



**HAL**  
open science

## Data-driven pattern-based constructs definition for the digital transformation modelling of collaborative networked manufacturing enterprises

Concetta Semeraro, Mario Lezoche, Hervé Panetto, Michele Dassisti, Stefano Cafagna

### ► To cite this version:

Concetta Semeraro, Mario Lezoche, Hervé Panetto, Michele Dassisti, Stefano Cafagna. Data-driven pattern-based constructs definition for the digital transformation modelling of collaborative networked manufacturing enterprises. 20th Working Conference on Virtual Enterprises (PRO-VE), Sep 2019, Turin, Italy. pp.507-515, 10.1007/978-3-030-28464-0\_44 . hal-02191335

**HAL Id: hal-02191335**

**<https://hal.science/hal-02191335>**

Submitted on 23 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Data-driven pattern-based constructs definition for the digital transformation modelling of collaborative networked manufacturing enterprises

Concetta Semeraro<sup>1,2,3</sup>, Mario Lezoche<sup>2</sup>, Hervé Panetto<sup>2</sup>, Michele Dassisti<sup>1</sup>,  
Stefano Cafagna<sup>3</sup>

<sup>1</sup> Department of Mechanics, Management & Mathematics (DMMM), Polytechnical University of Bari, Bari, Italy

<sup>2</sup> Université de Lorraine, CNRS, CRAN, Nancy, France

<sup>3</sup> Master Italy s.r.l, Conversano, Bari, Italy

**Abstract.** The digital transformation of collaborative networked manufacturing enterprises requires building and applying digital models representing the set of resources and processes knowledge. Modelling such digital copy of the physical system to perform real-time validation and optimization is quite complex and thus needs a big amount of data and some modelling patterns representing the operational semantics of the modelled elements. Generally, the modelling action has a specific application type. For this reason, the core challenge of the digital transformation modelling is to create a modular “digital model”, namely a decomposable and re-composable model, towards different applications. The authors propose an approach based on the combination of data-driven and model-driven approaches, to identify and formalize modelling patterns, that combine for developing a modular executable model of the studied system.

**Keywords:** Data-Driven Approaches, Model-Based Approaches, Modularity, Knowledge sharing.

## 1 Introduction

In the age of the so-called “Factory of the Future” also named “Industry of the Future” or Industry 4.0 to name only a few of those initiatives, the transformation from an Automated Factory to an Autonomous Factory [1] needs to overcome the information asymmetry amongst technology, processes, people and organizations [2] along the entire system lifecycle. This action requires creating a so-called collaborative network enterprise [3]. The interactions between entities should be supported by computer networks. The collaborative network, in the design phase, guides the designers to interact with customers. Iteratively it lets them adjusting the network according to their expectations, improving the design models and achieving personalized product design [4]. In the production phase, it enables the simulation of the plant in a virtual space, identifying and optimizing the actual production and predicting failures. In the service phase, the collaborative network can provide value-added services with the support of physical simulation and data driven intelligence [5]. In the context of manufacturing processes, the prerequisite of a manufacturing network collaboration is to share skills and core competencies[6]. A “digitalization” approach enables the virtual replication of the factory to monitor and simulate real time physical processes. It allows to connect the entire value chain [7] by merging sensor data acquired from the physical world into

virtual or simulation-based models. It is then necessary to emulate the system behaviour through realistic models. However, with the advent of cyber-physical systems (CPS) [8], it is hard to construct an accurate model by using traditional model-based approaches because of the complexity of the systems. On the other hand, recent advances in sensor technology [9] have enabled significant growth of data collection and analysis, leading researchers to focus on data-driven methods.

The paper proposes the combination of a data-driven approach with a model-based approach [10] [11] to build a modular approach in order to help designers to generate models from collected data, based on a set of patterns representing standard functions of the real manufacturing processes. The core elements of a digital model are implemented by the fusion of sensor-based data, physical-based patterns and data-driven models. They enable the creation of a reliable decision making system [12] for a collaborative network. At the same time, a modular approach needs to be developed for improving the efficiency of those digital models.

## 2 State of Art

The increasing complexity of manufacturing systems requires new approaches to detect production failures and diagnose operating profiles [5]. Model-based and data-driven approaches cope with such issue [10].

