Formalisation models and knowledge extraction: Application to heterogeneous data sources in the context of the Industry of the Future

Mario Lezoche

Collegium Sciences et Technologie
Pôle AM2I
École Doctorale IAEM Lorraine
Comission de Mention Automatique

Modèles de formalisation et extraction de connaissance : Application aux sources de données hétérogènes dans le contexte de l 'Industrie du futur

HABILITATION A DIRIGER DES RECHERCHES

présentée et soutenue publiquement le 12 Janvier 2021

pour l'obtention d'une

Habilitation de l'Université de Lorraine

(mention Automatique, Traitement du Signal et des Images, Génie Informatique)

par

Mario LEZOCHE

Maître de Conférences
Docteur de l'Université ROMA TRE

Composition du jury

Président :

Hind BRIL-EL HAOUZI	Professeur des Universités, CRAN, Université de Lorraine

Rapporteurs :

Marianne HUCHARD	Professeur des Universités, LIRM, Université de Montpellier II
Néjib MOALLA	Professeur des Universités, DISP, Université Lumière Lyon II
Angel ORTIZ BAS	Professeur des Universités, ETSII, Universitat Politecnica de Valencia, Espagne

Examinateurs :

Michele DASSISTI	Professeur des Universités, DMMM, Politecnico di Bari, Italie
Alejandro FERNANDEZ	Professeur des Universités, LIFIA, Universidad Nacional de La Plata, Argentina
Virginie GOEPP	Maître de Conférence HDR, ICUBE, INSA de Strasbourg
Hervé PANETTO	Professeur des Universités, CRAN, Université de Lorraine

Invités :

William DERIGENT	Maître de Conférence HDR, CRAN, Université de Lorraine

Centre de Recherche en Automatique de Nancy — UMR 7039
Université de Lorraine — CNRS

"Any fool can know. The point is to understand."
— Albert Einstein

« Carpe diem »

À Lisa, la Merveille que la Vie m'a donné la chance de rencontrer

Acknowledgements

Acknowledgements in an HDR manuscript are an essential part. Although I have often been advised to use the personal pronoun "I" when reporting the results of the research I have worked on, nothing would have been possible without the intervention of the many colleagues and collaborators with whom I have had the pleasure and honour to work.

The path I have travelled since reaching my first academic goal, the PhD, has been studded with stimulating, pleasant, enriching encounters that have made me the university professor I am today and that are still working to allow me to change and improve in order to reach new evolutionary goals in the future.

The first person I would like to thank is without doubt Hervé Panetto, my scientific godfather! Without him, I would not be here to write my manuscript for the qualification to lead research. He was the one who welcomed me to CRAN, the one with whom I learned to direct the PhDs of my students, the one with whom I learned to weave the strong professional relationships that have added so much to my academic views. He enabled me to fully understand a domain that was unknown to me: Interoperability.
Hervé and I share many Italian, Serbian, Spanish, English, Brazilian and Argentinean friends, and his great communication skills have taught me a lot.
At CRAN I was welcomed by the many colleagues from the ISET department with whom I have shared my life as a researcher since 2010! Alexis, Pascale, William, Alex, David, Phuc, Benoit, Hind, Eric, JP, Thierry, Patrick, Vincent, Didier.
A separate chapter of my researcher story has been the great teamwork with the PhD students I have had the pleasure of framing: Yongxin, Silvana, Yasamin, Concetta and Mickael! They have all been fantastic research partners who have helped me to simplify scientific communication and to go deeper and deeper into the research domains I care about.

The HDR process is lengthy and has involved many researchers who have invested a lot of time in evaluating the results of my teaching and research activities and eventually the future project. I would like to thank Hind Bril El Haouzi and Néjib Moalla for agreeing to evaluate my manuscript in its initial form, which resulted in the University of Lorraine's authorisation to conduct research. Special thanks to Marianne Huchard, Néjib Moalla and Angel Ortiz Bas for being the main rapporteurs on my manuscript and for their reports which were a great source of improvement. Thanks are due to Hervé Panetto, Michele Dassisti, Virginie Goepp, Alejandro Fernandez and William Derigent who completed the jury.
The fruitful exchanges we had while defending HDR is one of the reasons why I chose this profession!

I have the enormous good fortune to do a job that I have chosen and that I am passionate about, but all this would be meaningless if there was no Life outside the professional sphere.
Many people have been at my side and have accompanied me to make me who I am today. This HDR also stems from their presence.
My parents, my mother who has always pushed me towards research and my father who is always proud of the path I have taken. My brother Diego who, despite being geographically distant, has always been at my side in important moments.
The family of my beloved wife, who has long since become my family. My birth family who have always cheered me on and have always supported me even when my choices have taken me far away from them.
My friends with whom I share the other passion of my life, martial arts, and those with whom I share the rest of my life: Guf, Raffi and the whole wonderful gang of troublemakers with whom it is nice to share the passing of time.

My soulmate, Lisa, who has put up with my very long days, my white nights and all the work I have brought home over many weekends. Her encouragement has always been an inexhaustible source of inspiration for me. She is my companion on the journey, that magical path that I have been 'lucky' to share with her and that takes us far away together. Thank you for helping me to remember what the priorities of our existence are.

I dedicate this qualification to my partner in Life, the one who

helps me along this wonderful path that is Life:

***Lisa**, the incessant Wonder.*

List of Images

List of Tables

Foreword

The document you are about to read describes my activities as a teacher and researcher since September 2012. On that date I integrated as associate professor (Maître de Conférences) the CRAN (Research Centre for Automatic Control, CNRS UMR 7039) and the IUT (University Institute of Technology) Hubert Curien d'Épinal in the QLIO (Quality, Industrial Logistics and Organization) department which is a component of the Université de Lorraine.

The professional life of the researcher is studded with the creation of different reports, scientific articles, balance sheets, administrative articles, documents related to projects, this document that you are about to read could be counted as another document that must be generated to continue on his professional path. In reality this report is something unique, not because it is the document necessary for my qualification to lead research (habilitation à diriger la recherche), but because it has helped me to better understand who I am as a researcher and where I want to go.
Writing, reflecting and asking myself many questions, I found myself reflecting on the present of my research and pedagogy, it made me rethink the path I took to get to this point and finally it basically helped me to put in order the multiple ideas to create my research project for the years to come.

In the last 8 years the taste for study and research has been getting more and more refined and thanks to the immense fortune of being able to do a job that I love and that allows me to be free in my choices I had the opportunity, always growing, to move towards themes that attracted me intellectually more and more. The formalisation of knowledge, its extraction and reuse in a variety of contexts completely attracts the interest of my intellect. From my point of view this theme is central to the very definition of humanity and as such I am completely involved in it.

This qualification document has above all a vocation to demonstrate the ability to lead a research team. Through its general introduction and the following three chapters it will show on overview of my research domain interests (introduction), my research management skills (chapter 1), my technical-scientific skills (chapter 2) and will present my research and future teaching project (chapter 3).

General Introduction

The speed and measure of the changes coming about by the fourth industrial revolution are not to be ignored. These changes will bring about shifts in power, shifts in wealth, and knowledge. Only in being knowledgeable about these changes and the speed in which this is occurring can we ensure that advances in knowledge and technology reach all and benefit all [Min, 2018].

The first industrial revolution marked a period of development in the latter half of the 18th century, precisely in 1760, that transformed largely rural, agrarian societies in Europe and America into industrialized, urban ones. Goods that had once been painstakingly crafted by hand started to be produced in mass quantities by machines in factories, thanks to the introduction of new machines and techniques in textiles, iron making and other industries. This transition included the use of coal as the main energy while trains were the main means of transportation. Textile and steel were the dominant industries in terms of employment, value of output, and capital invested. Fuelled by the game-changing use of steam power, the first industrial revolution began in Britain and spread to the rest of the Europe and the entire world, including the United States, by the 1830s and ‘40s.

The second industrial revolution began in 1900 with the invention of the internal combustion engine. This led to an era of rapid industrialization using oil and electricity to power mass production. Technology has changed the world in many ways, but perhaps no period introduced more changes than the second industrial revolution. From the late 19th to early 20th centuries, cities grew, factories sprawled, and people’s lives became regulated by the clock rather than the sun. Rapid advances in the creation of steel, chemicals and electricity helped fuel production, including mass-produced consumer goods and weapons. It became far easier to get around on trains, automobiles and bicycles. At the same time, ideas and news spread via newspapers, the radio and telegraph.

Another century passes and we bear witness to the third industrial revolution. It started in 1960 and was characterized by the emergence of yet another source of untapped, at the time, energy. Nuclear energy. The third revolution brought forth the rise of electronics, telecommunications and of course computers. Through the new technologies, the third industrial revolution opened the doors to space expeditions, research, and biotechnology. In the world of the industries, two major inventions, Programmable Logic Controllers (PLCs) and Robots helped give rise to an era of high-level automation.

For many people, Industry 4.0 is the fourth industrial revolution [Prisecaru, 2016], although there is a large portion of people that still disagree. If we were to view Industry 4.0 as a revolution, then we would have to admit that it is a revolution happening right now. We are experiencing it every day and its magnitude is yet unknown. Industry 4.0

started in the dawn of the third millennium with the one thing that everyone uses every single day. The Internet. We can see the transition from the first industrial revolution that rooted for technological phenomenon all the way to Industry 4.0 that develops virtual reality worlds, allowing us to bend the laws of physics. The fourth industrial revolution shapes the world. Worldwide economies are based on it.

All industrial revolutions have substantially changed the way humans live. Each of them has increased the specific value that knowledge has in allowing the management of processes (marketing, finance, logistics, production, human resources, etc.). The more revolutions followed one another, the more companies have equipped themselves with systems to manage data, information and knowledge, which has become the key resource in a world where systems are increasingly complex, and information is increasingly numerous and comes from heterogeneous sources.

Industry 4.0 and Knowledge

In the fourth industrial revolution, the main goal of the implementation of new technologies is related to the effective and efficient customer-oriented adaptation of products (and thus production) and services in order to increase the value added for companies, raising their competitive position, while for customers improving satisfaction and loyalty [Roblek, 2016]. In order to achieve this goal, manufacturing companies need to develop and manage new knowledge that is crucial for the organization's decision-making process and the achieving of the related business goals [Abubakar, 2019].

Industry 4.0 reflects a combination of digital and manufacturing technologies, Specifically the new technological transformation embraces technological advances that concern the production process (i.e., advanced manufacturing systems, autonomous robots, additive manufacturing), the use of smart products and/or data tools and analytics [Porter, 2015]. Within the manufacturing process, the adoption of autonomous and/or collaborative robotics [Adamson, 2017] or 3D printing is opening up new opportunities to create new knowledge concerning products and processes [Anderson, 2012].

At the same time, smart products and "data-driven technologies" enable the successful acquisition of useful data from several sources within the organizational boundaries as well as from customers and suppliers [Klingenberg, 2019]. Therefore, Industry 4.0 stresses the huge potentialities of data that can be used in real time, enriching contextual knowledge or generating new one in the way products can be produced and used, as well as in the practices concerning value generation (from product to service), allowing firms to take actions and make decisions based on such knowledge [Tao, 2018]. Moreover, it is essential to consider a holistic view of manufacturing processes, integrating data from different sources to achieve the business benefits of new technologies [Schneider, 2018].

Therefore, products and services are highly influenced by this new industrial paradigm. Products become more complex, modular, and configurable supporting mass customization to meet specific customer needs.

Cyber Physical System as information sources

The technological advances have led to some examples of a new systems generation. Cyber-Physical Systems (CPSs) represent more than networking and information technology, information and knowledge being integrated into physical objects. By integrating perception, communication, learning, behaviour generation, reasoning into such systems a new generation of intelligent and autonomous systems may be developed. Industry 4.0 technologies related to cyber-physical systems, IoT, and cloud computing can generate benefits from a circular economy point of view since they allow design for circularity based on the information gathered from customers as well as through the whole production process [De Sousa, 2018].

A component is a core concept in the Industry 4.0 context. An Industry 4.0 component constitutes a specific case of a Cyber-Physical System (CPS). It is used as a model to represent the properties of a CPS, for instance, real objects in a production environment connected with virtual objects and processes. An Industry 4.0 component can be a production system, an individual machine, or an assembly inside a machine.

Some core concepts in CPS can be traced back to the sensor network research and technologies related to sensor nodes and sensor networks. A sensor node integrates sensors, actuators, computing elements (e.g. processor, memory, etc.), communication modules, and a battery. The sensor network interconnects many small sensor nodes via wireless or wired connection [Golatowski, 2003] as can be seen in the Figure 1. Called as Wireless Sensor Networks (WSN), a large number of sensor nodes equipped with wireless network connection can be deployed in the environment of the physical phenomenon. Those sensor nodes may provide raw data to the nodes responsible for data fusion or they may process the raw data by means of their computing capabilities and relay the required part of it to the other sensor nodes.

A large-scale CPS can be envisioned as millions of networked smart devices, sensors, and actuators being embedded in the physical world, which can sense, process, and communicate the data all over the network. Proliferation of technology-mediated social interactions via these highly featured and networked smart devices has allowed many individuals to contribute to the size of Big Data available. Depending on the size of data sets and number of smart devices involved, Big Data may be in the range of multiple terabytes to many petabytes (i.e. 1024 terabytes) [Sheth, 2013].

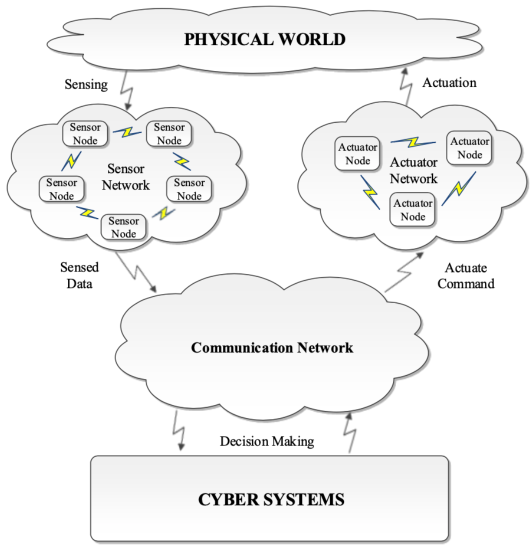


Figure 1- CPS Holistic view [Gunes, 2014]

The data generated by CPSs are contextualised, which makes them information. This makes CPSs, in the context of Industry 4.0, a huge source of information that brings with it, often implicitly, relationships about the environment and the working domain. This information and relationships are a potential source of knowledge that must be extracted, formalised and, potentially, reused.

