar to Smart Home environments, Smart Offices provide applications and services to improve the motivation, work satisfaction and overall well-being of employees and preventing certain illnesses. Examples for such services are the personalization of room settings, e.g. light color, music or room scent, based on the personal preferences of the persons in a room. Functionalities like the automated adjustment of height settings of furniture like tables or chairs depending on the height of the person, who is using it, create an ergonomic office environment in order to avoid postural defects and accompanying diseases.

In summary, the application and utilization of digital twins in smart environments such as smart homes, smart buildings or smart offices offers various benefits and advantages over previous systems. The usage of personal and context related data of individuals, stored and processed in the digital twin can lead to better, more personalized, safer and more efficient applications and use cases of smart environments powered by IoT devices.

4 Identity Management and Data Protection Functions

As described in section 2, human digital twins clearly leverage virtual connections in the sense of C2PS's reference architecture. Thus, we assume identity information, such as observations from sensors and other inputs as well as derived knowledge is transmitted between the digital twins, and not between humans or devices on the physical layer. Standard federated identity management protocols, such as OpenID Connect or SAML are suitable for this [Ro14]. The information can be linked to an identifier of the specific human digital twin to enable traceability. The human digital twin can also enable pseudonymization, acting as a privacy-protecting identity intermediary [Ra07] in such a case. This also enables mapping between distributed clusters of human digital twins with different identifier regimes. Information can also be de-aggregated or even anonymized by functional units within the human digital twin [Ra07].

In general, it is notable that human digital twins are especially suitable for enabling key data protection goals via their functional components, leveraging the holistic perspective on users' personal information immanent in the paradigm. To illustrate this, consider the implementation of the privacy dashboard pattern for information in a human digital twin depicted in Figures 2 and 3.

Privacy dashboard are widely used for offering transparency, intervenability, and accountability for users' personal information in corporations' systems [ZAM14], and thus support the majority of privacy goals [HJR15]. As the digital twin gives access to both all observations and derived knowledge about a user, and functions to modify this information, all that is needed is an interface connecting the dashboard frontend to the human digital twin's functional units and data stores. The functions in such a privacy interface would include retrieving observations and/or derived information about a user from a digital twin's data stores. The interface also connects to the digital twin's functional modules for updating the information – or at least correcting it, if it is wrong, and deleting personal information on request of the user – or at least blocking it in case deletion is not feasible [ZHK19].

In scenarios where personal information is not linked to identifiers, human digital twins can be equipped with functionality to identify users from observations about them, in essence making the instance of human digital twin an observation is assigned to and associated identifier derived knowledge in the sense of the reference architecture. For example, a system can use biometrics based on camera information to identify a user, and then direct observations from other sensors also capturing information about the user to the appropriate human digital twin. This can enable more precise analytics, but can also improve users' privacy as it enables targeted transparency, enabling e.g. a dashboard implementation that does not give all users access to all information [ZAM14], which is a privacy issue in itself.