The model-based approaches compare simulated results with known information, represented by mathematical or physical equations. The model approach is based on a set of different models to represent the structure, the behaviour and the interactions of a physical system to be monitored or predicted [13] [14]. There are various models focusing on the representation of different characteristics of the reality here a not exhaustive list of the most studied and used:

- A geometric model defines shapes, sizes, positions and it assembles the relations of machine components [15]. It reflects the geometry, the kinematics, the logic and the interfaces of the real system [16].
- A physical model analyses the phenomena, such as deformation, cracking and corrosion. It simulates the physical properties (e.g. function/capacity, cutting force, torque and wear) and loads (e.g. stress, resistance and temperature) [17].
- A behaviour model describes the way the physical system is governed by driving factors (e.g. control orders) or disturbing factors (e.g. human interferences).
- A collaborative information model [18] defines how different components interact and simulates the collaborative behaviour among several assets.
- A decision-making model [18] makes the model capable of evaluating, reasoning, and validating. It consists of variable input, algorithms and a collection of constraints and rules. It includes rules of constraints, associations and deductions [13] and it stores and analyses the running status data, then it makes decisions using machine learning algorithm.

The geometric, physical, behaviour and collaborative models are descriptive models [19], while the decision-making model is an intelligent model [19] and it could be related to data-driven approach. The data-driven approaches refer to models derived from processed data (e.g. sensor/actuator data) which represent the current state of the system [20]. Data driven models are designed to mine the hidden patterns and knowledge through the analysis of a huge amount of historical data [21]. The patterns

mining methods make use of the context data and they unveil the complex coupling relationships. The data driven approaches can be classified in supervised, unsupervised and reinforcement learning approaches [22]. The supervised learning develops model based on input and output data [10]. The algorithms of this class consist of a target/outcome variable (or dependent variable) which is to be predicted by a given set of predictors (independent variables). This approach enables to classify and determine a list of system's defaults [23] with health indicators for each part of it [24]. The unsupervised learning [25], instead, discovers an internal representation from input data only. It has not any target or outcome variable to predict and estimate and are used to create autonomously clusters for different working regimes and machine conditions [26]. Finally, the reinforcement learning trains the machine to make specific decisions [27]. The machine is exposed to an environment where it trains itself continually using trial and error. Compared to supervised and unsupervised learning methods, the machine learns from past and tries to capture the best possible knowledge to make accurate business decisions [28].

On the one hand, the data-driven approaches allow to integrate parameters across different domains of a collaborative network (e.g. product, process and logistics) into models that would be difficult to build with the traditional model-based modelling approaches [29]. The data-driven, in fact, aim at transforming the data into relevant information but the quality and scope of the data play a critical role [30].

On the other hand, the model-based approaches rely on the use of mathematical models to simulate the systems behaviour in different operating conditions but for complex systems, these models are not easy to develop and keep updated during the system life-cycle [30]. At the same time, a single approach cannot be adapted to all different applications of a collaborative network because of the complexity and the variety that characterize manufacturing systems. Hybrid approaches [10] are been developed to cope with defined problem such as fault detection and diagnosis [31], prediction or classification accuracy [32] only for specific application case.

It means that is necessary to develop and standardize modelling patterns for developing a modular approach. Modularity, in fact, is concerned with shifting from rigid systems and inflexible production models toward an agile system. The modularity can be defined as the capability of system components to be separated and combined easily and quickly [33]. For the digital transformation modelling, the modularity is the ability to integrate, to add, and to replace models [34] based on the specific application. The idea behind a modular approach is to use, and especially re-use, predefined functional patterns, that are systematically developed and linked for the configuration of a holistic manufacturing system [35]. The paper presents a modular approach for discovering automatically data-driven pattern-based constructs in order to generate semi-automated models.

### **3 Data-driven pattern-based constructs definition**

The modular design approach is developed to capture patterns from data and to share and to reuse knowledge encapsulated in a pattern among systems or processes operating

