A Multi-Perspective Data Model for Cyber Physical Production Networks

Ada Bagozi[1], Devis Bianchini[1] and Anisa Rula[1]

[1]University of Brescia, Brescia, Italy

Abstract

Recently, the research on data management is moving towards the design of data models for Cyber Physical Production Networks. They constitute ecosystems where cyber and engineered physical elements record data (e.g., using sensors), analyse them using connected services (e.g., over cloud computing infrastructures) and interact with human actors using multi-channel interfaces, going beyond the boundaries of a single enterprise and spanning over the whole production network. In this paper, we propose a conceptual data model that relates the digital counterpart of a product with data collected over the phases of the product lifecycle, with special attention on the manufacturing process, and with data gathered to monitor the shop floor machines used during the production. This enables the development of multi-perspective data services that implement advanced business goals at the production network level, such as production scheduling, energy efficiency, product and process monitoring. We provide an architecture of the proposed solution and its preliminary validation in an ongoing industrial research project.

Keywords

Multi-perspective data model, service-oriented architectures, Cyber Physical Production Network, Industry 4.0

1. Introduction

The Industry 4.0 revolution is shifting the innovation focus from traditional crafts and engineering, such as mechanical engineering, material science and plants construction, to immaterial assets, like knowledge extraction and data analysis and exploration. The latter ones are still mainly focused on the design of Digital Twins of machines or parts in isolation, as virtual representations of physical assets in a Cyber-Physical System. Recently, research on Digital Threads is gaining momentum, conceived as the cyber side representation of a product and the context where the product is designed, produced, used and maintained, to enable the holistic view and traceability along its entire lifecycle [1, 2].

However, this vision is focused on the digital representation of the product only, while its full potential is still far from the sphere of awareness. Heterogeneous and not interconnected data comes from the shop floor level, all along the product lifecycle, to ensure the production of costly and complex products, high product quality levels, long lasting functioning, less frequent and efficient maintenance activities and scalable performance over time. A pressing demand has been raised to process and integrate those data by exploiting service-oriented architectures

SEBD 2022: The 30th Italian Symposium on Advanced Database Systems, June 19-22, 2022, Tirrenia (Pisa), Italy
ada.bagozi@unibs.it (A. Bagozi); devis.bianchini@unibs.it (D. Bianchini); anisa.rula@unibs.it (A. Rula)

and web-based systems, leading to the notion of Cyber Physical Production Network (CPPN). A CPPN is an ecosystem where cyber and engineered physical elements collect (e.g., using sensors) and analyse data using connected services (e.g., over cloud computing infrastructures). It allows the interaction with human actors using multi-channel interfaces, going beyond the boundaries of a single enterprise and spanning over the whole production network.

In this paper, we propose an information model that relates the digital counterpart of a product with data collected over the phases of the product lifecycle, with special attention on the manufacturing process, and with data gathered to monitor the shop floor machines (work centers) used during the production. The information model is represented according to multiple perspectives, namely the production process, the industrial assets and the product, used to organise the data collected along the product lifecycle. The approach also enables the development of a portfolio of data-oriented services on top of the model, properly tailored on the perspectives that can be composed to implement advanced supply chain functionalities such as production scheduling, energy efficiency, product and process monitoring. Finally, we provide an architecture of the proposed solution and its validation in an industrial research project, to demonstrate its effectiveness for a particular case of the remote monitoring service to detect near real-time anomalous events.

The paper is organised as follows: in Section 2 the multi-perspective model is described; Section 3 introduces data-oriented services design; in Section 4 implementation and experimental validation are discussed; comparison with related work is presented in Section 5; finally, Section 6 closes the paper.

2. Multi-perspective information model

The proposed multi-perspective information model is shown in Figure 1. The model is organised according to the following perspectives.