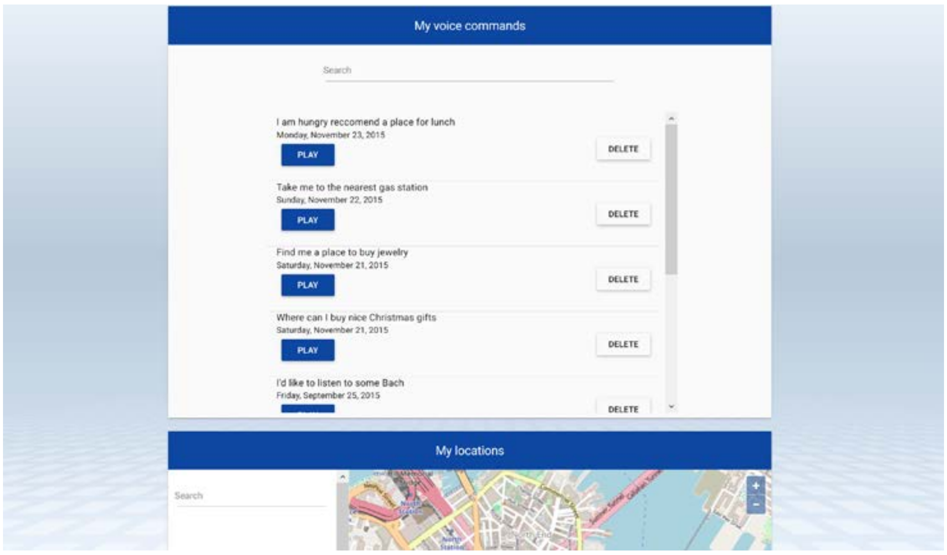

Fig. 2: Privacy dashboard for observations stored in a human digital twin

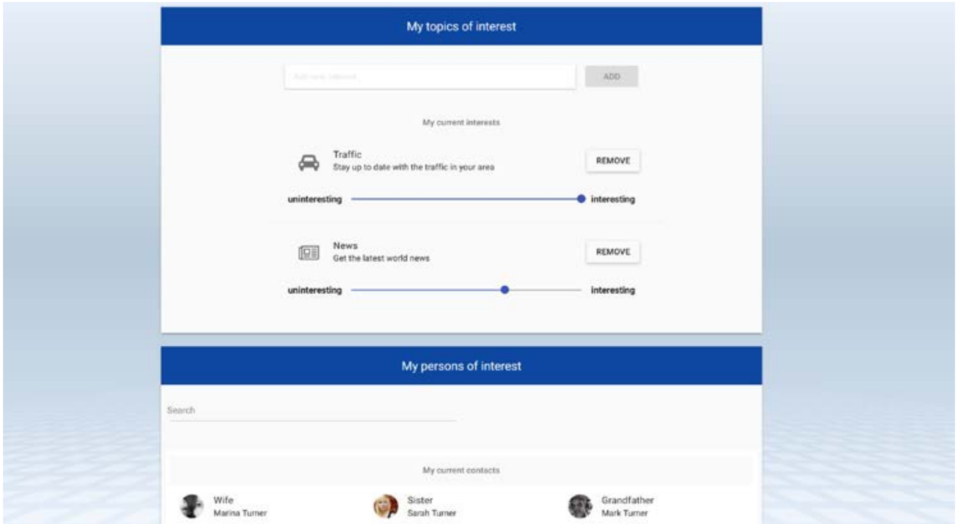

Fig. 3: Privacy dashboard for derived knowledge stored in a human digital twin

Beyond that, digital twins have also been successfully used for anomaly detection [GV17]. Human digital twins could likewise be used to spot unauthorized access to information or physical assets, or to detect anomalous human behaviour indicative of health issues, a common use case in ambient assisted living [Ja17].

5 Discussion

While a human digital twin for end users is a promising concept, some complexities are foreseeable in practice, leaving room for further research. In consumer scenarios such as home automation, a fragmentation of the digital twin by silos of different manufacturers [Ja17] is to be expected. Unlike in industrial and building scenarios, there is no obvious integrator role. This could coincide with the building management; it could be the end user, a device manufacturer, or a dedicated integrator. This also causes challenges in interoperability. The use cases implemented by intelligent home automation vary greatly for different manufacturers [Ja17], which makes it difficult to define certain functional components and data using the reference presented here, and calls into question whether we can reach semantic interoperability before the features offered by such systems have converged.

In many instances, a human digital twin in the field of intelligent household appliances poses significant data protection challenges, as an integrated representation can involve aggregating raw data from the most intimate spaces. As we described, they also enable a high degree of data sovereignty. Collected information is bound to be critical personal information, thus, data minimization and unlinkability, the protection goals not addressed by privacy dashboards, should be a target of system design. Depending on the characteristics of the integrator and the implemented use cases, different approaches are conceivable here. Therefore, solutions addressing the issue are out of the scope of this reference architecture.

Summing up, digital twins are a proven concept in networked production systems, which as we illustrated also holds a lot of promise for the field of intelligent home automation. Human digital twins are a logical extension of the concept, and address core challenges in the Internet of Things, enabling various use cases, interoperability, and key privacy functions.