The positioning of my research

The subject of my research has always been:

- the knowledge formalization in systems;
- the knowledge modelling activity in systems;
- the possibility to use this knowledge to characterize the semantic interoperability of the studied systems. Semantic interoperability can be defined as the ability for two or more systems to share, to understand and to consume information.

The increasing multiplication and complexity of the information necessary for the management of production processes pushes to the structuring of knowledge to accelerate its passage and optimize the interoperability of systems. In Figure 2 we can see all the different steps where the implicit knowledge of the systems is a brake to the knowledge passage itself between the various systems.

The first section of this introductory chapter highlighted the impact of the Fourth Industrial Revolution on the human economy and society. In the face of this new epochal change, two characterizations were highlighted:

- The importance of knowledge as a means of development and evolution. The information needed to manage production processes is increasingly numerous, more heterogeneous, more volatile and more distributed. This implies the use of business information systems increasingly linked to real processes in a continuous way in order to retrieve and process data, contextualize them into information and apply knowledge to improve performance.

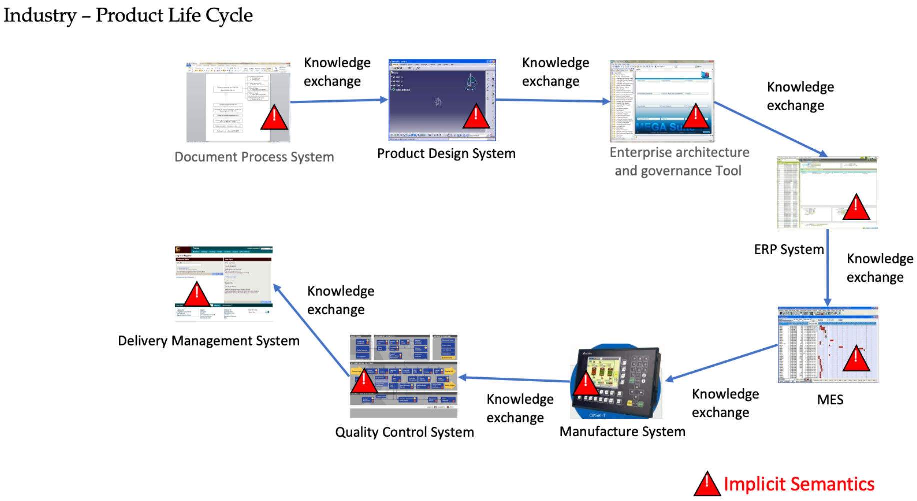


Figure 2 – Implicit semantics and semantic interoperability problem

- The key role that some technologies, such as cyber-physical systems, are playing in the restructuring of dominant roles in society.

The exploitation of the knowledge accumulated in the various systems involves two different issues. The first is the need to model systems so that they can semantically interoperate without problems of meaning. The second is to highlight methods to formalize and extract knowledge from all systems that are part of the value creation chain.

CPS modelling in cooperative information systems

The CPS, as we discussed, describes a broad range of network connected, multi-disciplinary, physically aware engineered systems that integrate embedded computing (cyber-) and technologies into the physical world (adapted from [Derler, 2013]). Inside this kind of network, each smart component (a sub-system of the CPS) is with sensing, data collection, transmission and actuation capabilities, and vast endpoints in the cloud, gathering and providing large amounts of heterogeneous data.
As presented, the CPSs are leading the Industry 4.0 with benefits from high flexibility of production, more accessible participation of all involved parties of business processes. The new manufacturing paradigm is characterized by autonomous behaviour and intercommunicating properties of its production elements across all levels of manufacturing processes.

At this regard the research directions, related to the CPS and the Industry 4.0 paradigm, take an important place, they focus on various scientific problems like the optimization of sensor networks organization by handling big datasets, challenges about the information and knowledge representation and processing. These research domains can benefit from scientific methods well known in the artificial intelligence domain, and machine learning. Focusing on this motivation I'm currently investigating the application of a mathematical approach named Formal Concept Analysis (FCA) for modelling and thus analysing a large-scale set of collaborative CPS.
One of the research interests I'm focusing on, is related to an extension of the FCA-based patterns for optimizing the interoperability in distributed systems, like the CPS, in the Industry 4.0.

Cooperative Enterprise Information Systems (CEIS) and interoperability issues
The Information Systems are systems whose activities are devoted to capture and to store data, to process them and produce knowledge, used by any stakeholders within an enterprise or among different networked enterprises. It is commonly agreed that Cooperative Information Systems (CIS) provide a backbone for the Integrated Information Infrastructure [Sheth, 1998].
The cooperative manufacturing systems involve large number of Information Systems distributed over large, complex networked architecture in relation to physical machines. Such cooperative enterprise information systems (CEIS) have access to a large amount of information and have to interoperate between them and with the machines to achieve their purpose. The CEIS architects and developers have to face a hard problem: interoperability at a large scale. There is a growing demand for integrating such systems tightly with organizational and manufacturing work so that these information systems can be fully, directly and immediately exploited by the intra and inter-enterprise processes [Izza, 2009].
Although the progress made in information technology considerably improved the efficiency of applications development, its drawbacks and limitations are obvious and serious. The components technologies are heterogeneous, platform- and machine-dependent. The above-mentioned limitations and barriers measurably hinder the development and the maintenance process.
There is a growing demand to integrate such systems tightly with organizational work so that these information systems can be directly and immediately used by the business activity.
Some work [Chen, 2006] in the INTEROP NoE project has identified three different levels of barriers for interoperability: technical, conceptual and organisational. Organisational barriers are still an important issue but out of scope of this paper. The technological barriers are strongly studied by researchers in computer science and the solution is generally based on model transformation [Frankel, 2003].
My, past and actual, research [Lezoche et al, 2011] focuses on the conceptual level of interoperability that is the ability to understand the exchanged information. A concept is

a cognition unit of meaning [Vyvyan, 2006], an abstract idea, a mental symbol. It is created through the action of conceptualisation, that is, a general and abstract mental representation of an object. During the history of human effort to model knowledge, different conceptualisation approaches regarding different application domains were developed [Aspray, 1985].

When trying to assess the understanding of an expression coming from a system to another system, there are several possible levels of interoperability [Euzenat, 2001]:

- *encoding:* being able to segment the representation in characters;
- *lexical:* being able to segment the representation in words (or symbols);
- *syntactic:* being able to structure the representation in structured sentences (or formulas or assertions);
- *semantic:* being able to construct the propositional meaning of the representation;
- *semiotic:* being able to construct the pragmatic meaning of the representation (or its meaning in context).

This structure is coherent, each level cannot be achieved if the previous levels have not been completed [Euzenat, 2001]. The encoding, lexical and syntactic levels are the most effective solutions for removing technical barriers for interoperability, but not sufficient, to achieve a practical interoperability between computerized systems. Dealing with trying to enable a seamless data and model exchange at the semantic level is still a big issue that needs conceptual representation of the intended exchanged information and the definition of the pragmatic meaning of that exchanged information in the context of the source and destination applications.

To achieve the purpose of the cooperation between the different Information Systems, information must be physically exchanged (technical interoperability), must be understood (conceptual interoperability) and must be used for the purpose that they have been produced (conceptual and organizational interoperability).

Classifying interoperability problems [Panetto, 2007] and [Panetto, 2008] may help in understanding the degree of development needed to solve, at least partially, these problems but conceptualization and semantics extraction is still an important issue because of the various contextual understanding of tacit knowledge embedded into those applications. The main prerequisite for achievement of interoperability of information systems is to maximize the amount of semantics which can be used and make it increasingly explicit [Panetto, 2008], and consequently, to make the systems semantically interoperable.

The main prerequisite for achieving the interoperability of information systems (and thus a set of collaborative CPSs is to maximize the amount of semantics that can be used and to enact it by making it increasingly explicit [Obrst, 2003]. There are different approaches in conceptual modelling and these differences are reflected in the conceptual languages used for the modelling action. Entity-Relationship approaches (E-R) have been widely used and extended. They led to the development of different languages for data modelling [Barker, 1990], [Czejdo, 1990] and [Hohenstein, 1991], Object-Oriented

Modelling (OOM)) [Rumbaugh, 1991] approach addresses the complexity of a problem domain by considering the problem as a set of related, interacting Objects. However, the abstract semantics inherent to these approaches imposes the modeller to make subjective choices between entities, attributes and relationships artefacts for modelling a universe-of-discourse [Lezoche et al, 2012b]. In order to cope with such heterogeneous modelling patterns, we focus our interest on approaches that enable their normalization to a fine-grained semantic model by fragmenting the represented knowledge into atoms called formal concepts.

A Meta-Model of a Cyber-Physical System

The components of a CPS: lets denote as P_i and C_i respectively the set of physical and cyber components of a system CPS_j. CPS_j is a structural agglomerate of these elements P_i and C_i which can also include other subsystems CPS_k into a composite cyber-physical system.

There are two relations of different nature between these components:

R^P - the relation between subsystems to be physically connected (e.g. in a production line) and signifies transmission of any kind of physical object between systems.

R^C - the relation of the connection between cyber components which signifies presence of an information/control channel between the components.

The components of a system perform certain functions depending on their role in the system, and according to that they have input I_i and output O_i, that capture the flows between this element and the elements that it is related to by R^P and R^C. As an example, for a sensor, input and output reflects transformation of mechanical or physical alterations of the physical world into quantitative measurements of a particular property. The source and destination of the exchange can be either other components of the system or the environment or some external source. To cover the latter case, we introduce, into the sets of all physical and all cyber elements of CPSs model, two elements e_P and e_C to stand for those kinds of sources or destinations.

We define a system of CPSs as a tuple CPSs = $\langle \mathcal{P}, \mathcal{C}, \mathcal{CPS}, R^P, R^C \rangle$, where $\mathcal{P} = e_P \cup \bigcup_i P_i$ is a set uniting physical components of individual CPS, $\mathcal{C} = e_C \cup \bigcup_i C_i$ is a set of cyber components, and $\mathcal{CPS}$ which is set of CPSs. Each individual CPS of the set $\mathcal{CPS}$, as was defined before, is a tuple of subset of cyber components, physical components and other CPS that it consolidates. Here we assume that every element of $\mathcal{P}, \mathcal{C}$ and $\mathcal{CPS}$ has its corresponding input I_i and output O_i. In general, I_i and O_i can be of any type and have any values.

Compositionally, different CPSs $c_1, c_2 \in \mathcal{CPS}$ could share some of their components: $c_1 \cap c_2 \neq \emptyset$. For example, as in systems utilizing the same computational node to supervise physical production activities, or as an actuator such as light switch which can be considered as a part of two systems: one is local electrical circuit of an apartment, and the other is a smart-home system for automating and controlling the household electronics. Following figure is an example of two simple CPSs consisting of one physical and one cyber component each forming a composite CPS. Where the communication between CPS_1 and CPS_2 is done through the C_1 and C_2; and the composite CPS_3 has its own actuator components P_3 and C_3.

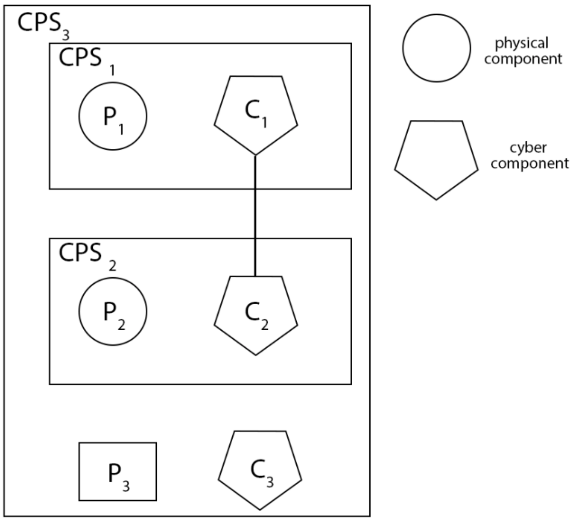


Figure 3 - Composite CPS_3 consisting of two subordinate systems CPS_1 and CPS_2, and providing its proper functionality through components P_3 and C_3

The proposed meta-model CPSs = $\langle \mathcal{P}, \mathcal{C}, \mathcal{CPS}, \mathrm{R}^{\mathrm{P}}, \mathrm{R}^{\mathrm{C}} \rangle$ that we have elaborated is presented in UML 2.0 notation on Figure 4. In the scientific literature, some authors have proposed different results related to CPS meta models from different points of view. [Jeon, 2012] present the CPS Meta Modeller tool for designing complex and large-scale systems using the Electronics and Telecommunication Research Institute (ETRI1) CPS Modelling Language. [Mezhuyev, 2013] show the geometrical meta-metamodel, allowing to link the physical properties of domains with its spatial structure. Some other authors [Son, 2012] show how to transform a Simulink model into an ETRI CPS Modeling Language (ECML) model for modelling CPS for simulating its behaviour and [Klimeš, 2014] shows how to control cyber-physical systems deriving behavioural specifications from user inputs. All these researches focused on the design and the internal behaviour of a CPS. The work I presented is a formal meta-model of the structure of any CPS, for proposing a common formal foundation of a composite CPS, aggregating the broad definitions found in the literature.