Product perspective. The product perspective concerns product configuration over the Product Lifecycle Management (PLM). Each product (in turn described through product details) is composed of a set of parts which are identified by a part code and are hierarchically organised into different BoM, depending on the specific phase of the product lifecycle that is being considered (e.g., Engineering BoM, Manufacturing BoM). Part codes belonging to different hierarchies can be connected to each other, e.g., in order to manage changes propagation over the product lifecycle. We focus on the MBoM, which concerns the manufacturing process and reports the distinction between: (i) part codes that must be produced internally, connecting to a Production Order (PO); (ii) part codes that must be produced externally, by one of the suppliers of the supply chain, connecting to a Contract Work Order (CWO); (iii) part codes to be bought and then assembled in the final product, connecting to a Purchase Order. The final product, corresponding to the root of the BoM hierarchy, is associated with the PO as a Sales Order item.

Process perspective. In the process perspective, the main entity is represented by the production phase, that is in turn organised according to a hierarchy, where each phase is composed of a set of sub-phases, reaching a level of detail ranging from macro to micro industrial processes. Production phases implement: (i) the PO of the final product, connected to the Sales Order item; (ii) the PO of one or more components of the final product, connected to

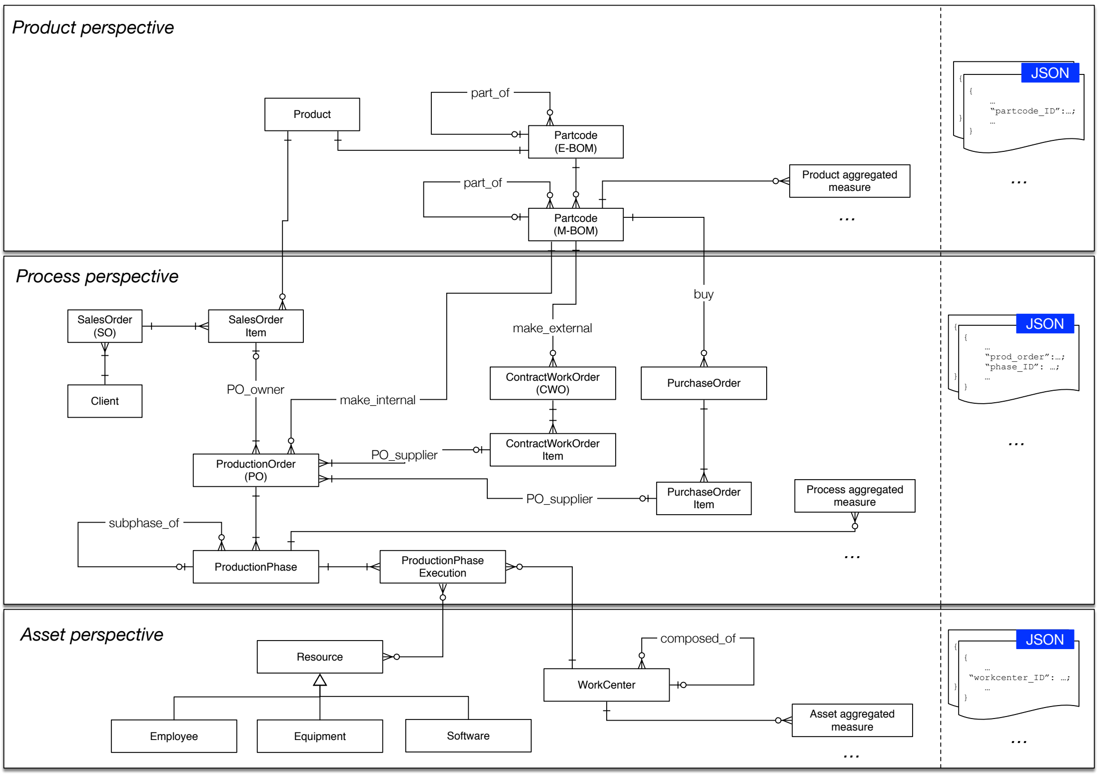

Figure 1: Multi-perspective information model of the Cyber Physical Production Network.

the CWO item or to the Purchase Order item and corresponding to the point of view of one of the suppliers in the supply chain. In this way, the information model brings together the viewpoints of all the supply chain actors. In this case, the hierarchy of process phases is focused on the production step in the PLM, thus corresponding to the MBoM. However, this modelling can be seamlessly extended to the other phases of the PLM (e.g., engineering, maintenance) and the other kinds of the BoMs.