in a similar condition. It enables the virtual replication of an enterprise and then monitoring, simulating and predicting failures in manufacturing processes. The idea is to build knowledge-based modelling constructs representing data-driven patterns contextualised in different process' situations. The critical point is discovering those data-driven patterns. FCA (Formal Concept Analysis) is a mathematical theory oriented at applications in knowledge representation [36]. It provides tools to group the data and to discover formal patterns by representing it as a hierarchy of formal concepts organised in a semi ordered set named lattice. Given a set of objects, a set of properties, and defined the relations between object and properties, a formal concept represents a subset of objects sharing the same sub-set of properties. A concept is constituted by two parts: its extension which consists of all objects belonging to the concept, and its intension which comprises all properties shared by those objects. This understanding allows a formal discovering of associations among concepts and consequently recognizing which concepts are closely related based on the set of shared properties. In this context, FCA is applied to discover automatically patterns from data. The patterns generated are selected and evaluated for detecting relationships, trends, associations, and anomalies that characterize a system in analysis. A pattern helps to discover useful knowledge from a collection of data. It can describe recurrent behaviours of the system or it can codify tacit associations that can be used to predict the future behaviour. The patterns discovered are modelled in System Modelling Language (SysML) [37]. SysML provides nine interrelated types of diagrams to describe the function, the structure, the behaviour and the system requirements. It supports the specification, analysis, and verification of systems' models.

The modular approach developed is articulated in four different stages. The first stage is to define the knowledge structure in a data table (Fig.1). The data table presents the objects on the rows and the properties on the columns. The cross indicates that exists a relation between an object and an attribute. In second stage, FCA converts automatically the data table into a lattice (Fig.2). The lattice presents nodes, connecting line and names. Nodes represent formal concepts and the lines connect objects and attributes belonging to a concept. The name of each objects is noted under nodes instead the name of each property above nodes. The third stage is to evaluate the patterns extracted from the lattice (Fig.3) and to detect the behaviours, the associations to model. It is possible also to identify the relationships among patterns. The last stage is to generate the model of the pattern in SysML diagrams (Fig.4), based on the features defined in the previous step. The objects are modelled as blocks and the properties as value properties using the block definition diagram (bdd). The behaviour of a pattern is modelled as a set of actions to describe how the inputs are transformed into outputs. In particular, the state diagram (stm) models the behaviour defining states and events of an object during its lifetime. It simulates how the states change based on internal or external events. The parametric diagram (par) models the associations discovered in terms of constraints. Constraints represent physical laws or mathematical and logical operators or decisions that evaluate input parameters to return a result. The modular approach enables to create semi-automated models based on the patterns discovered.

#### 4 Case Study

An Italian SME, the Master Italy s.r.l, that produces small accessories for civil window frames, is here considered to test how to construct data-driven pattern. The process and the product in analysis are respectively the die casting aluminium and the steel corner. Die casting is a metal casting process that is characterized by forcing molten metal under high pressure into a mould cavity. The injection cycle of die casting aluminium process is composed by four different phases:

1. Melting: the aluminium enters at the solid state and exits at the molten state.
2. Injection: the molten aluminium is transferred, through a plunger, into a chamber where it is injected into the mould.
3. Moulding: the molten aluminium solidifies in the mould cavity.
4. Extraction: an ejection mechanism pushes the casting out of the mould cavity.

The data set presents the injection cycles as objects and technological parameters as properties (Fig.1). The technological parameters monitored are: plunger course 'C1'(m), 'C2'(m), and time 'T1'(sec) 'T2'(sec), in first and second stage of injection phase, multiplied course 'CC' (m), multiplied pressure 'PM' (bar), cavity pressure 'PS' (bar), clamping force 'FC' (N). The FCA is applied to determine the hierarchical lattice (Fig.2). There are five different range of values for each parameter. The first pattern extracted (Fig.3a) has a set of objects sharing the properties C2-1, T1-2, C2-5, T2-1, PM-1, FC-3. The second pattern (Fig.3b) instead has the properties C2-1, T1-2, C2-5, T2-1, PM-1, FC-4, PS-2. The objects of the patterns share most of properties (C2-1, T1-2, C2-5, T2-1, PM-1) but FC operates in the range 4 and it depends also on PS-2 in the pattern b). These are two different but correlated patterns. The parameters C1,C2,T1,T2,PM are parameters related to the injection stage, while FC and PS are related to the moulding stage. It means that there are physical equations between the technological parameters of the two phases. Clamping force (FC), in fact, refers to the force applied to a mould by the clamping unit of the injection moulding machine. In order to keep the mold closed, this force must oppose the separating force, caused by the injection parameters (C1,C2,T1,T2,PM). The required clamping force depends on the cavity pressure (PS) inside the mould and the projected area, on which the pressure acts. For this reason, the patterns are respectively the clamping pattern (a) and the pressure control pattern (b). The model of the compression pattern (Fig.4) presents the process stage as blocks and the technological parameters as value properties on the block definition diagram. The physical laws are modelled on the parametric diagram. The state diagram can simulate if all injection cycles (objects) of the clamping pattern are conformed to the quality product. In the clamping pattern, all parameters (properties) operate into the defined ranges. In the pressure control pattern, the state of the property FC-4 closes out of range. It means that PS can be analysed to detect and to implement a strategy capable of preventing quality defects caused by FC.