The elaborated meta model finds its focus in the interaction of the cyber component, which naturally stands for its computational functionality, and the physical component, which models its physical behaviour. The existence of those two entities let emerge the

[1] https://www.etri.re.kr/eng/main/main.etri

concept of Cyber-Physical System. If one of those components doesn't exist there is no possibility to have a CPS. The Physical component is modelled as an abstract class, it could be composed by a sensing component, an actuating component or by a component that merges the two capabilities. A CPS component needs an input and an output. It cannot exist a CPS component that has got only one of those two properties. An atomic CPS is the one that does not have any subsystems, but his own functional elements. This definition is created to show, with the best detail possible, the relationships between the two different parts of the entire CPS system. It stops at the presented "atomic" level because of the scope of my actual scientific interest that focuses on the relationships of the CPSs and the possibility to improve their interoperability.

I didn't specify the difference between physical and cyber types of communication and the corresponding types of input and output interfaces, although it could be a worthwhile extension for a future model.

The relation 'is part of' (physically) is introduced into the model to represent physical structure of systems and their inclusion into one another on the physical level. As an extension to this type of composition of complex CPS we also introduce the aggregation relation 'logically includes'. Together with inheritance relation between classes Composite CPS and Atomic CPS it complies to the structural Composite pattern. With the help of this aggregation relation, I modelled the property of CPS of dynamical reconfiguration and adaptation. Any system can lend its functionality to many super-systems (I borrow the utilisation of sub- and super- prefixes from mathematics, by analogy with subsets and supersets), although probably not at the same time. Inversely, any system can accommodate multiple subsystems.

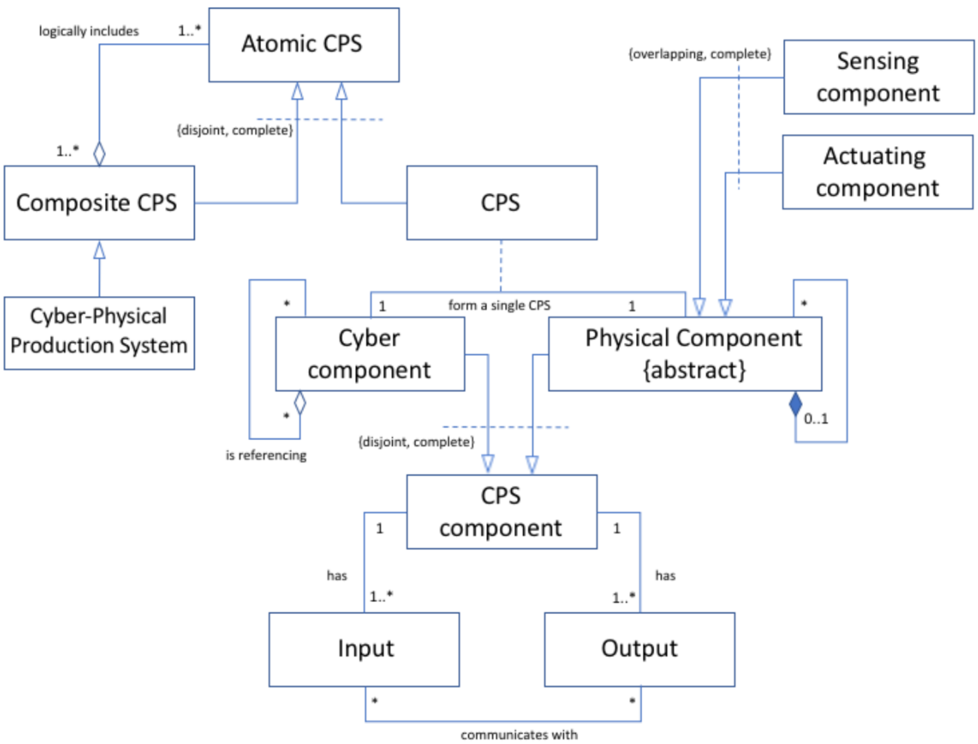


Figure 4 – CPS Meta-Model from the paper [IJ7]

The class of Cyber-Physical Production Systems can be viewed, in the proposed meta-model, as a subclass of Composite CPS. This interpretation goes in the same direction of the Monostori definition [Monostori, 2014]. There is a tight connection between these two relations 'is physically part of' and 'logically includes', in the sense that whenever a system is in the relation 'is physically part of' this also entails that it is being 'logically included' in that system, but not in the other direction.

Hierarchical structures of the meta-model and corresponding algebraic lattice representation

Unlike our previous approach [Morozov et al, 2015] where we modelled CPS using Formal Concept Analysis in a standard object-attribute fashion, currently we extend modelling approach to also account for links that exist between components and also for hierarchical inclusion of systems one into another according to their composite structure.

In this way CPSs can be modelled independently in the physical and cyber perspectives using corresponding relations, each one defining an algebraic lattice. The hierarchical structure gives rise to the third lattice that can be used for tacit knowledge recognition and further explicitation. The relations between the two Lattices are related to the context. The merging of the two Lattices is computationally expensive.

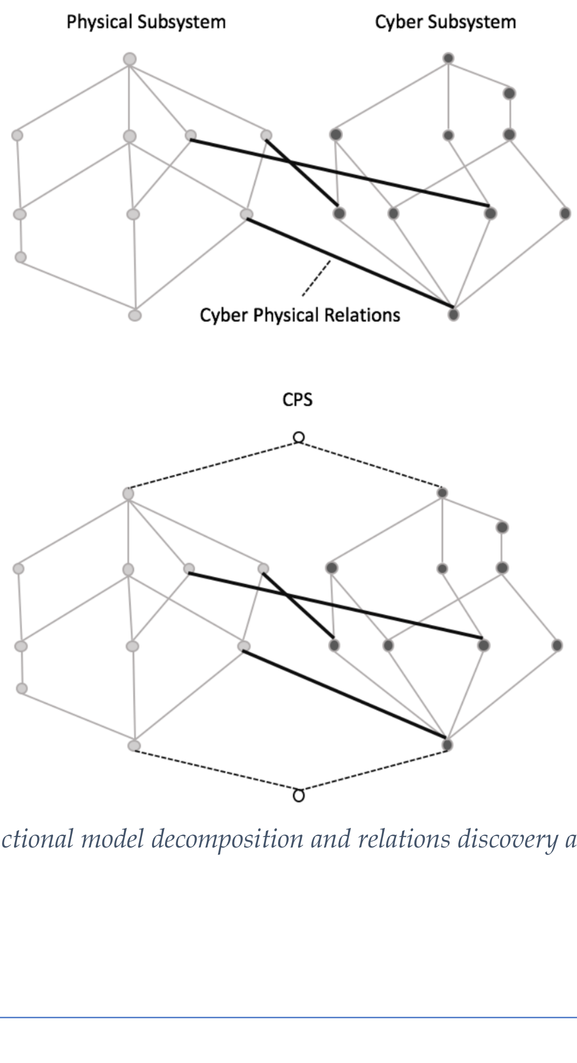


Figure 5 - CPS functional model decomposition and relations discovery and representation

FCA and RCA: Knowledge extraction methods

Understanding the domain behind the data is a key to business growth and competitiveness. Knowledge discovery from data (KDD) helps addresses that concern by distilling trends and patterns that are intelligible to human experts [Dehaspe, 1999]. In industry, data objects are typically unlabelled and often comprise both proper features and object-to- object links. Such datasets fit the unsupervised multi-relational data mining (MRDM) mode [D'Aquin, 2011], i.e. clustering and association discovery. However, existing MRDM association miners [Buitelaar, 2005], [Dehaspe, 2001] and [Ferré, 2015] restrict their output format to singleton-premise rules, hence they fail to capture more subtle associations.
Formal concept analysis (FCA) [Džeroski, 2003] has been proven as a versatile framework for KDD [Kramer, 2001] in many practical applications [Baader, 2007]. It extracts knowledge as a compact set of association rules [Goethals, 2002]. Relational concept analysis (RCA) [Rouane-Hacene, 2013] is MRDM extension of FCA. However, straightforwardly defined relational association rules may easily contain circular references or references from conclusion to premise, thus preventing a meaningful interpretation.

Formal concept analysis [Džeroski, 2003] is an algebraic approach for eliciting the conceptual structure of a dataset. Input data format is a triple K = (O, A, I) called a (formal) context. O is a set of objects, A is a set of attributes and $I \subseteq O \times A$ an incidence relation listing valid pairs (o, a) (object o has the attribute a). FCA reveals all pairs of sets $(X, Y) \in \wp(O) \times \wp(A)$ strongly correlated, meaning that all objects having the attributes in Y are in X and vice-versa. Such pair is a (formal) concepts with an extent X and intent Y.

Relational concept analysis assumes datasets are made of several contexts, one per type of object, and context-to-context relations. Any relational intent can be described with only non-relational attributes. Such expansion avoids circular dependencies, even if one may exist between full intents. We will present FCA and RCA deeper in the Chapter 2.

Research issues

The main focus of my research has always been how to formalize the implicit knowledge in models representing systems of all kinds. Since 2012, the year of my integration at CRAN, semantic interoperability has been added to my interests, bringing a great openness of methods and possible developments.

An approach to interoperable systems engineering must be based on different types and levels of abstraction or models. These models must express and formalize not only the "structural" aspect of the system components but also their behaviour, which may be constrained by requirements specific to the system domain (business rules). Another type of constraint may be induced by the interoperation protocol(s), which may impose strict rules to endow interoperating systems with properties such as autonomy, confidentiality and transparency [Lezoche et al, 2012b].

The goal of my research is to study the problems posed by model-driven cooperative systems engineering, the cooperation concerning "actors" (organizations, design teams, software systems, etc.) willing to interoperate. Collaborative enterprises are now organized in enterprise networks, either as extended or virtual enterprises [Bititci, 2004], [Camarinha, 2008].

Defined as such, these collaborative enterprises can be likened to an open enterprise network system [Oberndorf, 1998], and by extension the Information System (IS) of this network can itself be considered an IS network system. The specification of such an IS network implies a shift from a single integration paradigm to an interoperation paradigm [Fisher, 2006].

One of the requirements of this need for collaboration concerns the capacity of these components to interoperate, i.e. their interoperability, which is more or less total. Ducq [Ducq, 2008] considers the interoperability of systems as a particular performance requirement of the company.

There are standards and reference tools providing practices and metrics to measure this interoperability [IJ4] [Liao et al, 2016]. Several studies have proposed maturity models and formal metrics to assess the potential or degree of semantic interoperability of enterprises wishing to establish a collaborative network [Ducq, 2008]. However, these results do not allow a complete automation of the evaluation process because they suffer from a computer formalization of their models.

The scientific challenge is a computable formalization of the models and to make available and extend mathematical and modelling languages and tools adapted to each enterprise modelling project, despite the heterogeneity of business skills and the multidisciplinarity of the domains.

The analysis of formal concepts is a useful and powerful tool to formally describe the links between any objects (which form a context), in particular, as said, between objects carrying knowledge. The RCA, the FCA extension, method is not limited to the extraction of knowledge from distinct contexts: it aims to express knowledge by interacting the semantics of the different contexts, i.e. in addition to extracting knowledge from one context, data contained in other contexts are used to enrich the knowledge extraction.

One of the possible applications is related to the extraction of knowledge from complex systems. As Obendorf [Oberndorf, 1998] indicates, companies can be likened to complex systems. Currently, companies are focusing their interest on the Industry 4.0 paradigm to find a dynamic and rapid way to optimize their production. Formalization and optimization actions to deal with their procedures according to the parameters of Industry 4.0 are highly time-dependent [Ocampo-Martinez, 2017].

Document structure

Chapter 1 of this document presents a summary of activities related to teaching and research. After a short curriculum vitae, the entire course of teaching activities is presented, including administrative responsibilities. To conclude, all research-related activities are presented, such as the summary of publications made, participation in national and international working groups, links with other international research groups, participation in research projects and responsibility in relation to the scientific world. This chapter should make it possible to quantitatively evaluate my teaching and research activities.

Chapter 2 details my research work since 2012, the year of my recruitment as associate professor (maître de conférences). This part of the document will present how the current vision of my research theme has been built up over the years by the possibility of supervising students who are eager to explore the fields of research most interesting to me. The semantic interoperability of models has been a fascinating domain with which I started my scientific career as a professor and it has declined, in the first two theses I supervised, in how semantic annotations are useful for semantic interoperability and how the contribution of semantics in the creation of models of decision-making processes can optimize the use of the resources of the examined systems. The following two theses were the watershed of my scientific interests. Together with semantic interoperability, I have combined a domain to which I am strongly attached, the formalization, extraction and reuse of knowledge in industry 4.0 contexts. The third thesis focused on the formalization of data-driven knowledge within invariant structures of cyber-physical systems. The fourth thesis focused on more theoretical issues by solving some questions for the optimization of association rules extraction that are useful for the explicit expression of implicit knowledge in enterprises related Big Data. This chapter should allow a qualitative evaluation of my research activities.

Chapter 3 describes my research project concerning the formalization and extraction of knowledge applied to heterogeneous data sources in the context of industry 4.0. First of all, this chapter takes up the concepts stated in the introduction on the need to optimise knowledge management in an environment like industry 4.0 which is completely based on the optimisation of sharing and implementation of actions deriving from knowledge generated from heterogeneous sources. The aim of this project is to contribute, through the results of my research, to define formalization and knowledge extraction methods specific to the multimodal environment defined by industry 4.0. The proposed project is strongly supported by the tools and results presented in chapter 2.
This project is then contextualised in the local, regional, national and international context. The chapter ends with the presentation of my teaching project which is characterized on the one hand by my responsibilities at IUT (University Institute of Technology) Hubert Curien d'Épinal level in the QLIO (Quality, Industrial Logistics and

Organization) department where I manage all the information technology paths at industrial level, both at DUT (University Technological Diploma) and Licence Professionnelle (Professional Faculty of the first University level course) level, and on the other hand by my contribution in the engineering school Telecom Nancy where I can structure courses that are in line with the use of tools directly related to knowledge formalization and management in industry 4.0.

