

The Use of Digital Twins to Overcome Semantic Barriers in Hyperconnected Ecosystems for Industry

Frank-Walter Jaekel¹, Patrick Gering¹ and Thomas Knothe¹

¹ Fraunhofer Institute for Production Systems and Design Technology IPK, Pascalstr. 8-9, D-10587 Berlin, Germany

Abstract

To establish business networks a match between businesses demands and potential partner information is required. Publicly available information on the Internet about companies, products and services usually don't follow a common standard. The concept of a digital twin could be used to organise the different information and, in the future, to harmonise the way company data is made available on the web. Every company usually has a web presence, related documents, web pages and a trace on the web, which can be used for an initial structure of the digital twin. Hereby, first services for the correlations between partner companies and requirements can be designed. But it requires the management of legal aspects e.g. the access of bots to the public available information. The paper provides initial ideas and feasibility checks and it propose an evolution of the current heterogeneous content and structure of the data into a well-structured digital twin including content related ontologies to describe the company characteristics.

Keywords

Hyperconnected, ecosystems, digital twin, ontology, supply chain, semantic

1. Introduction

Supply chains suspended, the availability of raw materials becomes more difficult, new laws demand high transparency across the entire supply chain. The need for high flexibility of sourcing and reaction on demands from the market requires not only agile, adapting processes within manufacturing, but also cross-functional cooperating manufacturing networks in particular.

The Hyperconnected Ecosystems for industrial networks should enable the networking of any relevant information and its accessibility at anytime from anywhere (hyperconnected). Barriers between network partners are reduced and aligned with current requirements. The evolution of the network will be enabled by its configurability and flexible integration of services. Ultimately, any required information should be immediately available in a usable form and without effort at any location. In this way, the network has similarities to social networks in which partners find each other and exchange data and services as needed.

This blurs the boundaries between the partners in the network and creates the need for the use of data from machines, products and processes across the network without delay. At the same time, the range for tracing the paths in the network increases, as the supplier's connections can also be identified and mapped as a network. This incorporates services providers for logistic, IT and finance as well as legal services. Assessments can be carried out quickly in this way, for example to identify deviations from the legal framework. However, the industrial requirements for security, sovereignty and transparency must be met.

Proceedings of the Workshop of I-ESA '22, March 23–24, 2022, Valencia, Spain

EMAIL: frank-walter.jaekel@ipk.fraunhofer.de (F.-W. Jaekel); patrick.gering@ipk.fraunhofer.de (P. Gering); thomas.knothe@ipk.fraunhofer.de (T. Knothe)

ORCID: 0000-0003-4846-005X (F.-W. Jaekel); 0000-0002-3055-7155 (T. Knothe)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

The construction of the "hyperconnected ecosystem for industry" faces, among other things, the challenge of the existence of a difficult-to-manage number of platforms, digital services, cloud solutions and services, which also include enterprise platforms for products and services. Interoperability usually only takes place within a solution and can only be realized across the board with a high translation effort. This makes the targeted search for and networking with new providers time-consuming and difficult. Solutions for specific markets already exist and use smart methods to interrelate demands and offers of their partners [1, 2].

The "hyperconnected ecosystems for industrial networks" aim at fast communication and flexible data use in dynamic enterprise networks and thus in particular at transparent and flexible supply relationships. The technological basis for this is provided by approaches from the field of social networks but requires effective tendering mechanism and interoperability [3]. Especially the correlation between demand and offer needs to be supported as well as the understanding of the provided data within communication along the network. Even if there is no consensus of digital twin definitions, the formalisation in terms of ontologies, models related data sources and managing bidirectional changes of data might improve of interoperability between the network partners.

This incorporates not only digital twins of products and services but also digital twins of the factory or company in general. The twin might be represented in a first attempt by a semantical enriched webpage like considered by semantic web approaches [4]. In the future it should evolve in terms of more behaviour details and cross references to experiences and evaluations of the partner e.g. by others references and sources describing the detailed work of the company from their customer side. This way would allow an evolution from a very simple twin provided by the company to a more and more enriched twin.

The Hyperconnected Ecosystem for industrial networks assumes a networking of independent companies. This creates semi-automatic connections between the network partners. For example, a partner can find other partners via defined properties and then address them specifically. The retention of sovereignty over the data to be transmitted (data ownership) must be ensured. In addition, services can be provided as well as searched for and embedded in one's own environment. Here the role of the digital twin can support the clear data ownership on the partner company side in case the company provides a digital twin. This includes the update of the data and control about the data published through the network.