#### 4 Conclusions and Future Works

The paper presents how to construct data-driven patterns in order to create a modular approach for digital modelling transformation. The approach defines how to build a semi-automated model based on the patterns discovered automatically in FCA.

Fig.1 Data Table

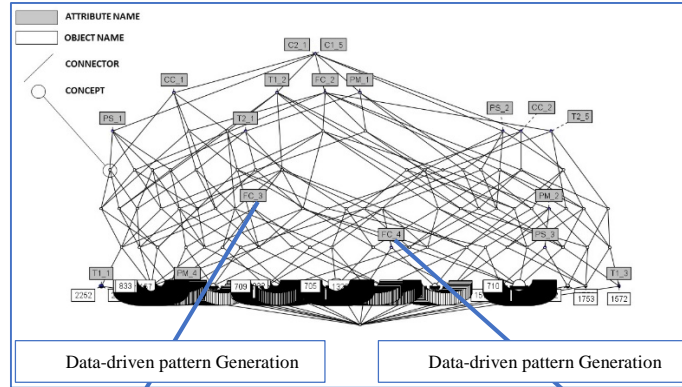


Fig.2. FCA Lattice

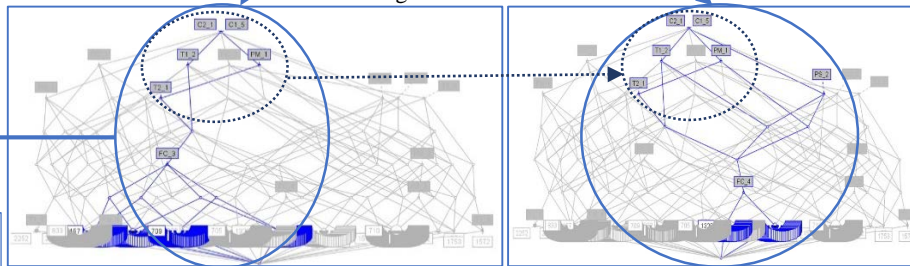


Fig.3a) Clamping Pattern

Fig.3b) Pression Control Pattern

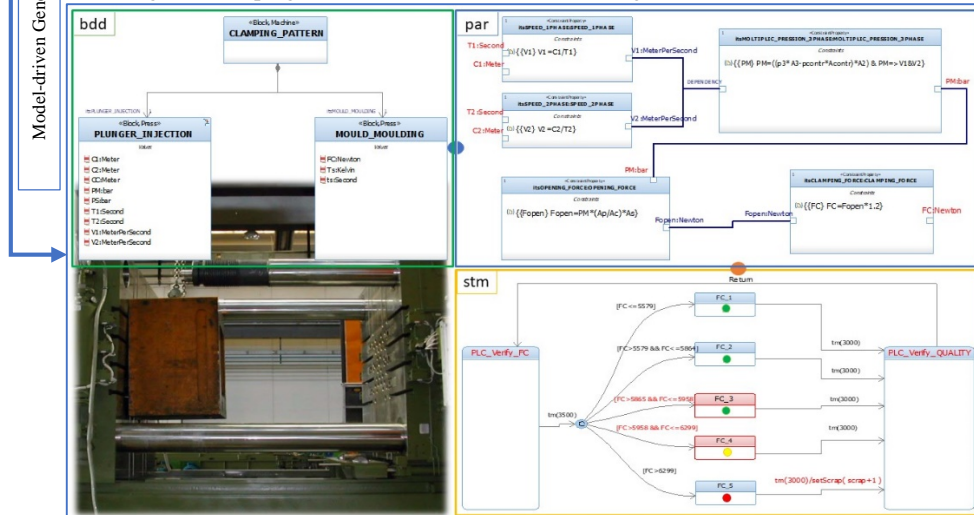


Fig.4. SysML Structure and Behaviour Model of the Clamping Pattern

The patterns are systematically modelled in SysML and linked for the configuration of a holistic manufacturing system. The approach shows also how it is possible to discover the behaviours and associations from data and how to analyse the relationships between patterns. The goal is to use the same patterns for modelling other systems. The future works is to enrich the pattern' semantics to create a comprehensive library of formalized data-driven patterns. In this way, data-driven patterns can be combined, based on the specific application, to create easily dynamic models. The approach tested is inductive. It means that the data suggests model about the meaning of their content. The future work is to develop a deductive approach. In this case, the modular design approach can be used also to verify if a certain model (or hypothesis) is consistent with the available data or if it necessary to implement a monitoring strategy to collect new data.