1 Summary of teaching and research activities

Consideré que incluso en el lenguaje humano no hay ninguna proposición que no implique a todo el universo; decir "el tigre" es decir los tigres que lo generaron, los ciervos y las tortugas que devoró, el pasto que alimentó a los ciervos, la tierra que fue la madre del pasto, el cielo que dio nacimiento a la tierra.

I considered that even in human language there is no proposition that does not imply the whole universe; to say "the tiger" is to say the tigers that generated it, the deer and tortoises it devoured, the pasture that fed the deer, the land that was the mother of pasture, the sky that gave birth to the earth.

Jorge Luis Borges – Aleph

1.1 Curriculum Vitae

Mario LEZOCHE

Personal details

- Born the 3th of April of 1974 in Santa Maria Capua Vetere (Italy)
- Double nationality: French and Italian
- Married the 1st of August 2020

Personal contact details

- 4, rue Joseph Piroux, 54140, Jarville-La-Malgrange, France
- Mobile : +33 (0)6 24 58 97 37
- Personal web site: http://www.lezoche.eu

Professional contact details

- CRAN, Campus Sciences, BP 70239, Faculté des Sciences, 54506 Vandoeuvre cedex
- Téléphone : +33 (0)3 72 74 53 23
- Email : mario.lezoche@univ-lorraine.fr
- Web: http://www.cran.univ-lorraine.fr

1.1.1 Academic Titles

2011	**Post-doc's Degree** in Computer Science Engineering, Université Henri Poincaré – Nancy I, CRAN laboratory
2009	**Ph.D Thesis** at CNR of Rome and University of "Roma Tre"
2006	**Research Master's Degree** in Computer Science Engineering, University of Roma TRE

Table 1 – Academic Titles

1.1.2 Administrative situation

- Associate professor (Maître de Conférences) normal class, echelon 6 on the national post 61MCF0693.
- I am a teaching member of the University Institute of Technology (IUT) Hubert Curien of Épinal at the department of Quality, Industrial Logistics and Organization (QLIO) since the 1st September 2012.
- I am researcher at
    - the Research Center for Automatic Control (CRAN) that is a joint research unit (UMR 7039) between the University of Lorraine and the French National Scientific Research Center (CNRS) - Institute for Information

Sciences and Technologies (INS2I) and Institute for Engineering and Systems Sciences (INSIS) directed by the Full professor Didier WOLF;
    - I am in the Eco-Technic systems engineering (ISET) department coordinated by Full Professor Hind BRIL-EL HAOUZI and Full Professor Benôit IUNG;
    - I collaborate in the coordination of the research project team Intelligent System and Objects in Interaction (S&O-2I) with the Associate Professor William DERIGENT and Associate Professor Philippe THOMAS;
- I am bound to the session 61 to the CNU

1.1.3 Career summary

Date	Highlight
12 / 2010 09 / 2012	Appointed as Post Doc at the University of Lorraine Post Doc Researcher at CRAN UMR 7039 CNRS
09 / 2012	Appointed as professor at the University of Lorraine at the IUT Hubert Curien d'Épinal Researcher at CRAN UMR 7039 CNRS
09 / 2013	Tenure
11 / 2013	PhD Thesis of Yongxin LIAO Title: "Semantic annotations for system interoperability in a PLM context" Directors: Hervé PANETTO and Nacer BOUDJLIDA Co-Tutor: Mario LEZOCHE
2013 - 2016	Coordinator of the Professional Licence Industrial Production and Development of the Innovation Approach (PIDDI)
2013 - 2016	Elected to the Institute Council of the IUT Hubert Curien d'Épinal Teacher-researcher representative
2016	Chair of the workshop Enterprise Integration, Interoperability and Networking (EI2N) in Rhodes
09 / 2017	Obtaining the doctoral supervision and research grant (PEDR) Level B (CNU 61 session)
11 / 2017	PhD Thesis of Silvana DETRO Title: "A framework for interoperability assessment in E-Health information systems using process semantics mining" Directors: Hervé PANETTO Co-director: Mario LEZOCHE
2017	Transformation of the PIDDI Professional Licence in E-Commerce and Digital Marketing (ECMN) Professional Licence New accreditation of the ECMN Professional Licence
2017	Chair of the workshop Enterprise Integration, Interoperability and Networking (EI2N) in Rhodes
2017 – 2020	Co-Direction of the ECMN Professional Licence

2017 - 2020	Member of the Directors' board at the IUT Hubert Curien d'Épinal
09 / 2017 03 / 2018	Leave for Research or Thematic Conversions (CRCT) at the Universitat Politècnica de Catalunya
06 / 2019	PhD Thesis of Yasamin ESLAMI Title: "A Modelling-Based Sustainability Assessment in Manufacturing Organizations" Directors: Hervé PANETTO, Michele DASSISTI Co-director: Mario LEZOCHE
06 / 2020	PhD Thesis of Concetta SEMERARO Title: "Contribution to the formalisation of data-driven invariant modelling constructs of Cyber-Physical Systems" Directors: Hervé PANETTO, Michele DASSISTI Co-director: Mario LEZOCHE
10 / 2020	Started the PhD Thesis of Yandé NDIAYE Title: "Knowledge discovery and formalisation for Additive manufacturing through Artificial Intelligence and Information Retrieval methods" Directors: Hervé PANETTO, Yan LU Co-director: Mario LEZOCHE
11 / 2020	PhD Thesis of Mickael WAJNBERG Title: "Relational Concept Analysis: a versatile knowledge extraction method" Directors: Hervé PANETTO, Alexandre BLONDIN MASSÉ Co-director: Mario LEZOCHE

Table 2 – Career summary

1.2 Teaching-related activities

1.2.1 General presentation

Since the day I was hired in 2012, I have made my teaching career mainly at the IUT Hubert Curien in Épinal. The majority of my teaching service is done in the Quality, Industrial Logistics and Organization (QLIO) department and in the two Professional Licences, the Industrial Logistic (MILI) one and the one I was managing, the Industrial Production and Development of the Innovation Approach (PIDDI) until 2016. The PIDDI Licence was then transformed into an E-Commerce and Digital Marketing Licence and which I co-directed until today (2020). During my Post-Doc I performed also the temporary teaching and research associates (ATER) during the years of Post-doc and thanks to that experience I continued to teach at Telecom Nancy computer engineering school.

My teaching is mainly focused on three topics, Databases, information management in enterprises through business information systems such as ERP, ESM and knowledge formalization through the teaching of logic and ontologies.

If the teaching in the QLIO department is quite standardized with a uniform program at national level, the teachings in Licence pro and engineering school have been shaped by my desire to teach through concrete projects that integrate the more theoretical parts. I have built the modules used for teaching the formalization of knowledge through Ontologies and also those related to algorithmic learning.

In 2016 I built up the Licence pro E-Commerce and Digital Marketing accreditation dossier together with the co-responsible of the Licence.

I have also invested heavily in accompanying students for the DUT 2A, TN 2A and TN 3A projects.

1.2.1 Administrative Responsibilities

1.2.1.1 Direction and animation of teaching courses

- From 2013 to 2016 responsible of the PIDDI Professional Licence
- From 2014 to 2017 responsible for the specialization in Enterprise Information Systems (SIE) at the School of Engineering Telecom Nancy.
- From 2017 to 2020 co-responsible for ECMN Professional Licence

1.2.1.2 educational administration management

- From 2012 member of the admission commission of the DUT QLIO
- As of 2012 member of the DUT QLIO improvement council
- From 2013 to 2016 member of the Admission Commission of the PIDDI Professional Licence

- From 2013 to 2020 member of the directive IUT Institute council
- From 2013 to 2016 Elected to represent teacher-researchers on the IUT Institute Council.
- From 2013 to 2016 member of the restricted Committee of the Institute Council IUT
- From 2013 to 2016 member of the PIDDI Professional Licence Development Council
- From 2018 to 2020 member of the admission commission of the Professional Licence ECMN
- From 2018 to 2020 member of the ECMN Professional Licence Development Board

1.2.2 Summary presentation of the teachings 2012 - 2020

All my teachings at DUT level, Licence and engineering school are in frontal mode and combine initial, continuing and professional training. On average my annual teaching service counts more than 300 hours H.Eq.TD without counting the hours calculated for administrative responsibilities.
The three diagrams of the Figure 6 present the division of my teaching between 2012 and 2020 according to the topic (Data Management, Information Systems, Knowledge Formalization, Computer Engineering), according to the nature of the teaching (CM, TD, TP) and finally according to the level of the diploma.
The hours associated with student placement are the paid hours and are therefore not representative of the associated investment. For this reason, they are not represented in the diagrams. The Table 3 does not count the hours related to administrative responsibilities in the various academic formations.

Theme		Level	CM	TD	TP	Total	Total Eq.TD	Total by Theme	Total Eq TD by Theme
Data Management		DUT 1	80	96	64	240	280	1496	1628
		DUT 2	32	32	176	240	256		
		LP (L3)	32	64	240	336	352		
		M1	120	160	400	680	740		
Informations Systems		DUT 1	96	112	32	240	288	568	640
		M1	32	64	96	192	208		
		M2	16	40	80	136	144		
Knowledge formalisation		LP (L3)	32	48	160	240	256	368	392
		M2	16	32	80	128	136		
Computer Engineering		M1	32	64	64	160	176	160	176
Student Supervision	Projets	DUT 2	0	48	0	48	48	192	192
		LP (L3)	0	48	0	48	48		
		M1	0	96	0	96	96		
	Stage	DUT 2	0	48	0	48	48	360	360
		LP (L3)	0	168	0	168	168		
		M1	0	72	0	72	72		
		M2	0	72	0	72	72		
		Total	**488**	**1264**	**1392**	**3144**	**3388**		

Table 3 - Teaching and student supervision between 2012 and 2020.

More than half of my teaching (58%) concerns data management, which is one of the axes of my research, together with the formalisation of knowledge (14%). Half of my teaching, during these eight years, has been for the school of computer engineering students, therefore for a level M1 (40%) and M2 (10%). The other half of my teaching was offered to the university primary cycle students. This has allowed me to develop a pedagogy approach that allows me to get in touch with the whole range of levels of university students.

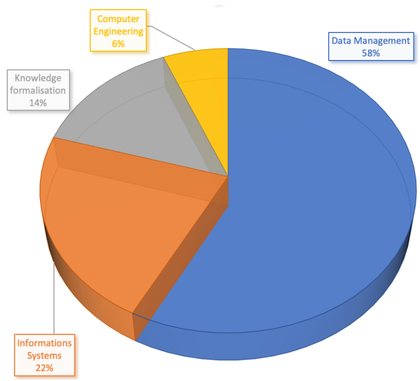


Figure 1a

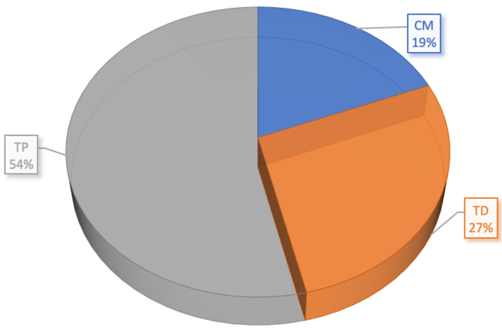


Figure 1b

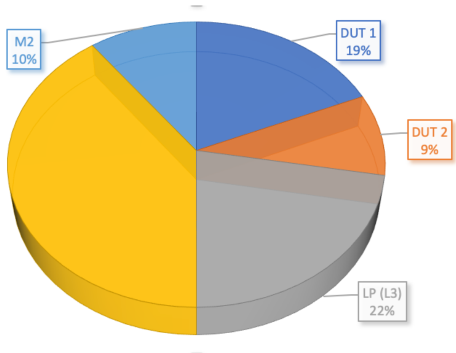


Figure 1c

Figure 6 (a,b,c) present the teachings division between 2012-2020 according to the topic, nature and level

1.2.3 Participation in academic training

- 2012 - 2013
  - DUT QLIO 1 (L1): Data Management, Computer engineering
  - DUT QLIO 2 (L2): Information Systems
  - Engineering School Telecom Nancy (M1): Data management
- 2014 – 2016
  - DUT QLIO 1 (L1): Data Management, Computer engineering
  - DUT QLIO 2 (L2): Information Systems
  - PIDDI Professional Licence (L3): Data management
  - MILI Professional Licence (L3): Data management
  - Engineering School Telecom Nancy (L3): Computer engineering
  - Engineering School Telecom Nancy (M1): Data management
  - Engineering School Telecom Nancy (M2): Data management, Information systems, Knowledge formalisation

    - Master ISC (M1): Knowledge formalisation
- 2017 – 2020
    - DUT QLIO 1 (L1): Data Management, Computer engineering
    - DUT QLIO 2 (L2): Information Systems
    - ECMN Professional Licence (L3): Data management, Knowledge formalisation
    - MILI Professional Licence (L3): Data management
    - Engineering School Telecom Nancy (M1): Data management
    - Engineering School Telecom Nancy (M2): Data management, Information systems, Knowledge formalisation

1.2.4 Responsibilities of Teaching Units

- Design of information systems (L1 - DUT 1)
- Algorithms (L1 – DUT 1)
- Artificial Intelligence (M1 – Telecom Nancy)
- Integrated business management software (M2 – Telecom Nancy)
- Enterprise 4.0 – Blockchain (M1 – Telecom Nancy)
- Organize and direct the production of knowledge (L3 – ECMN LP)
- Information Systems (L3 – MILI LP)
- Database management (L3 – ECMN)
- Use of an ERP (L2 – DUT 2)
- Database management (L2 – DUT 2)
- Database systems (L1 – DUT 1)

1.3 Research-related activities

1.3.1 Summary of publications

The Table 4 contains a summary of the publications to which I have contributed. The following section creates an exhaustive list of publications in which I have participated. My name will be in bold. Publications that refer to work with students that I have personally followed are recognisable by the fact that the students will be underlined. The impact factor of publications in JCR science journals will be in bold where known. It will be in reference to the year of publication. I will use the identifiers of these references throughout the document to cite them.