Asset perspective. The asset configuration perspective includes the resources involved in the realisation of the product (machinery, equipment, information systems, human resources). Work centers are the machines used in the manufacturing process and are hierarchically organised according to the RAMI 4.0 Reference Architectural Model for Industry 4.0 [3]. A work center or a resource can be involved in many production phases during its lifecycle. A production phase can be executed and distributed on several work centers and may consume/require different resources.

Parameters. Parameters are distinguished among aggregated measures (e.g., KPI) and data streams collected as-is from the field (e.g., sensors data acquisition). In case of aggregated measures a new entity is created for each measure containing records with the calculated value, a time reference (depending on the granularity of the aggregation, e.g., monthly/weekly/daily and so forth) and relationships with the target entities of the multi-perspective data model. The data stream collected from a vibration sensor on a specific work center or component is an example of this kind of parameter.

3. Data-oriented services design

A service repository is designed on top of the multi-perspective data model. Services are categorised according to different service types and over the three perspectives of the product, process and assets. Concerning service types, we distinguish among: (i) *collect services*, used to transfer data from the physical side of the production network towards the data storage on cloud; (ii) *monitor services*, used to proactively identify/predict warnings/errors about the phenomena occurring on the physical production network [4]; (iii) *dispatch services*, used to share data across the actors of the production network; (iv) *display services*, exposed to visualise data on the dashboard. We formally define a multi-perspective data-oriented service as follows:

Definition 1 (Data-oriented service). *A data-oriented service $\mathcal{S}_i$ is described as a tuple*

$$S_i = \langle n_{S_i}, t_{S_i}, P_{S_i}, url_{S_i}, m_{s_i}, IN_{S_i}, OUT_{S_i} \rangle \quad (1)$$

where: (i) n_{S_i} is the name of the service; (ii) t_{S_i} is the service type, namely: collect, monitor, dispatch, display; (iii) P_{S_i} is the set of perspectives on which the service is focused (product, process, assets); (iv) url_{s_i} is the endpoint of the service for it invocation; (v) m_{s_i} is the HTTP method (e.g., `get`, `post`*) used to invoke the service; (vi) IN_{S_i} is the representation of the service input; (vii) OUT_{S_i} is the representation of the service output. We denote with $\mathcal{S}$ the overall set of data-oriented services.*

Services in $\mathcal{S}$ are implemented as RESTful services; therefore, they are also described through HTTP method (e.g., `post`, `get`, `put` and `delete`) implemented in the service, corresponding to the Create-Read-Update-Delete actions, and the resource that is being created, read, updated or deleted, respectively. A multi-perspective service S_i deals with different types of data, based on its input/output parameters. Among the inputs/outputs of S_i we could mention entities from the process, product and assets perspectives of the data model. Moreover, input/output data can be parameters (i.e., calculated measures or data-streams collected from the physical side of the production network). Services dealing with data-streams are called data-intensive services and need advanced techniques to deal with data-streams. For instance, a predictive maintenance service is a data-intensive monitor service if it detects concepts drifts from massive data streams collected on the monitored machines. Service inputs/outputs are represented in JSON format and modelled in terms of information model entities and their relationships.