Continued monitoring, data collection and analysis provide up-to-date information about the behaviours of the system in a continuous stream. Those actions enable to collect and convert data in information, share the information acquired, formalize the knowledge, joint performance measurements and leverage the skills and the knowledge. These benefits are in line with the collaboration-associated benefits of a collaborative network.

## References

- [1] R. Rosen, G. von Wichert, G. Lo, e K. D. Bettenhausen, «About the importance of autonomy and digital twins for the future of manufacturing», *IFAC-Pap.*, vol. 48, n. 3, pagg. 567–572, 2015.
- [2] A. Padovano, F. Longo, L. Nicoletti, e G. Mirabelli, «A Digital Twin based Service Oriented Application for a 4.0 Knowledge Navigation in the Smart Factory», *IFAC-Pap.*, vol. 51, n. 11, pagg. 631–636, 2018.
- [3] L. M. Camarinha-Matos e H. Afsarmanesh, «Collaborative networks: Value creation in a knowledge society,(invited keynote paper)», 2006.
- [4] Q. Qi e F. Tao, «Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison», *IEEE Access*, vol. 6, pagg. 3585–3593, 2018.
- [5] J. Wang, L. Ye, R. X. Gao, C. Li, e L. Zhang, «Digital Twin for rotating machinery fault diagnosis in smart manufacturing», *Int. J. Prod. Res.*, vol. 0, n. 0, pagg. 1–15, dic. 2018.
- [6] S. Min *et al.*, «Supply chain collaboration: what's happening?», *Int. J. Logist. Manag.*, vol. 16, n. 2, pagg. 237–256, 2005.
- [7] M. Ghobakhloo, «The future of manufacturing industry: A strategic roadmap toward Industry 4.0», *J. Manuf. Technol. Manag.*, vol. 29, n. 6, pagg. 910–936, 2018.
- [8] J. Lee, B. Bagheri, e H.-A. Kao, «A cyber-physical systems architecture for industry 4.0-based manufacturing systems», *Manuf. Lett.*, vol. 3, pagg. 18–23, 2015.
- [9] M. Dassisti *et al.*, «Industry 4.0 paradigm: The viewpoint of the small and medium enterprises», in *7th International Conference on Information Society and Technology, ICIST 2017*, 2017, vol. 1, pagg. 50–54.
- [10] K. Tidriri, N. Chatti, S. Verron, e T. Tiplica, «Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges», *Annu. Rev. Control*, vol. 42, pagg. 63–81, 2016.
- [11] H. Khorasgani, A. Farahat, K. Ristovski, C. Gupta, e G. Biswas, «A Framework for Unifying Model-based and Data-driven Fault Diagnosis», in *PHM Society Conference*, 2018, vol. 10.
- [12] Z. Liu, N. Meyendorf, e N. Mrad, «The role of data fusion in predictive maintenance using digital twin», presentato al AIP Conference Proceedings, 2018, vol. 1949.
- [13] F. Tao, M. Zhang, Y. Liu, e A. Y. C. Nee, «Digital twin driven prognostics and health management for complex equipment», *CIRP Ann.*, 2018.