Type of publication	Number
International peer-reviewed journals (JCR)	9
International peer-reviewed journals (No-JCR)	1
Special issues journals	1
International scientific vulgarization journals	3
Conferences, congresses, international peer-reviewed conferences with proceedings	32
Direction of work and proceedings	2
Peer-reviewed journals with national circulation	1
Book chapter	3
Others	3
Total	**55**

Table 4 - summary of publications

1.3.1.1 International peer-reviewed journals (9)(JCR)

[IJ9] **Mario Lezoche**, Jorge Hernandez, Maria del Mar Alemany Diaz, Hervé Panetto, Janusz Kacprzyk. Agri-food 4.0: a survey of the supply chains and technologies for the future agriculture. Computers in Industry, Elsevier, 2020, 117:103187, ⟨10.1016/j.compind.2020.103187⟩. ⟨hal-02395411⟩

[IJ8] Yasamin Eslami, **Mario Lezoche**, Hervé Panetto, Michele Dassisti. On analysing sustainability assessment in manufacturing organizations: A survey. International Journal of Production Research, Taylor & Francis, 2020, ⟨10.1080/00207543.2020.1755066⟩. ⟨hal-02524117⟩

[IJ7] **Mario Lezoche**, Hervé Panetto. Cyber-Physical Systems, a new formal paradigm to model redundancy and resiliency. Enterprise Information Systems, Taylor & Francis, 2020, ⟨10.1080/17517575.2018.1536807⟩. ⟨hal-01895093⟩

[IJ6] Yasamin Eslami, Michele Dassisti, **Mario Lezoche**, Hervé Panetto. A survey on sustainability in manufacturing organisations: dimensions and future insights. International Journal of Production Research, Taylor & Francis, 2019, 57 (15-16), pp.5194-5214. ⟨10.1080/00207543.2018.1544723⟩. ⟨hal-01911366⟩

[IJ5] Silvana Pereira Detro, Eduardo Portela Santos, Hervé Panetto, Eduardo Loures de Freitas, **Mario Lezoche**. Applying process mining and semantic reasoning for process model customization in healthcare. Enterprise Information Systems, Taylor & Francis, 2018, ⟨10.1080/17517575.2019.1632382⟩. ⟨hal-02155320⟩

[IJ4] Yongxin Liao, **Mario Lezoche**, Hervé Panetto, Nacer Boudjlida. Semantic annotations for semantic interoperability in a product lifecycle management context. International Journal of Production Research, Taylor & Francis, 2016, 54 (18), pp.5534-5553. ⟨10.1080/00207543.2016.1165875⟩. ⟨hal-01286475⟩

[IJ3] Yongxin Liao, **Mario Lezoche**, Hervé Panetto, Nacer Boudjlida, Eduardo Rocha Loures. Semantic annotation for knowledge explicitation in a product lifecycle management context: a survey. Computers in Industry, Elsevier, 2015, 71, pp.24-34. ⟨10.1016/j.compind.2015.03.005⟩. ⟨hal-01123854⟩

[IJ2] **Mario Lezoche**, Esma Yahia, Alexis Aubry, Hervé Panetto, Milan Zdravković. Conceptualising and structuring semantics in Cooperative Enterprise Information Systems Models. Computers in Industry, Elsevier, 2012, 63 (8), pp.775-787. ⟨10.1016/j.compind.2012.08.006⟩. ⟨hal-00722419⟩

[IJ1] Esma Yahia, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Semantics enactment for interoperability assessment in Enterprise Information Systems. Annual Reviews in Control, Elsevier, 2012, 36 (1), pp.101-117. ⟨10.1016/j.arcontrol.2012.03.008⟩. ⟨hal-00671856⟩

1.3.1.2 International peer-reviewed journals (1)(No-JCR)

[IJN1] Yongxin Liao, **Mario Lezoche**, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. A semantic annotation framework to assist the knowledge interoperability along a product life cycle. Advanced Materials Research, Trans Tech Publications, 2014, 945-949, pp.424-429. ⟨10.4028/www.scientific.net/AMR.945-949.424⟩. ⟨hal-01026591⟩

1.3.1.3 Special issues of journals (1)

[IJSI1] Hervé Panetto, **Mario Lezoche**, Jorge Hernandez, Maria del Mar Eva Alemany Diaz, Janusz Kacprzyk. Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. Computers in Industry, Elsevier, 2020, 116:103188, ⟨10.1016/j.compind.2020.103188⟩. ⟨hal-02450378⟩

1.3.1.4 International scientific vulgarization journals (3)

[IJV3] Concetta Semeraro, Hervé Panetto, **Mario Lezoche**, Michele Dassisti, Stefano Cafagna. A monitoring strategy for industry 4.0: Master italy s.r.l case study. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2019, Systems engineering research at French Universities, 22 (4), pp.20-22. ⟨10.1002/inst.12269⟩. ⟨hal-02423272⟩

[IJV2] Mickael Wajnberg, **Mario Lezoche**, Blondin Alexandre Massé, Petko Valtchev, Hervé Panetto. Complex system tacit knowledge extraction trough a formal method. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2017, 20 (4), pp.23-26. ⟨10.1002/inst.12176⟩. ⟨hal-01673069⟩

[IJV1] Silvana Pereira Detro, Eduardo Portela Santos, Hervé Panetto, Eduardo Rocha Loures, **Mario Lezoche**. Configuring process variants through semantic reasoning in systems engineering. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2017, 20 (4), pp.36-39. ⟨10.1002/inst.12179⟩. ⟨hal-01673070⟩

1.3.1.5 Conferences, congresses, international peer-reviewed conferences with proceedings (32)

[C32] Mickael Wajnberg, Petko Valtchev, **Mario Lezoche**, Hervé Panetto, Alexandre Blondin Masse. Mining process factor causality links with multi-relational associations. 10th International Conference on Knowledge Capture, K-CAP'19, Nov 2019, Marina Del Rey, CA, United States. pp.263-266, ⟨10.1145/3360901.3364446⟩. ⟨hal-02377662v2⟩

[C31] Mickael Wajnberg, Petko Valtchev, **Mario Lezoche**, Alexandre Blondin Masse, Hervé Panetto. Concept analysis-based association mining from linked data: A case in industrial decision making. 2nd International Workshop on Data meets Applied Ontologies in Open Science and Innovation, DAO-SI 2019, Sep 2019, Gratz, Austria. ⟨hal-02455243⟩

[C30] 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⟩

[C29] Shaofeng Liu, Guoqing Zhao, Huilan Chen, Alejandro Fernandez, Diego Torres, **Mario Lezoche,** Knowledge mobilisation crossing boundaries: a multi-perspective framework for agri-food value chains. 6th Model-IT International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain, Jun 2019, Molfetta, Italy. ⟨hal-02191439⟩

[C28] Yasamin Eslami, Michele Dassisti, Hervé Panetto, **Mario Lezoche**. Sustainability assessment of manufacturing organizations based on indicator sets: A formal concept analysis. OTM/IFAC/IFIP International Workshop on

Enterprise Integration, Interoperability and Networking, EI2N 2018, Oct 2018, Valletta, Malta. pp.36-44, ⟨10.1007/978-3-030-11683-5_4⟩. ⟨hal-02047367⟩

[C27] Dmitry Morozov, **Mario Lezoche**, Hervé Panetto. Multi-paradigm modelling of Cyber-Physical Systems. 16th IFAC Symposium on Information Control Problems in Manufacturing, INCOM 2018, Jun 2018, Bergamo, Italy. pp.1385-1390, ⟨10.1016/j.ifacol.2018.08.334⟩. ⟨hal-01813386⟩

[C26] Mickael Wajnberg, **Mario Lezoche**, Alexandre Blondin Masse, Petko Valtchev, Hervé Panetto, Semantic interoperability of large systems through a formal method: Relational Concept Analysis. 16th IFAC Symposium on Information Control Problems in Manufacturing, INCOM 2018, Jun 2018, Bergamo, Italy. pp.1397-1402, ⟨10.1016/j.ifacol.2018.08.330⟩. ⟨hal-01813398⟩

[C25] Silvana Detro, Eduardo Santos, Hervé Panetto, Eduardo Freitas Rocha Loures, **Mario Lezoche**. Managing business process variability through process mining and semantic reasoning: an application in healthcare. 18th Working Conference on Virtual Enterprises, PRO-VE 2017, Sep 2017, Vicenza, Italy. pp.333-340, ⟨10.1007/978-3-319-65151-4_31⟩. ⟨hal-01674899⟩

[C24] Pierre Monnin, **Mario Lezoche**, Amedeo Napoli, Adrien Coulet. Using formal concept analysis for checking the structure of an ontology in LOD: the example of DBpedia. 23rd International Symposium on Methodologies for Intelligent Systems, ISMIS 2017, Warsaw university of technology, Jun 2017, Warsaw, Poland. pp.674-683, ⟨10.1007/978-3-319-60438-1_66⟩. ⟨hal-01511909⟩

[C23] Michele Dassisti, Hervé Panetto, **Mario Lezoche**, Pasquale Merla, Concetta Semeraro, Industry 4.0 paradigm: The viewpoint of the small and medium enterprises. 7th International Conference on Information Society and Technology, ICIST 2017, Mar 2017, Kopaonik, Serbia. pp.50-54. ⟨hal-01526397⟩

[C22] Dmitry Morozov, **Mario Lezoche**, Hervé Panetto. FCA modelling for CPS interoperability optimization in Industry 4.0. 7th International Conference on Information Society and Technology, ICIST 2017, Mar 2017, Kopaonik, Serbia. pp.55-59. ⟨hal-01526383⟩

[C21] Giuseppina Uva, Michele Dassisti, Francesco Iannone, Franco Maddalena, Michele Ruta, **Mario Lezoche**, Modelling framework for sustainable co-management of multi-purpose exhibition systems: the “Fiera del Levante” case. International High-Performance Built Environment Conference, iHBE 2016 (A Sustainable Built Environment Conference 2016 Series - SBE16), Nov

2016, Sydney, Australia. pp.812-821, ⟨10.1016/j.proeng.2017.04.242⟩. ⟨hal-01637828⟩

[C20] Gabriel Leal, Wided Guédria, Hervé Panetto, Erik Proper, **Mario Lezoche**. Using formal measures to improve maturity model assessment for conceptual interoperability. 11th OTM/IFAC Enterprise Integration, Interoperability and Networking Workshop, EI2N, OM, Oct 2016, Kallithea, Rhodes, Greece. pp.47-56, ⟨10.1007/978-3-319-55961-2_5⟩. ⟨hal-01387780⟩

[C19] Gabriel Leal, Wided Guedria, Hervé Panetto, **Mario Lezoche**. Towards a comparative analysis of interoperability assessment approaches for collaborative enterprise systems. 23rd IPSE International Conference on Transdisciplinary Engineering, Oct 2016, Curitiba, Brazil. pp.45-54, ⟨10.3233/978-1-61499-703-0-45⟩. ⟨hal-01376442⟩

[C18] Silvana Pereira Detro, Dmitry Morozov, **Mario Lezoche**, Hervé Panetto, Eduardo Portela Santos, Enhancing semantic interoperability in healthcare using semantic process mining. 6th International Conference on Information Society and Techology, ICIST 2016, Feb 2016, Kopaonik, Serbia. pp.80-85. ⟨hal-01298125⟩

[C17] Milan Zdravković, Miroslav Trajanović, João Sarraipa, Ricardo Jardim-Gonçalves, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Survey of Internet-of-Things platforms. 6th International Conference on Information Society and Techology, ICIST 2016, Feb 2016, Kopaonik, Serbia. pp.216-220. ⟨hal-01298141⟩

[C16] Dmitry Morozov, Sergei O. Kuznetsov, **Mario Lezoche**, Hervé Panetto. Formal methods for process knowledge extraction. 10ème Colloque sur la Modélisation des Systèmes Réactifs, MSR 2015, Nov 2015, Nancy, France. ⟨hal-01224489⟩

[C15] José Marcelo Cestari, Eduardo Rocha Loures, Eduardo Santos, Yongxin Liao, Hervé Panetto, **Mario Lezoche**. An overview of attributes characterization for interoperability assessment from the public administration perspective. Third Workshop on Industrial and Business Applications of Semantic Technologies and Knowledge-based information systems, INBAST 2014, Oct 2014, Amantea, Italy. pp.329-338, ⟨10.1007/978-3-662-45550-0_33⟩. ⟨hal-01074447⟩

[C14] Yongxin Liao, **Mario Lezoche**, Hervé Panetto, Nacer Boudjlida, Eduardo Rocha Loures. Formal semantic annotations for models interoperability in a PLM environment. 19th IFAC World Congress, IFAC'14, International

Federation of Automatic Control, Aug 2014, Cape Town, South Africa. pp.2382-2393, ⟨10.3182/20140824-6-ZA-1003.02551⟩. ⟨hal-01082576⟩

[C13] José Marcelo Cestari, **Mario Lezoche**, Eduardo Rocha Loures, Hervé Panetto, Eduardo Portela Santos. A method for e-government concepts interoperability assessment. 4th International Conference on Information Society and Technology, ICIST 2014, Mar 2014, Kopaonik, Serbia. pp.79-85. ⟨hal-00993431⟩

[C12] Yongxin Liao, **Mario Lezoche**, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. Semantic enrichment of models to assist knowledge management in a PLM environment. 21st International Conference on Cooperative Information Systems CoopIS 2013, Sep 2013, Graz, Austria. pp.267-274. ⟨hal-00850145⟩

[C11] Yongxin Liao, **Mario Lezoche**, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. Formalization of Semantic Annotation for Systems Interoperability in a PLM environment. OTM Federated conferences and worlshops, 2nd Workshop on Industrial and Business Applications of Semantic Web Technologies (INBAST), Sep 2012, Rome, Italy. pp.207-218, ⟨10.1007/978-3-642-33618-8-207⟩. ⟨hal-00722740⟩