4. Approach implementation in a real case study

The approach presented in this paper guided the implementation of the architecture presented in Figure 2. The proposed architecture is composed of four main parts: (i) the production network, including the process phases and actors involved in each phase; (ii) the services repository, integrating the set of services aimed at implementing the objectives of actors in the supply

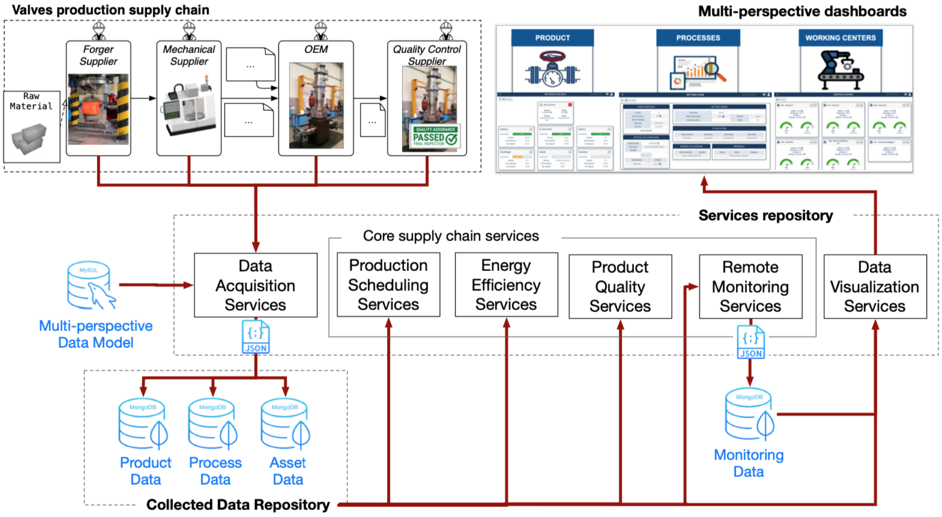

Figure 2: Approach architecture.

chain; (iii) the data repository, containing parameters values coming from different sources and owned by different actors; (iv) a multi-perspectives dashboard, for detailed exploration of data according to the multiple perspectives.

For the production network, we considered the real case study about the production of valves to be used in deep and ultra-deep water applications. Valves are placed in prohibitive environments and, once installed, are difficult to remove and maintain over time and require high quality levels. Valves supply chain involves several actors to produce different sub-parts of the valve: the valves producer, who represents the OEM of the production process; the forger, who is in charge of transforming the raw metals into workable pieces; the mechanical supplier, who is in charge of modelling pieces provided by the forger to be assembled in the final product; the product quality test company, who is in charge of performing quality tests on the final valves. The parameters values are collected from the production network and stored in data repository, thanks to sensors and IoT technologies, or from legacy systems owned by involved actors (e.g., ERP, PDM). Data streams are saved within a NoSQL database (e.g., Product Data, Process Data and Asset Data) as JSON documents using MongoDB technology. Each JSON document represents a record of collected measures within a set of parameters and includes the timestamp and a reference to the perspective in which the measures have been collected. Supply chain actors have different goals, that are achieved through the deployment of combinations of data-oriented:

- `Production Scheduling Services`, that interact with the schedulers of each actor involved in the production process and support the actors in scheduling the production in an efficient way;
- `Energy Efficiency Services`, to monitor the efficiency of energy harvesting assets in the supply chain;
- `Product Quality Services`, in charge of correlating anomalies detected on the product along its lifecycle with anomalies raised on work centers involved in the production process;

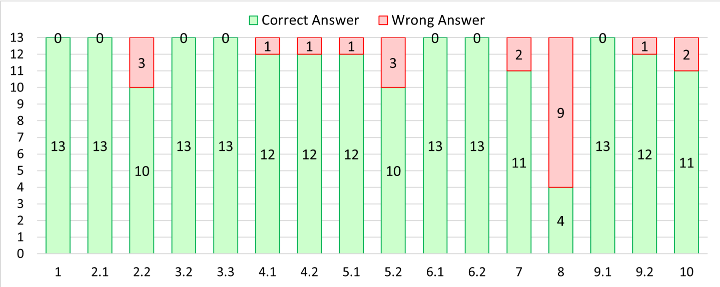

Figure 3: Comparison between correct and wrong answers in the usability experiment.