Approaches such as the Open Platform Communications Unified Architecture (OPC-UA) [5], the international data space (IDS) [6], Asset Administration Shell (AAS) [7] and BaSys/BaSyx [8] provides already interface definitions and related information model structures. In fact in BaSyx the AAS implementation is addressed as a digital twin [9]. However specific semantic or ontologies covering behaviour, reaction times, laws, business indicators as well as product information needs to be correlated across different sources such as standards and quasi standards e.g. ECLASS (eCI@ss) [10], OPC-UA companion specification.

2. The digital twin within the network

To date, no consensus definition of a digital twin has been established in scientific publications. While digital twins are defined on the one hand for final products [11, 12] or entire product life cycles [13, 14]. Gartner coined the term Digital Twin of an Organization (DTO) and no longer limits digital twins to physical objects. The vision is to create a digital image of an entire organization, its processes including infrastructure, roles and responsibilities, products, etc. [15, 16].

Following the idea, we refer to van der Aalst's definition of the digital twin and the classification of the digital shadow and master, which replaces the physical object with reality and removes the restriction to physical objects [16].

According to the Fraunhofer Institute for Production Systems and Design Technology (IPK), the digital master is described by reference models and information of products and process originating from development phases. The digital shadow results in the continuously recording, storing and provisioning of operational data. Only the intelligent linking of the digital master with the digital shadow results in the digital twin [17].

In the paper we see the “partner company twin” as a digital representation of the available information of a legal entity usually a company or parts of a legal entity such as factories of a group.

According to a digital twin within the paper, it consists of a digital master and the related digital shadow of one or more real objects and its relationship. This includes their products, legal status, business activities etc. (Figure 1) and comes close to the idea of the DTO. In this case the master provides a structure of the twin in terms of references to data sources, models, ontologies, etc.. The digital twin represents the digital representation of the company by providing the data requested by the digital master and linking it to the transaction data and traces on the web in the sense of a digital shadow. The twin should represent the current status of the company. The digital representation of the company can evolve from a simple web page to a more comprehensive twin. A challenge is the bidirectional interaction between company data and the digital twin, because changes in the real world should directly trigger a feedback in the digital world.

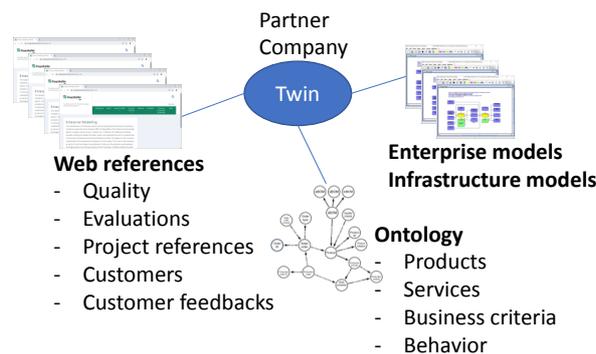


Figure 1: Concept of company twin representing network partner

The twin should allow the interrelation of this information. In an initial attempt this is just a profile covering characteristics of a network partner. The characteristics are just defined terms with structure, synonyms and relations, which might evolve to ontology but also consider different existing ontologies.

The profile can be filled in manually, but in discussions with stakeholders, especially small and medium-sized enterprises, they did not agree with the effort involved. Therefore, we also considered an automatic search and data retrieval mechanism that is described in the technical part of the paper.

The “partner company twin” is now composed of instantiated characteristics and references to data sources. This creates a library of potential partners of a business network. They can be located in different web or cloud infrastructures. Network relationships and connections can be established by business demands or by data requests.

This should be supported by a tendering service to find appropriate business partners. A demand needs to be described e.g. a required part of a product or a required service. Afterwards a search takes place to find correlation between the different partners on the characteristics described. In the future this might be automatically if a network partner disappears or is not able to deliver in the timeframe. Also new partners can be considered easily and reconfigured into the network (Figure 2).