[C10] Yongxin Liao, **Mario Lezoche**, Hervé Panetto, Nacer Boudjlida. Semantic Annotation Model Definition for Systems Interoperability. OTM 2011 Workshops 2011 - 6th International Workshop on Enterprise Integration, Interoperability and Networking (EI2N), Oct 2011, Hersonissos, Crete, Greece. pp.61-70, ⟨10.1007/978-3-642-25126-9_14⟩. ⟨hal-00627909⟩

[C9] Milan Zdravković, Miroslav Trajanović, Hervé Panetto, Alexis Aubry, **Mario Lezoche**. Ontology-based supply chain process configuration. 34th International Conference on Production Engineering, ICPE 2011, Sep 2011, Nis, Serbia. pp.399-402. ⟨hal-00627531⟩

[C8] Esma Yahia, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Semantics enactment in Enterprise Information Systems. 18th IFAC World Congress, IFAC WC'2011, Aug 2011, Milan, Italy. pp.13064-13073, ⟨10.3182/20110828-6-IT-1002.03597⟩. ⟨hal-00582507⟩

[C7] Yongxin Liao, **Mario Lezoche**, Hervé Panetto, Nacer Boudjlida. Why, Where and How to use Semantic Annotation for Systems Interoperability. 1st UNITE Doctoral Symposium, Jun 2011, Bucarest, Romania. pp.71-78. ⟨hal-00597903⟩

[C6] **Mario Lezoche**, Hervé Panetto, Alexis Aubry. Conceptualisation approach for cooperative information systems interoperability. 13th International Conference on Enterprise Information Systems, ICEIS 2011, Jun 2011, Beijing, China. pp.101-110, ⟨10.5220/0003508401010110⟩. ⟨hal-00583866⟩

[C5] Esma Yahia, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Extraction de la sémantique dans les modèles de systèmes d'information d'entreprises collaboratifs. Journée Nationale du GT Easy-DIM, Apr 2011, Lyon, France. ⟨hal-00586779⟩

[C4] Francesco Taglino, **Mario Lezoche**. Business Process driven solutions for innovative enterprise information systems. itAIS 2008 - 5th Conference of the Italian Chapter of AIS, Dec 2008, Paris, France. pp.ISBN:978-88-6105-076-1. ⟨hal-00656678⟩

[C3] **Mario Lezoche**, Antonio de Nicola, Tania Di Mascio, Francesco Taglino. Semantic Lifting of Business Process Models. 12th Enterprise Distributed Object Computing Conference Workshops, Sep 2008, Munich, Germany. ISBN: 978-0-7695-3720-7. ⟨hal-00656690⟩

[C2] **Mario Lezoche**, Michele Missikoff, Leonardo Tininini. Business Process Evolution: a Rule-based Approach. 9th Workshop on Business Process Modeling, Development, and Support (BPMDS'08), Jun 2008, Montpellier, France. pp.ISSN 1613-0073 vol-335. ⟨hal-00656687⟩

[C1] Antonio de Nicola, **Mario Lezoche**, Michele Missikoff. An Ontological Approach to Business Process Modeling. 3th Indian International Conference on Artificial Intelligence 2007, Dec 2007, India. pp.ISBN 978-0-9727412-2-4. ⟨hal-00656686⟩

1.3.1.6 Direction of work and proceedings (2)

[DW2] **Mario Lezoche**, Hervé Panetto. New challenges and Advances in MBSE in French Universities. France. INSIGHT - International Council on Systems Engineering (INCOSE), 20 (4), Willey, pp.8-10, 2017, INSIGHT, 2156-4868. ⟨10.1002/inst.12172⟩. ⟨hal-01673066⟩

[DW1] Hervé Panetto, **Mario Lezoche**, Jorge Hernandez, Maria del Mar Eva Alemany Diaz, Janusz Kacprzyk. Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. Computers in Industry, Elsevier, 2020, 116:103188, ⟨10.1016/j.compind.2020.103188⟩. ⟨hal-02450378⟩

1.3.1.7 Peer-reviewed journals with national circulation (1)

[NJ1] Esma Yahia, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Extraction de la structure de la sémantique dans les modèles de système d'information d'entreprises collaboratives. Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information, Lavoisier, 2012, 17 (4), pp.49-77. ⟨10.3166/ISI.17.4.49-77⟩. ⟨hal-00722247⟩

1.3.1.8 Book Chapter (3)

[BS3] Concetta Semeraro, Hervé Panetto, **Mario Lezoche**, Michele Dassisti. Digital twin paradigm for collaborative intelligent manufacturing. ICIM Coallition. 2020 Intelligent Manufacturing Report, In press. ⟨hal-03025641⟩

[BS2] **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Formal Fact-Oriented model transformations for Cooperative Information Systems semantic conceptualisation. Enterprise Information Systems. Lecture Notes in Business Information Processing, 2012, LNBIP 102, pp.117-131. ⟨10.1007/978-3-642-29958-2_8⟩. ⟨hal-00660047⟩

[BS1] **Mario Lezoche**, Taglino Francesco. Business Process Driven Solutions for Innovative Enterprise Information Systems. Information Systems: People, Organizations, Institutions, and Technologies, Physica Verlag Heidelberg, pp.Part 7, 407-414, 2009, ⟨10.1007/978-3-7908-2148-2_47⟩. ⟨hal-00656681⟩

1.3.1.9 Other publication (1)

[OP1] Esma Yahia, **Mario Lezoche**, Alexis Aubry, Hervé Panetto. Extraction de la sémantique dans les modèles de systèmes d'information d'entreprises collaboratifs. Journée Nationale du GT Easy-DIM, Apr 2011, Lyon, France. 2011

1.3.2 Research mentoring activities

Since 2012 I co-supervised 5 doctoral theses. I also supervised two post-doc and 2 research master students. During these eight years I also supervised a PhD student who unfortunately withdrew shortly before the final exam. Currently I'm supervising a PhD student who works at the NIST in USA. Table 5 gives a summary of all the supervision. In the near future I will have two new PhD students to supervise, a CIFRE thesis with the SNMSF company (50% of direction) and a PhD student related to the ANR project ISOBIM that will start in January 2021.

Supervision Type	Total number	Of which ongoing	Total percentage of management
Post-Doc	2	2	150%
PhD Thesis	6	1	204%
Future PhD Thesis	2	2	84%
Master's research	2	0	200%

Table 5 – Mentoring activity

1.3.2.1 List of mentored Post-doc

[PD2] Concetta SEMERARO:

- Title: Cyber Physical Systems metamodeling for Digital Twin application
- July 2020 – July 2021
- Management at 50% - Co-management: Prof. Hervé Panetto (50%)
- Post-Doc financed on the DIH4CPS project.

[PD1] Yasamin ESLAMI:

- Title: Cyber Physical Systems metamodeling definition for knowledge formalisation process
- March 2020 – March 2021
- Management at 100%
- Post-Doc financed on the DIH4CPS project.
- Scientific production: One international journal paper [IJ8].

1.3.2.2 List of mentored (and mentoring) theses

[DT6] Yandé Ndiaye:

- Title: Découverte et formalisation des connaissances pour la fabrication additive par des méthodes d'intelligence artificielle et de fouille de données (Knowledge discovery and formalisation for Additive manufacturing through Artificial Intelligence and Information Retrieval methods).
- 2020 – 2023.
- PhD in co-supervision Université de Lorraine / NIST (USA). Scholarship from the NIST.
- Co-Director at 34%. Directors: Prof. Hervé Panetto (33%) and Prof. Yan Lu (33%).

[DT5] Mickael WAJNBERG:

- Title: Analyse relationnelle de concepts: une méthode polyvalente pour l'extraction de connaissance (Relational concept analysis: a versatile method for knowledge extraction)
- 2017 – 2020.
- PhD in co-supervision Université de Lorraine / Université du Québec à Montréal (Canada). Scholarship from the Canadian government.
- Management at 34%. Co-directors: Prof. Hervé Panetto (33%) and Prof. Alexandre Blondin Massé (33%).
- Scientific production: One international journal paper [IJV3], Three conference papers [C32], [C31] and [C26].

[DT4] Concetta SEMERARO:

- *Thesis Title:* Contribution à la formalisation qui est dirigé par les données d'invariants de modélisation de systèmes cyber physiques (Contribution to the formalization which is driven by the data of modeling invariants of cyber-physical systems)
- *Period:* 2016 – 2020.
- *Grant:* PhD in co-supervision Université de Lorraine / Politecnico di Bari (Italy). Scholarship from the Italian government.
- ***Personal Management:*** **34%**. Co-directors: Prof. Hervé Panetto (33%) and Prof. Michele Dassisti (33%).
- *Scientific production:* One international journal paper [IJV4], Two conference papers [C30] and [C23].
- *Actual Position:* Post-doc at the Université de Lorraine on a European Project contract titled DIH4CPS.

[DT3] Yasamin ESLAMI:

- *Thesis Title:* A Modelling-Based Sustainability Assessment in Manufacturing Organizations
- *Period:* 2015 – 2019.
- *Grant:* PhD in co-supervision Université de Lorraine / Politecnico di Bari (Italy). Scholarship from the Italian government.
- **Personal Management: 34%**. Co-directors: Prof. Hervé Panetto (33%) and Prof. Michele Dassisti (33%).
- *Scientific production:* One international journal paper [IJ6], One conference paper [C28].
- *Actual Position:* Post-doc in the Chamber of Commerce and Industry of the Vosges, on a European Project contract titled DIH4CPS.

[DT2] Silvana PEREIRA DETRO:

- *Thesis Title:* A framework for interoperability assessment in E-Health information systems using process semantics mining
- *Period:* 2015 – 2017.
- *Grant:* PhD in co-supervision Université de Lorraine / PUCPR (Brazil). Scholarship from the Brazilian government.
- ***Personal Management:* 34%**. Co-directors: Prof. Hervé Panetto (33%) and Prof. Eduardo Portela (33%).
- *Scientific production:* Two international journal papers [IJ5], [IJV2], Two conference papers [C25] and [C18].
- *Actual Position:* Associate professor at the Universidade Federal do Paraná

[DT1] Yongxin LIAO:

- *Thesis Title:* Annotations sémantiques pour l'interopérabilité de systèmes dans un contexte PLM (Semantic annotations for system interoperability in a PLM context).
- *Period:* 2010 – 2013.
- *Grant:* Scholarship paid by the Charles Hermite Federation.
- ***Personal Management:* 34%**. Co-directors: Prof. Hervé Panetto (33%) and Prof. Nacer Boujlida (33%).
- *Scientific production:* Three international journal papers [IJ4], [IJ3], [IJV1], Six conferences papers [C15], [C14], [C12], [C11], [C10] and [C7].
- *Actual position:* Associate Professor at Huaqiao University

1.3.2.3 List of future PhD tutorship

[FT2] Selected Student: Nicolas Neutwyler:

- *Thesis Title:* Formal methods for extracting and reusing knowledge from heterogeneous sources for the semantic interoperability of distributed architectures.
- *Expected Period:* February 2021 – February 2023.
- *Grant:* Thesis ANRT CIFRE with the enterprise SNMSF.
- ***Personal Management:* 50% co-director**. Co-directors: Prof. Hervé Panetto.

[FT1] Selected Student: Subject definition:

- *Thesis Title:* Definition of a formal knowledge repository related to the concept of digital twin using formal methods and automatic learning algorithms for the extraction and structuring of knowledge from heterogeneous data sources: application to the ISOBIM project.

- *Expected Period:* February 2021 – February 2023.
- *Grant:* ANR Project ISOBIM.
- ***Personal Management:*** **34% co-director**. Co-directors: Prof. Hind Bril El Haouzi and Prof. William Derigent.

1.3.2.4 List of mentored research master's students

[RM2] Gabriella Anna TEUTONICO:

- Title: Intéroperabilité des systèmes (Systems interoperability)
- 2016
- Management at 100%.

[RM1] Luigi CASCELLA:

- Title: Enterprise Knowledge formalisation
- 2012
- Management at 100%.

1.3.3 Scientific Outreach

1.3.3.1 Participation in faculty selection committees

2016: Selection committee member MCF0644. IUT Épinal – LERMAB. CNU 60 / 62

1.3.3.2 Conference Programme Committee

Since 2014: Member of the International Program committee of the annual International Conference on Information Society and Technology (ICIST).

From 2013 to 2016: Member of the International Program committee of the annual International Conference PRO-VE.

From 2016 to 2017: co-chair of the annual workshop EI2N "Enterprise Integration, Interoperability and Networking. 2 editions

1.3.3.3 Editorial responsibility

2020: Panetto H., Lezoche M., Hernandez J., Alemany Diaz MME, Kacprzyk J. (Editors). Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. Computers in Industry, Elsevier, 2020

1.3.3.4 Organization of symposiums, conferences, study days

2017: Member of the organising committee for the organisation of the 24th GDR MACS STP Days

2015: Co-organizer of a scientific day of the Charles Hermite Federation on the 16th November 2015: "Knowledge management using formal methods for seizure prediction in epilepsy".

2014: Co-organizer of a scientific day of the Charles Hermite Federation on the 18th February 2014: "Knowledge management using formal methods".