- [14] C. Zhuang, J. Liu, e H. Xiong, «Digital twin-based smart production management and control framework for the complex product assembly shop-floor», *Int. J. Adv. Manuf. Technol.*, vol. 96, n. 1–4, pagg. 1149–1163, 2018.
- [15] M. Ayani, M. Ganebäck, e A. H. C. Ng, «Digital Twin: Applying emulation for machine reconditioning», *Procedia CIRP*, vol. 72, pagg. 243–248, gen. 2018.
- [16] L. Damiani, M. Demartini, P. Giribone, M. Maggiani, R. Revetria, e F. Tonelli, «Simulation and Digital Twin Based Design of a Production Line: A Case Study.», in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2018, vol. 2.
- [17] J. Post, M. Groen, e G. Klaseboer, «PHYSICAL MODEL BASED DIGITAL TWINS IN MANUFACTURING PROCESSES», *Opt. Lett.*, vol. 34, n. 13, pagg. 1982–4, 2009.
- [18] J. Bao, D. Guo, J. Li, e J. Zhang, «The modelling and operations for the digital twin in the context of manufacturing», *Enterp. Inf. Syst.*, vol. 0, n. 0, pagg. 1–23, ott. 2018.
- [19] W. Luo, T. Hu, W. Zhu, e F. Tao, «Digital twin modeling method for CNC machine tool», in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, 2018, pagg. 1–4.
- [20] B. A. Talkhestani, N. Jazdi, W. Schlögl, e M. Weyrich, «A concept in synchronization of virtual production system with real factory based on anchor-point method», presentato al *Procedia CIRP*, 2018, vol. 67, pagg. 13–17.
- [21] Y. Zhang, S. Ren, Y. Liu, T. Sakao, e D. Huisingh, «A framework for Big Data driven product lifecycle management», *J. Clean. Prod.*, vol. 159, pagg. 229–240, 2017.
- [22] S. Jain, G. Shao, e S.-J. Shin, «Manufacturing data analytics using a virtual factory representation», *Int. J. Prod. Res.*, vol. 55, n. 18, pagg. 5450–5464, 2017.
- [23] R. M. Asimov, S. V. Chernoshey, I. Kruse, e V. S. Osipovich, «Digital twin in the Analysis of a Big Data», 2018.
- [24] K. Ding, F. T. S. Chan, X. Zhang, G. Zhou, e F. Zhang, «Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors», *Int. J. Prod. Res.*, vol. 0, n. 0, pagg. 1–20, gen. 2019.
- [25] T. Sutharssan, S. Stoyanov, C. Bailey, e C. Yin, «Prognostic and health management for engineering systems: a review of the data-driven approach and algorithms», *J. Eng.*, vol. 2015, n. 7, pagg. 215–222, 2015.
- [26] J. Lee, H.-A. Kao, e S. Yang, «Service innovation and smart analytics for industry 4.0 and big data environment», *Procedia Cirp*, vol. 16, pagg. 3–8, 2014.
- [27] L. Zhou, S. Pan, J. Wang, e A. V. Vasilakos, «Machine learning on big data: Opportunities and challenges», *Neurocomputing*, vol. 237, pagg. 350–361, 2017.
- [28] M. van Otterlo e M. Wiering, «Reinforcement learning and markov decision processes», in *Reinforcement Learning*, Springer, 2012, pagg. 3–42.
- [29] A. Kusiak, «Smart manufacturing», *Int. J. Prod. Res.*, vol. 56, n. 1–2, pagg. 508–517, 2018.
- [30] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, e G. Tripot, «A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models», *IEEE Trans. Reliab.*, vol. 61, n. 2, pagg. 491–503, 2012.
- [31] X. Zhang e K. A. Hoo, «Effective fault detection and isolation using bond graph-based domain decomposition», *Comput. Chem. Eng.*, vol. 35, n. 1, pagg. 132–148, 2011.
- [32] K. Ghosh, Y. S. Ng, e R. Srinivasan, «Evaluation of decision fusion strategies for effective collaboration among heterogeneous fault diagnostic methods», *Comput. Chem. Eng.*, vol. 35, n. 2, pagg. 342–355, 2011.
- [33] M. Mabkhot, A. Al-Ahmari, B. Salah, e H. Alkhalefah, «Requirements of the Smart Factory System: A Survey and Perspective», *Machines*, vol. 6, n. 2, pag. 23, 2018.
- [34] B. Schleich, N. ANWER, L. Mathieu, e S. Wartack, «Shaping the digital twin for design and production engineering», *CIRP Ann. - Manuf. Technol.*, vol. 66, n. 1, 2017.
- [35] J. Guo, N. Zhao, L. Sun, e S. Zhang, «Modular based flexible digital twin for factory design», *J. Ambient Intell. Humaniz. Comput.*, pagg. 1–12, 2018.
- [36] B. Ganter, G. Stumme, e R. Wille, *Formal concept analysis: foundations and applications*, vol. 3626. springer, 2005.
- [37] S. Friedenthal, A. Moore, e R. Steiner, *A practical guide to SysML: the systems modeling language*. Morgan Kaufmann, 2014.