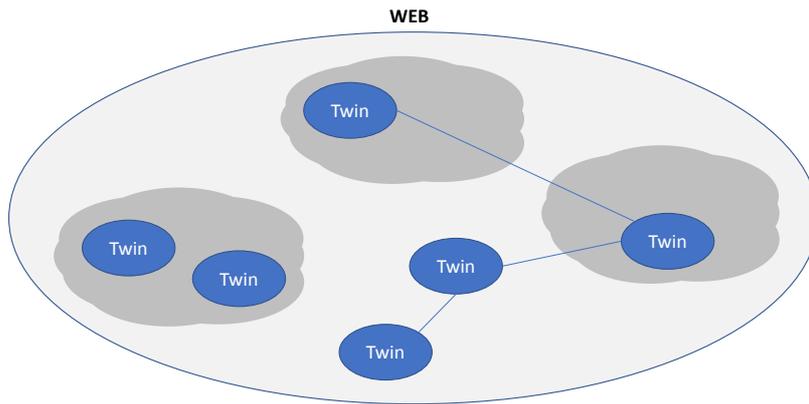


Figure 2: Rough concept of network of twins across platforms and cloud

3. Technologies and current work

Available technologies to realize the describe concept of “hyperconnected ecosystems for industry” using “partner company twins” are communication technologies such as OPC-UA, MQTT or just REST protocols, cloud infrastructures, concepts such as international data space, asset administration shell, FIWARE[18], I4Q [19], GAIAX [20], BASYx which we analysed. They provide a very good basis and also initial and open toolsets such as BASYx [9]. However, we did not find “the approach” because of lots of technical dependency related to programming languages, developing environments and platforms. A good connectivity between machines and enterprise application can be achieved if OPC-UA is used and a common companion specification but every system needs these OPC-UA interface. BASYx allows different interface options such as OPC-UA, MQTT etc. but you need a BASYx implementation following the AAS concept. Therefore, it is difficult to decide which technology will be persistent in the future.

So, we followed a more generic approach with small exchangeable services and connecting different “partner company twins”. The feasibility is still a challenge and we work currently on related feasibility tests. In parallel with the user interfaces being developed with industry partners, we have identified the following services for initial feasibility testing:

1. Matching of demand and partner characteristics
2. Semi-automatic filling of partner profiles
3. Identification of requested data in text

The matcher correlates simply a demand with different characteristics across a set of profiles. The design of the matcher relies on REST interfaces. The demand is based on logical statements but also the characteristics within the profiles can have logical interdependencies. The matcher is configured by properties and the related logic. It takes a parameter set combined with logical equations and check it against a set of profiles. A future extension of the matcher is the use of ontologies for the correlation. This is especially related to different descriptions of products and materials.

The service to fill the profiles relies currently on technologies such as web crawler to analyses web addresses and scraping. After initial tests we encounter that most of the supplier pages are permitted concerning automatic data analysis (robots.txt). Therefore, we have requested a special approval for the scraping tool before the analysis proceed. The crawler was finally less needed because we will get specifically the web address of the partner. The interest of the partner companies is the reduction of work about maintaining a profile. This supports the decision to take the actual data directly from their web pages.

The work with the scraper and the different web pages provides initial results. However, formats are quite different on the web pages, semantic web guidelines are less considered and additional information are provided in text form such as product specifications. Therefore, the text analysis with the Industrial-Strength Natural Language Processing (spaCy) [21] has been considered. It provides a grammatical analysis of the text to identify for example relationships between criteria and related values. The example in Figure 3 illustrates the usage and the language independency. The sentence on

the left side “The diameter is only 20 mm despite its great performance” is analyzed using the profile template. The result is a row of the profile with the requested data.