1.3.3.5 Proofreading for scientific journals or conferences

Journals and internationals conferences papers reviewing:

- Journal of Intelligent Manufacturing (JIM), Enterprise Information Systems (EIS), International Journal of Production Research (IJPR), Journal of Engineering Manufacture (JEM), Future Generation Computer Systems Elsevier (FGCS), Data & Knowledge Engineering (DATAK), Computer In Industry (CII), Facta Universitatis (FU), Industrial Electronics Society (IES), International Journal of Cooperative Information Systems (IJCIS), IEEE control, IoS Journal, IEEE IoT,

Internationals conferences papers reviewing:

- CAISE'08, OTM'2009, IFAC/IFIP EI2N'2010, ATOP'2010, PAASC'2010, IFIP IWEI'2011, BPM2011, JD Macs 2011, ProVe'2011, OTM'2011, IWEI'2011, MOSIM'2012, BPM'2012, Bustech2012, IWEI2012, COOPIS2012, ODBASE 2012, PAASX2012, IWEI 2013, BPM 2013, PRO-VE 2013, COOPIS 2013, EI2N 2013, EMMSAD 2013, MIM2013, IFACWC 2014, ICIST 2014, Bustech 2014, EMMSAD 2014, BMP 2014, CSDM 2014, Pro-Ve 2014, WCCS14, MOSIM 2014, EI2N 2014, AICCSA 2014, ComSIS 2014, IWEI 2015, EMMSAD 2015, ISI 2015, Pro-Ve 2015, ICT-DM15, EI2N 2015, IFAC WC 2017, Pro)Ve 2017, INCOM 2018, ICIST 2018, Mosim 2018, PLM 2018, Pro-Ve 2018, EI2N 2018, ICIST 2019, MIM 2019, CAISE 2019, JOWO 2019, EI2N 2019, ICPS, PLM 2020

1.3.3.6 Publications with foreign researchers

More than half of my international journal articles are the result of collaboration with fellow researchers from a foreign university (8/15). These universities are from 7 different countries: Brazil, Canada, Italy, Poland, Serbia, Spain and UK.

1.3.4 Participation in working groups and learned societies

Since 2013: Member of the Scientific Interest Group Interoperability-GR of the Greater Region.

Since 2012: Member of the WG Easy-DIM (currently INE) of GDR MACS

Since 2014: Member of IFAC TC 5.3

Since 2020: Member of IFIP TC-12 (WG12.1 and WG12.6)

1.3.5 Participation in research projects

1.3.5.1 International projects

- **STIC AmSud** AGROFAIR "Agro-Knowledge Integration: Developing a FAIR data science approach for adding value to the agricultural supply chain" 2021-2022: 24 months
  Role: Coordinator
  Coordinator: Université de Lorraine (CRAN), FR

- **ANR PRCE** ISOBIM "Proposition d'un processus collaboratif pour la rénovation par ISOlation extérieure basé sur les paradigmes Lean et BIM" 2021-2024: 42 months
  Role: Definition of the roles the 13th of January 2021
  Coordinator: Université de Lorraine (CRAN), FR

- **H2020 IA** DIH4CPS "Fostering DIHs for Embedding Interoperability in Cyber-Physical Systems of European SMEs" 2020-2022: 36 months
  Role: Task 3.2 leader
  Coordinator: UNINOVA, PT

- **H2020 Marie Curie RISE** RUC-APS "Enhancing and implementing Knowledge based ICT solutions within high Risk and Uncertain Conditions for Agriculture Production Systems" 2016-2020: 48 months
  Role: Agriculture Ontology creation leader
  Coordinator: University of Liverpool, UK

- **H2020 COST** TD1406 I2MHB "Innovation in Intelligent Management of Heritage Buildings" 2014-2019.
  Role: Researcher
  Coordinator: Universidade Nova de Lisboã, Portugal

- **PICS SICIM** with University of Nis (Serbia) "Semantic interoperability approach to Computer Integrated Manufacturing" (24 months 2012-2013)
  Role: Researcher
  Coordinator: Université de Lorraine

- **PICS IA4PSS** with Politecnico di Bari, Italy "Interoperability assessment for production systems sustainability" (36 months 2014-2016)
  Role: Researcher
  Coordinator: Université de Lorraine

- **Pavle Savic PHC** Project with the University of Nis (Serbia): "Runtime Model-Driven Software for Smart Cyber Physical Systems. (2016-2017)
  Role: Researcher
  Coordinator: Université de Lorraine

- **Project "Sciences sans frontières"** with PUCPR (Brazil) "Model and Formal Approach on Interoperability Engineering applied to the Systems Life Cycle". (2013-2015)
  Role: Researcher
  Coordinator: Université de Lorraine

- **CNRS PICS** project with the University of Nis (Serbia). "Semantic interoperability approach for Computer Integrated Manufacturing". (2013-2014)
  Role: Researcher
  Coordinator: Université de Lorraine

- PHC Galileo project with Politecnico di Bari (Italy). (2011-2012)
  Role: Researcher
  Coordinator: Université de Lorraine

1.3.5.2 Submitted International projects

- **ANR PRCI** KERMIT "Knowledge extraction and machine learning algorithms to improve industrial digital assistants" 2021-2024: 36 months
  Role: Coordinator
  Submission deadline: End of April 2021
  Coordinator: Université de Lorraine (CRAN), FR

- **ANR PRCI** AI4C2PS "Artificial Intelligence for Cognitive Cyber-Physical Systems interoperability" 2021-2024: 36 months
  Role: Participating in the writing of the detailed proposal
  Submission deadline: End of April 2021
  Coordinator: Université de Lorraine (CRAN), FR

1.3.6 Collective Responsibilities

Since 2017: Co-Head of the project team Intelligent System and Objects in Interaction (S&O-2I) with the Associate Professor William DERIGENT and Associate Professor Philippe THOMAS.
The Project team is a sub entity of the Eco-Technic systems engineering (ISET) department of the Research Center for Automatic Control (CRAN) laboratory.

2 Synthesis of research work 2012 - 2020: Semantic interoperability through the knowledge formalization in enterprises IT systems. Industry 4.0 Context.

In this chapter I will present my research work whose theme can be summarized as how the formalisation of knowledge can improve the semantic interoperability of the information systems of companies producing services and goods in Industry 4.0 context. The Figure 7 show the contributions that come mainly from the theses that I supervised [DT1] of Yongxing LIAO, [DT2] of Silvana PEREIRA DETRO, [DT3] Yasamin ESLAMI, [DT4] of Concetta SEMERARO and [DT5] of Mickael WAJNBERG.

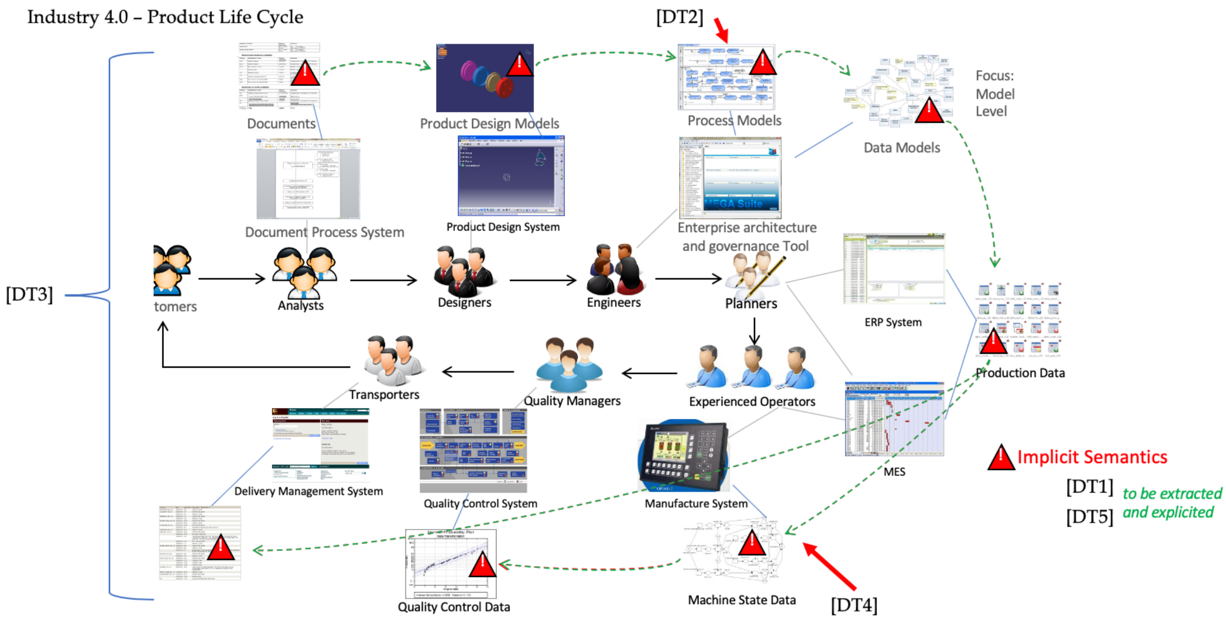


Figure 7 Contribution of supervised PhD theses in the issue of formalization of knowledge in an Industry 4.0 context (this image is adapted from DT1 image)

Yongxin Liao's thesis [DT1] was interested in the problem of semantic annotations in a product life cycle context.

Silvana Pereira Detro's thesis [DT2] contributed to solving the problem of semantic interoperability through the study of structuring semantic processes.

Yasamin Eslami's thesis [DT3] focused its contribution on formalising the solution to assure a sustainability assessment in the manufacturing Organizations.

Concetta Semeraro's thesis [DT4] attacked the problem of the semantic definition invariants of cyber-physical systems.

Eventually Mickael Wajnberg's thesis [DT5] focused on knowledge extraction through Multi Relational Data Mining (MRDM) methods and specially focusing on the Relational Concept Analysis (RCA).

2.1 Yongxin Liao's thesis: Semantic annotations for system interoperability in a PLM context.

Yongxin Liao's thesis relates to deal with the semantic interoperability by proposing a formal semantic annotation method to support the mutual understanding of the semantics inside the shared and exchanged information in a Product Life Cycle (PLM) environment.

The concept of PLC has been revealed for more than sixty years [CIMdata, 2003]. It describes every stage of a product of interest (such as imagining, defining, realising, using/supporting and retiring/disposing of). In the meantime, along with the advent and the evolution of Computer Aided Design (CAD) systems, the problems of locating the required data and losing control of change process associated with these data have gradually appeared [CIMdata, 2003]. As a solution, Product Data Management (PDM) systems have been developed and introduced for supporting easy, quick and secured access to valid data during the product design phase [Ameri, 2005]. However, as it is pointed out in [Elgueder, 2010], the data produced by CAD systems do not cover all the information that related to the whole product life cycle (from the requirement specification to dismantling information). The PDM systems are not able to give enough support for non-engineering data. In order to fill this gap, during the 1990s, the PLM solution is proposed to support the processes of capturing, representing, retrieving and reusing both engineering and non-engineering aspects of knowledge along the entire product life cycle. It intends to facilitate the knowledge management in or across enterprises [Ameri, 2005]. Therefore, the knowledge concerning the product life cycle, which we named PLC-related knowledge, has become one of the critical concepts in a PLM solution.

Knowledge is an awareness of things that brings to its owner the capability of grasping the meaning from the information [Ackoff, 1989]. This definition is included in a more structured presentation of the DIKW Pyramid [Zeleny, 1980], a hierarchical model for representing the structural relationships between Data, Information, Knowledge and Wisdom (DIKW). In this research work, we consider knowledge as a kind of intangible thing that is only explicit to its owner but remains tacit to the external world [Leibniz, 1989]. One of the main purposes of knowledge management is to make knowledge accessible and reusable [O'Leary, 1998]. Knowledge Representation is the result of embodying the knowledge from its owner's mind into some explicit forms. It gives a possibility for external entities to perform some specific operations for achieving their particular needs. Knowledge representations act as the carriers of knowledge to assist collaboration activities.

Interoperability serves as a foundational role to support collaboration. It is the ability that diverse entities can exchange knowledge representations and make use of those knowledge representations that they have exchanged. Five possible levels of interoperability have been categorized by Euzenat [Euzenat, 2001]: encoding level, lexical level, syntactic level, semantic level and semiotic level. While encoding, lexical and syntactic issues are now can be formally solved by many technical standards, enabling a seamless semantic interoperability remains a huge challenge [Panetto, 2007]. In order to cope with the semantic interoperability issue, two important obstacles still need to be overcame:

(1) The implicit semantics that is necessary for understanding a knowledge representation that is not made explicit.

(2) The lack of mechanisms to verify the correctness of explicit semantics in the exchanged knowledge representation.

A mutual understanding of the semantics inside the shared and exchanged knowledge representations is the cornerstone in the quest for semantic interoperability. Due to the essence of ontology, which is a kind of common agreement on the conceptualization of terms in a specific domain of interest, semantic annotations [Boudjlida, 2008] are usually considered as a possible solution to deal with these two obstacles.

In this context, semantic enrichment is considered as a process that makes any implicit semantics more explicit through the use of semantic annotations. Some research questions are then emerging from the needs of semantic enrichment in a PLM environment:

(1) What are the semantic interoperability problems that exist during the cooperation in a PLM environment?

(2) What kinds of knowledge representation in a PLM environment need semantic enrichment?

(3) What kinds of ontology can be used to support the semantic enrichments of those knowledge representations?

(4) What are the essential elements of a semantic annotation and how to formally represent a semantic annotation in a suitable format?

(5) How to semantically enrich a knowledge representation and how can these enriched semantics contribute to the semantic interoperability in a PLM environment?

2.1.1 The proposed solution

Compared with the other types of annotations, a semantic annotation has two important features: (1) it can be read and processed by a machine [Boudjlida, 2008]; (2) it contains a set of formal and shared terms in a specific context [Oren, 2006]. As a way to realize the semantic enrichment, semantic annotations use ontology to capture annotator's knowledge and then act as a knowledge carrier to enrich annotated object's semantics. It can then be widely used in many contexts for various purposes. In this work, there are two important aspects of the semantics that are made explicit by a semantic annotation:

(1) The domain semantics, which describes the context and the meaning of an annotated object in a specific domain.

(2) The structure semantics, which describes the interrelations between the annotated objects and the other objects related to them in a knowledge representation.

Based on the survey and exploration of current semantic annotation research, three shortcomings have been identified:

(1) The formalization of semantic annotations is not the focus of most of the semantic annotation research. They only considered the semantic annotation as a kind of "is a" association between one annotated object and one ontology concept. Even if there are some specific formalizations, they are difficult to be reused in other research but the studied ones.

(2) The domain semantics of the annotated objects is the only concern in most of the semantic annotation research, where the structure semantics is ignored, or vice-versa.