Bibliography

[AE17] Alam, K.M.; El Saddik, A.: C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. IEEE Access 5, pp. 2050– 2062, 2017.

[Ba19] Bauer, J.; Hoffmann, H.; Feld, T.; Runge, M.; Hinz, O.; Mayr, A.; Förster, K.; Teske, F.; Schäfer, F.; Konrad, C.; Franke, J.: ForeSight - Platform Approach for Enabling AI-based Services for Smart Living. In: How AI Impacts Urban Living and Public Health. Bd. 11862, Springer International Publishing, Cham, pp. 204–211, 2019.

[Ch19] Chakshu, N.K.; Carson, J.; Sazonov, I.; Nithiarasu, P.: A semi-active human digital twin model for detecting severity of carotid stenoses from head vibration—A coupled computational mechanics and computer vision method. International Journal for Numerical Methods in Biomedical Engineering 35/5, 2019.

[GV17] Grieves, M.; Vickers, J.: Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Transdisciplinary Perspectives on Complex Systems. Springer International Publishing, Cham, pp. 85–113, 2017.

[Ha20] Hafez, W.: Human Digital Twin: Enabling Human-Multi Smart Machines Collaboration. In: Intelligent Systems and Applications. Bd. 1038, Springer International Publishing, Cham, pp. 981–993, 2020.

[He16] Hernández-Serrano, J.; Muñoz, J.L.; Bröring, A.; Esparza, O.; Mikkelsen, L.; Schwarzott, W.; León, O.; Zibuschka, J.: On the Road to Secure and Privacy-preserving IoT Ecosystems. In: Interoperability and Open-Source Solutions for the Internet of Things. Springer, Cham, pp. 107–122, 2016.

[HJR15] Hansen, M.; Jensen, M.; Rost, M.: Protection Goals for Privacy Engineering. In: 2015 IEEE Security and Privacy Workshops, San Jose, CA, pp. 159-166, 2015.

[Ja17] Jakobi, T.; Ogonowski, C.; Castelli, N.; Stevens, G.; Wulf, V.: The Catch(es) with Smart Home: Experiences of a Living Lab Field Study. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM Press, pp. 1620–1633, 2017.

[Kh19] Khajavi, S.H.; Motlagh, N.H.; Jaribion, A.; Werner, L.C.; Holmstrom, J.: Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings. IEEE Access 7, pp. 147406–147419, 2019.

[KBB18] Karatzoglou, A.; Koehler, D.; Beigl, M.: Purpose-of-Visit-Driven Semantic Similarity Analysis on Semantic Trajectories for Enhancing The Future Location Prediction. In: 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, pp. 100-106, 2018.

[KHB18] Klostermeier, R.; Haag, S.; Benlian, A.: Digitale Zwillinge – Eine explorative Fallstudie zur Untersuchung von Geschäftsmodellen. HMD Praxis der Wirtschaftsinformatik 55/2, pp. 297–311, 2018.

[Ra07] Radmacher, M.; Zibuschka, J.; Scherner, T.; Fritsch, L.; Rannenberg, K.: Privatsphärenfreundliche topozentrische Dienste unter Berücksichtigung rechtlicher, technischer und wirtschaftlicher Restriktionen. In: Wirtschaftsinformatik (1), pp. 237-254, 2007.

[Ro14] Roßnagel, H.; Zibuschka, J.; Hinz, O.; Muntermann, J.: Users' willingness to pay for web identity management systems. European Journal of Information Systems 23/1, pp. 36–50, Jan. 2014.

[Ta18] Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F.: Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology 94/9-12, pp. 3563–3576, Feb. 2018.

[ZHK19] Zibuschka, J.; Horsch, M.; Kubach, M.: The ENTOURAGE Privacy and Security Reference Architecture for Internet of Things Ecosystems. In: Open Identity Summit 2019. Gesellschaft für Informatik, Bonn, pp. 119–130, 2019.

[ZAM14] Zimmermann, C.; Accorsi R.; Müller, G.: Privacy Dashboards: Reconciling Data-Driven Business Models and Privacy. In: 2014 Ninth International Conference on Availability, Reliability and Security, Fribourg, pp. 152-157, 2014.