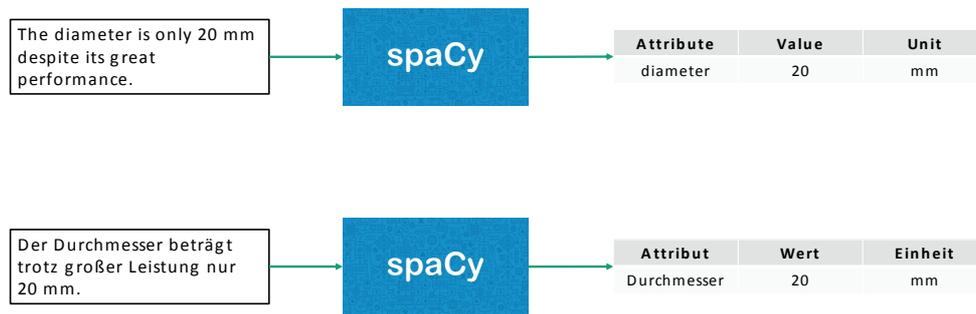


Figure 3: Text analysis

The tool chain realized by the described services above are expressed in Figure 4. The Crawler might be used if the partner urls are not available. Otherwise the target page is already selected by the url of the considered partner company. The scraper uses the target page to extract the detailed information of the partner company. In case of text documents spaCy will be in place to support detailed analysis regarding missing values for the profile. The data are finally introduced into the profile. However, if data is missing further activities might be initialized otherwise they will be kept empty. In case of further characteristics which might be encounter because of property tables they will be added to the profile for further use. The analysis requires a minimum structure of the text on the web page in terms of property tables. In terms of text the grammar needs to be correct and the text should be easy to read. The provided information should be consistent because currently no constancy check is in place but in the future it should be implemented. This tool chain will fill the profiles and also keep the profiles actual as far as the provided information will act like a digital twin of the partner company and actualizes the data accordantly to changes in the real work of the company.

Having the profiles available they can be used for tendering or communication access between the partners. This is supported by the matcher which can be configured with actual business rules and parameter types. The matcher takes the demand and find matches to create partnerships. Logical relationships and use of the characteristics are defined in the configuration of the matcher.

The described technology is one possible usage of “partner company digital twins”. It requires a clear description of company information as well as the connection between real world company business and its digital representation. This can be done through specific processes for regularly updating company information on the web in the form of manual activities or interfaces to the company's IT systems.

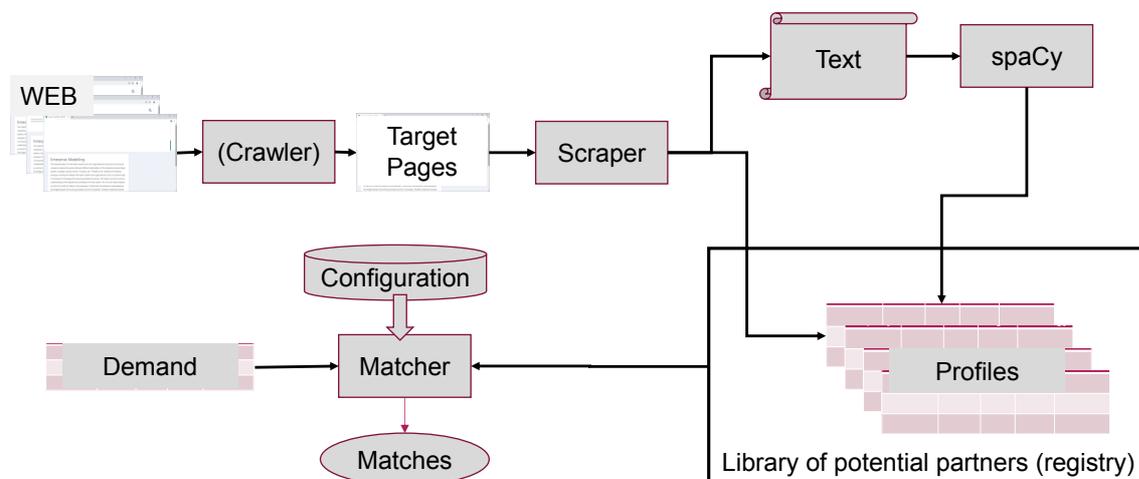


Figure 4: Tool chain about data collection and matching

4. Conclusion and outlook

Digital twins can provide a formal description in terms of ontologies, models and data relationships describing the company and its products as well as in term of its shadow e.g. evaluations and project results in the web. An example of the usage of these Digital Twins has been described in chapter 3 in terms of a feasibility study for data collection and match making. This is an important feature within a hyperconnected ecosystems to find partners and establish networks.

Currently only few organisations provide the data in a form of the digital twin. Therefore, also unstructured and incomplete data has to be considered for the correlation between demands and providers. The unstructured data appears as one of the barriers for automatic data processing because finally it cannot process full automatically. It still requires a final check by experts if the data is correctly identified.

The described work in this paper represents concepts and feasibility checks related to a hyperconnected ecosystem for industry starting with initial services such as matching between demand and partner characteristics. The next steps will focus on the real usage of the technology in terms of getting information from supplier web pages demanded by customers. We already start cooperation with companies in terms of the related analysis of their web pages. Also, profiles on the customer side has been drafted. The target is to archive a technical readiness level which will allow an initial distribution of the services.

However, a presumption still needs further consideration, the “partner company digital twin” currently the available digital twins are poor. So, it will need more work on developing technologies to simplify the provision of partner company digital twins.