- `Remote Monitoring Service`, to monitor the work centers in order to detect anomalies that may lead to higher consumption or breakdown/damage of the monitored system or production process failure in the near future; to this aim, the IDEAaS approach, described in [4], has been integrated within the architecture; IDEAaS is a novel system used to detect anomalies within data streams in Industry 4.0 applications, by applying incremental clustering and data relevance evaluation techniques to manage volume and velocity of incoming data; the IDEAaS approach can benefit from the partition of incoming data streams according to different perspectives or contexts; for this reason, the application of IDEAaS techniques has been fostered through its integration on top of the multi-perspective information model of the CPM, as demonstrated in the next section on experimental results.

These services are strongly interleaved over a supply chain. For example, a peak in energy consumption does not always mean a low efficiency issue. Indeed, such peak may be due to a dense production scheduling. By correlating the output of the production scheduler with the energy efficiency optimisation module, these cases can be properly detected. For this reason, the mechanical supplier relies on the service, which is used for the anomaly detection. Similarly, problems raised during product quality tests may be correlated with anomalies raised on the machines of one of the participants in the production process. To implement all these services, acquisition and visualisation of data under the three perspectives of product, process and industrial assets are required for all the actors of the supply chain.

Finally, display services populate a web-based dashboard, that the actors use for data exploration. From the home page of the dashboard, it is possible to start the data exploration by following one of the three perspectives, namely, product, process and industrial assets. Each perspective brings to a UI component (tile) implemented using ReactJS libraries: i) the *product synoptic* tile allows an exploration from the product perspective; ii) the *process phases* tile allows an exploration from the process perspective; iii) the *working centers* tile allows exploration from the industrial asset perspective. The management of different access permissions through display services leads to tiles customisation. Therefore, it is possible through the dashboards to configure what to show and what to hide for each actor according to his/her role.

Preliminary validation of the approach. We performed a usability experiment with 13 participants grouped in three categories, with increasing expertise on the use of data visualisation tools, such as synoptic and dashboard, to explore data relate to industrial processes. The users have been requested to perform a set of exploration actions, that included multiple perspectives, (e.g., *What are the problems, if any, that occurred on industrial assets during the production of the valve VSS000799?*). The experiment is *significant*, since the experimental setup and the participants are representative of the considered industrial context, and *reliable* (in this particular context, poor reliability is linked to users' variability). The experiment comprised four steps: test preparation, its introduction to participants explaining the purpose of the test, its execution, in which the examiner observes without influencing the participant, and debriefing, to collect through questionnaires the participants' opinions about the perceived difficulty of the experiment. Figure 3 shows the correctness of the answers, having a high rate of 89%, thus demonstrating the effectiveness of the multi-perspective approach in supporting data exploration in industrial environments.