5. Acknowledgements

Foundations of the described work has been developed in the scope of a project of the Werner von Siemens Centre (<https://wvsc.berlin/>) WvSC.EA “Electric motors 2.0” supported by the European Regional Development Fund (EFRE). The described services benefit in terms of usage partially from discussions within WvSC.EA.DI05 with the following partners Siemens AG, pi4_robotics GmbH, Chair of Service-centric Networking (SNET) from the Technische Universität Berlin, TresCom Technology, 5thIndustry and FhG IPK.

6. References

- [1] Globality, Smart sourcing, 2022. URL: <https://www.globality.com/en-us>.
- [2] Facturee, Online Manufacturing, 2022. URL: <https://www.facturee.de/en/>.
- [3] ASSURANT, Enabling Interoperability the key to a Connected Ecosystem, 2020. URL: <https://www.assurant.co.uk/newsroom-detail/Features/2020/September/enabling-interoperability-the-key-to-a-connected-ecosystem>.
- [4] W3C, Semantic Web, 2022. URL: <https://www.w3.org/standards/semanticweb/>.
- [5] OPC-UA, Welcome to the World of OPC, 2022. URL: <https://opcfoundation.org/>.
- [6] International Data Space Association, Reference Architecture Model, Version 3.0, 2019. URL: <https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf>.
- [7] BMWI, Details of the Asset Administration Shell, 2022. URL: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.html.
- [8] BASYS4.0, Basissystem Industrie 4.0: Eine offene Plattform für die vierte industrielle Revolution, 2022. URL: <https://www.basys40.de/>.
- [9] The Eclipse Foundation, BASYSx, 2021. URL: <https://www.eclipse.org/basyx/?target>
- [10] eCl@ss, Enable your global business and digitalization, 2022. URL: <https://www.eclass.eu/en/index.html>.
- [11] G. N. Schroeder, C. Steinmetz, C. E. Pereira, D. B. Espindola, Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange, IFAC-PapersOnLine 49 (2016) 12-17. doi: 10.1016/j.ifacol.2016.11.115.

- [12] M. Abramovici, J. C. Göbel, H. B. Dang, Semantic data management for the development and continuous reconfiguration of smart products and systems, *CIRP Annals* 65 (2016) 185-188. doi: 10.1016/j.cirp.2016.04.051.
- [13] R. Rosen, G. von Wichert, G. Lo, K. D. Bettenhausen, About The Importance of Autonomy and Digital Twins for the Future of Manufacturing, *IFAC-PapersOnLine* 48 (2015) 567-572. doi: 10.1016/j.ifacol.2015.06.141.
- [14] T. Gabor, L. Belzner, M. Kiermeier, M. T. Beck, A. Neitz, A Simulation-Based Architecture for Smart Cyber-Physical Systems, in: 2016 IEEE International Conference on Autonomic Computing (ICAC), IEEE, New York, 2016, 374-379. doi: 10.1109/ICAC.2016.29.
- [15] M. Kerremans, J. Kopcho, Create a Digital Twin of Your Organization to Optimize Your Digital Transformation Program, 2019. URL: <https://www.gartner.com/en/documents/3901491/create-a-digital-twin-of-your-organization-to-optimize-y>.
- [16] W. M. P. van der Aalst, Concurrency and Objects Matter! Disentangling the Fabric of Real Operational Processes to Create Digital Twins, in: A. Cerone, P. C. Ölveczky (Eds.), *Lecture Notes in Computer Science, Theoretical Aspects of Computing*, Springer, Cham, 2021, pp. 3-17. doi: 10.1007/978-3-030-85315-0_1.
- [17] T. Knothe, K. Lindow, C. Geisert, Gut vernetzt ist halb gewonnen. Um Digitale Zwillinge gewinnbringend einzusetzen, müssen produzierende Unternehmen wissen, wie sie diese sinnvoll vernetzen können, 2021. URL: <https://www.ipk.fraunhofer.de/de/publikationen/futur/futur-2021-2/gut-vernetzt-ist-halb-gewonnen.html>.
- [18] FIWARE, FIWARE: The Open Source Platform for Our Smart Digital Future, 2021. URL: <https://www.fiware.org/>.
- [19] I4Q, Industrial Data Services for Quality Control in Smart Manufacturing, 2021. URL: <https://www.i4q-project.eu/>.
- [20] GAIAX, About Gaia-X, 2021. URL: <https://www.gaia-x.eu/what-is-gaia-x>.
- [21] SpaCy, Industrial-Strength Natural Language Processing, 2022. URL: <https://spacy.io/>.