5. Related Work

The modelling approach addressed in this paper has been compared with the literature about the design of information models for managing the data flow from the physical world and the cyber one, where solutions such as Digital Twins and Digital Threads have been investigated. Authors in [5] describe a model-based approach (and a corresponding web-based GUI) to compose predefined building blocks, abstracted as Smart Services, connected each other and hierarchically organised. In [6] authors proposes an architecture where users specify a goal and take advantage of technologies such as digital twins to automatically compose the corresponding physical processes. Ontologies have been also proposed in [7] to face interoperability issues. In these papers the focus is on CPS within a single production line. Authors in [8] propose a generic architecture based on digital twin that relies on four layers (physical, network, virtual and application layer). In [9] authors model Digital Twins for product customisation. An information model is proposed to provide data about product, process, plan, plant, resource. The approach provides five services: production planning, automated execution, real-time monitoring, abnormal situation notification and dynamic response. In our information model, process and plan are associated to the different concepts of production phases and phase execution. Moreover, resources are properly organised within plants in a hierarchical way and services are designed on top of such multi-perspective model. The modern version of Digital Thread can be seen as a step forward with respect to Digital Twins. In [1] the authors focus on the use of models for designing smart products along their lifecycle, being agnostic about the specific technologies, and binding to specific implementations of such features only when needed. A case study for the design of Universal Robots UR3 controller is proposed in [10] within a model-driven integrated development environment independently from any implementation programming language, operating system, or runtime platform. The focus in the latter papers is on the smart product, whose representation evolves during the product lifecycle, but no data is collected on the design, production or maintenance stages. Similarly, the notion of Digital Thread proposed by commercial solutions such as PTC Windchill is implemented as a sequence of interleaved BOMs (e.g., EBOM, MBOM, as-built, as-maintained) without collecting data on the process and on the work centers during the product lifecycle. Our approach, although mainly focused on the production phase, could be fruitfully extended also during the other stages of the product lifecycle, such as the design or the maintenance ones.

6. Concluding Remarks

In this paper, an information model has been proposed that relates the digital counterpart of a product with data collected over the phases of the product lifecycle, according to three perspectives, namely the production process, the industrial assets and the product. The approach goes beyond the boundaries of a single actor, embracing all the participants in the production network, thus including also the definition of a portfolio of services on top of the model and data and services access rules for different actors involved in the supply chain. Future efforts are required to test the proposed solution in an intensive way when the massive collection of data will start in the project. Furthermore, a methodology is being developed to extend the portfolio of services to other kinds of manufacturing production networks and with proper access control policies to manage the authorisation of supply chain actors on data and services.

References

[1] T. Margaria, A. Schieweck, The Digital Thread in Industry 4.0, in: IFM 2019, 2019, pp. 3–24.

[2] P. Loucopoulos, E. Kavakli, N. Chechina, Requirements engineering for cyber physical production systems, in: CAiSE 2019, volume 11483, Springer, 2019, pp. 276–291.

[3] M. Hankel, B. Rexroth, The Reference Architectural Model Industrie 4.0 (RAMI 4.0), in: ZVEI, 2015.

[4] A. Bagozi, D. Bianchini, V. D. Antonellis, M. Garda, A. Marini, A relevance-based approach for big data exploration, FGCS 101 (2019) 51 – 69.

[5] D. Stock, D. Schel, T. Bauernhansl, Middleware-based Cyber-Physical Production System Modeling for Operators, Procedia Manufacturing 42 (2020) 111–118.

[6] T. Catarci, D. Firmani, F. Leotta, F. Mandreoli, M. Mecella, F. Sapio, A conceptual architecture and model for smart manufacturing relying on service-based digital twins, in: E. Bertino, C. K. Chang, P. Chen, E. Damiani, M. Goul, K. Oyama (Eds.), 2019 IEEE International Conference on Web Services, ICWS 2019, Milan, Italy, July 8-13, 2019, IEEE, 2019, pp. 229–236.

[7] D. Stock, D. Schel, Cyber-Physical Production System Finger-printing, Procedia CIRP 81 (2019) 393–398.

[8] H. Zhang, Q. Yan, Z. Wen, Information modeling for cyber-physical production system based on digital twin and automationml, The international journal of advanced manufacturing technology 107 (2020) 1927–1945.

[9] K. Park, J. Lee, H. Kim, S. Noh, Digital twin-based cyber physical production system architectural framework for personalized production, International Journal of Advanced Manufacturing Technology 106 (2020) 1787–1810.

[10] S. Bosselmann, M. Frohme, D. Kopetzki, M. Lybecait, S. Naujokat, J. Meubauer, D. Wirkner, P. Zweihoff, B. Steffen, DIME: A Programming-Less Modeling Environment for Web Applications, in: ISoLA 2021, 2016, pp. 809–832.