(3) In most of the semantic annotation research, there is lack of mechanism to support the inconsistency detection of the semantic annotations and the conflict identification between the annotated objects. [IJ4]

Therefore, we focus our research work on: (a) clearly identifying the essential elements of a semantic annotation by proposing a formalization that can be used to enrich different types of models; (b) proposing two mechanisms to detect the possible inconsistencies of semantic annotations and to identify possible conflicts between the annotated objects for facilitating and assisting the knowledge management in a PLM environment.

Based on this research focus, in this thesis, a semantic annotation is used as a way to employ one or several ontologies for making explicit both structure and domain semantics of an annotated target knowledge that needs to be made explicit. For this reason, a semantic annotation is considered as semantic relationships between the Target Knowledge Representations (TKRs) and the Ontology-based Knowledge Representations (OKRs). These relationships are formally defined in a Semantic Annotation Structure Model (SASM).

TKRs are the targets of semantic enrichment, namely the targets that semantic annotations are attached to. They contain implicit or possibly ambiguous explicit semantics, which is not easily intelligible. In a PLM environment, all the different types of models throughout the entire PLC are considered as embedding elements of TKRs (different kinds of modelling constructs represent different kinds of knowledge): such as data models, process models, state models, resource models, decision models.

OKRs are the ontologies that capture different aspects of knowledge and provide the common and shared conceptualizations for supporting the semantic enrichment of TKRs. In this research work, two types of ontologies are used: (1) The PLC-related ontologies that formalize the domain semantics of TKRs. They are normally collected by domain specialists and formalized by knowledge engineers. (2) The meta-model ontologies that formalize the structure semantics of TKRs. They are normally generated by modelling language experts. [C14]

The SASM contains a set of definitions that formalize the structure of a semantic annotation. It acts as a bridge to formally describe the semantic relationships between the TKRs and the OKRs. It identifies the essential elements of the semantic annotation, which can be used as the basis for developing the common semantic annotation schema for different kinds of models along the product life cycle. Differently from other semantic

annotation proposals, we propose to annotate an object in a TKR with a set of selected ontology elements. These elements belong to one or more OKRs, which describe the domain and structure semantics of the annotated objects.
Moreover, in order to detect the inconsistencies between the semantic annotations and to identify the conflicts between the annotated objects, we proposed a semantic annotation suggestion mechanism and two verification mechanisms that rely on the SASM. The former verification is based on the domain semantics comparisons of two or more related semantic annotations of a common annotated object. The later verification is based on the results of the former verification. Both structure and domain semantics, which are made explicit by semantic annotations, contribute to these three mechanisms. The domain semantics acts as the data that are used for similarity comparison. The structure semantics supports the creation of reasoning rules that are used for the inference.
Finally, in order to apply the proposed solution into a PLM environment, a semantic annotation framework, which is composed of a general semantic annotation procedure and the architecture of the framework, are proposed.

2.1.2 The contribution

This thesis presents a formal semantic annotation method that supports the semantic enrichment of models in a PLM environment for facilitating semantic interoperability issues. The contributions have been published in four papers ([C7], [C10], [C11] and [C12]), which are listed as follows:
(1) We surveyed the literature that relates to the Product Lifecycle Management, System Interoperability and Semantic Annotation. A comparison of some current semantic annotation researches is made for identifying the exiting drawbacks and potential challenges ([C7], [C10], [C11] and [C12]).
(2) We proposed a formalization of semantic annotations that can be used as a basis to create a semantic annotation schema for supporting the semantic enrichment of models ([C7], [C10], [C11] and [C12]).
(3) We proposed a semantic annotation suggestion mechanism and two verification mechanisms to support inconsistency detection between semantic annotations and conflict identification in model contents ([C12]).
(4) We proposed a guideline that contains a number of procedures to guide an engineer in applying the proposed solution and a semantic annotation framework for enriching semantics of models in a PLM environment ([C11] and [C12]).
(5) We designed, implemented and validated a prototype annotation tool for the semantic annotation of process models based on the proposed formalization, annotation procedures and framework ([C12]).

2.1.3 Formalization of semantic annotation

2.1.3.1 *Meta Model of the Semantic Annotation*

The comprehension of the knowledge that is represented by a model needs not only the domain semantics that is embedded in the contents of the model, but also the structure semantics that is embedded through the modelling constructs. Therefore, the relevant semantics is supposed to be contained in the employed ontologies. The ideal situation is that there exists equivalent semantics in ontologies for every annotated element in a model. However, because of the complexity of the reality, in most of the cases this situation rarely appears. Therefore, a more reasonable relation definition is required for describing the semantic relationships between an element of a target knowledge representation and its corresponding domain and structure semantics. To be more specific, in order to define the meta-model of the semantic annotation, based on the research context and the investigation of existing researches in the previous chapters, several important concepts that are used throughout this chapter need to be reviewed.

2.1.3.2 *Target Knowledge Representation (TKR)*

Models in a PLM environment act as an important role to enable the capturing and representation of the relevant knowledge from each product life cycle stage. These models are always expressed in some kinds of modelling languages or notations with designer's specific peculiarities, such as, different backgrounds, unique knowledge, heterogeneous expertise, particular needs and specific practices. This results in the implicit, or possibly ambiguously explicit, semantics that is not easily intelligible by the humans or the machines. The interoperation process between enterprise systems and stakeholders not only requires that these models can be exchanged and operated on, but also demands the unambiguous understandings of the exchanged models. In this research work, all different kinds of models throughout the product life cycle are considered as Target Knowledge Representations (TKRs).

Ontology-based Knowledge Representation (OKR)

Ontology represents a real-world semantics that enables human to use meaningful terminologies as machine processable contents. It formalizes the common and shared understandings in a human and machine interpretable way [C14], which is frequently chosen as a candidate procedure to formalize knowledge. According to the research context two kinds of ontologies (PLC-related ontologies and Meta-model ontologies) are used to support the semantic enrichment of models in a PLM environment. These ontologies are considered as Ontology-based Knowledge Representations (OKRs) in this research work.

2.1.3.3 *Semantic Annotation Structure Model (SASM)*

The Semantic Annotation is acting as a bridge to formally describe the semantic relationships between TKRs and OKRs. Two aspects of semantics are made explicit through a semantic annotation:

(1) The domain semantics that describes the context and the meaning of an annotated object in a certain domain.
(2) The structure semantics that describes the interrelations of the annotated object and the other related objects in the TKR.
The meta-model of the semantic annotation is presented in the Figure 3-1, which describes the main components of a semantic annotation and their relationships.

- A "Target Knowledge Representation" is the composition of one or more "Element of a TKR".
- The "Ontology-based Knowledge Representation" is the generalization of the "Meta-model Ontology" and the "PLC-related Ontology".
- A "Meta-model Ontology" is the composition of one or more "Element of a Meta-model Ontology".
- A "PLC-related Ontology" is the composition of one or more "Element of a PLC-related Ontology".
- An "Element of a TKR" can be annotated by zero or more "Semantic Annotation".
- A "Semantic Annotation" contains one "Structure Semantics".
- A "Semantic Annotation" contains zero or more "Domain Semantics".
- A "Structure Semantics" is the aggregation of one "Element of a Meta-model Ontology"
- A "Domain Semantics" is the aggregation of one or more "Element of a PLC-related Ontology".

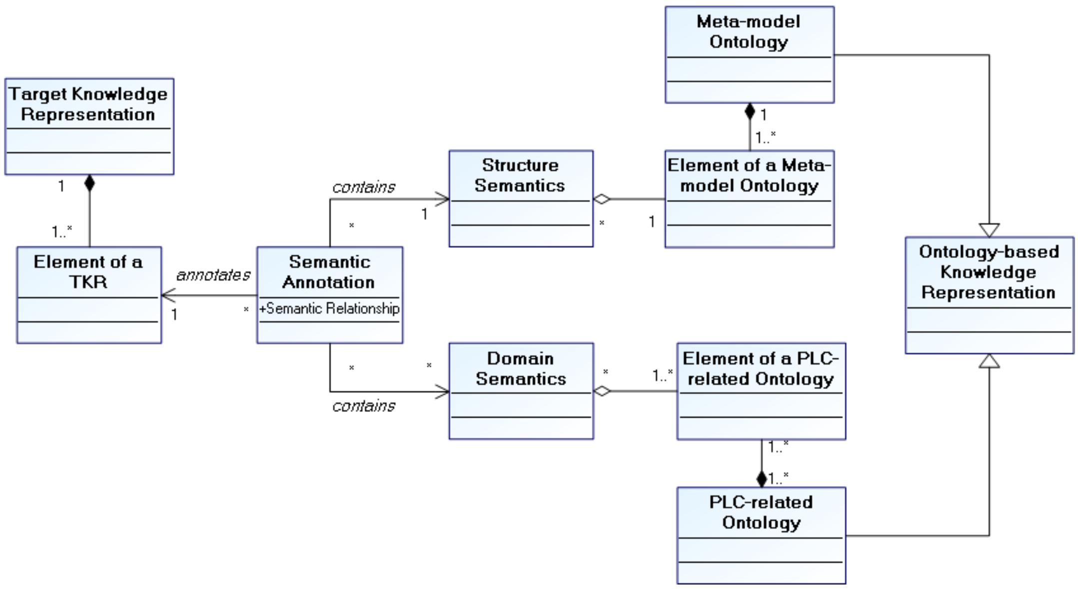


Figure 8 – Semantic annotation metamodel from [C14]

Based on this meta-model, in the next section, we propose a semantic block delimitation method. This method will be used as a basis to support the proposition of formal definitions and the creation of reasoning rules.

2.1.4 The Semantic Annotation Framework

In this section, the semantic annotation framework for capturing, representing and managing the knowledge related to the system of interest through the semantic annotation is presented.

2.1.4.1 Semantic Annotation Procedure

Taking advantages from the formal definitions and the three main mechanisms presented in previous sections, as shown in the *Figure 9*, a general overview of the procedures for applying the semantic annotations is presented. This workflow is divided into three main phases: The Preparation Phase, The Annotation Phase and The Reasoning Phase.

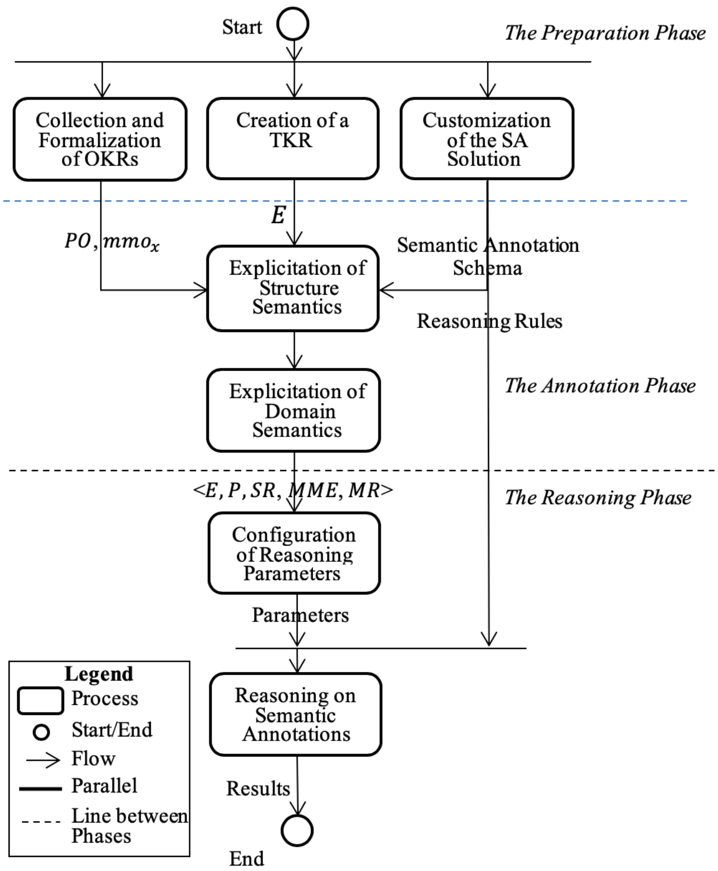


Figure 9 – Semantic annotation procedure [C14]

The Preparation Phase: during this phase, all the elements that are needed by both the annotation phase and the reasoning phase are prepared.

Creation of a TKR, in which, a model, namely a Target Knowledge Representation (TKR), is created by a modelling system. The set of elements E in this TKR are the output of this process.

Collection and Formalization of OKRs, in which, the ontologies, namely Ontology-based Knowledge Representations (OKRs), for making explicit the domain semantics and structure semantics are captured and formalized. The output of this process is a number of PLC-related ontologies (PO) and a meta-model ontology. The selection of an ontology can be based on some ontology evaluation methods [Sabou, 2006] and [Vrandevcić, 2009].

Customization of the SA Solution, in which, the formal definitions of semantic annotations and the reasoning mechanisms are used as the basis for customizing a semantic annotation model and the reasoning rules. The output of this process is divided into two parts: the semantic annotation schema that can be used as a repository to conserve the objects of semantic annotation (E), contents of semantic annotation (P and MME) and the semantic relationships (SR and MR) between them; the Reasoning Rules that can be used to support both delimitation of semantic block and the inference process in the reasoning phase, such as annotation suggestion rules, inconsistency detection rules, conflict identification rules and so on.

The Annotation Phase: during this phase, a number of semantic annotations are generated for supporting the reasoning phase.

Explicitation of Structure Semantics, in which, the structure semantics of a TKR, namely the interrelations between the model elements, are made explicit. The customized E and MME are used to keep selected the elements in the model and in the meta-model ontology respectively. The binary relations in MR are used to define the semantic relationships between elements in E and elements in MME.

Explicitation of Domain Semantics, in which, the domain semantics of a TKR, namely the meaning of model contents in a domain of interest, are made explicit. The customized P is used to keep the selected ontology elements in the PLC-related ontologies PO. The binary relations in the SR are used to define the semantic relationships between the elements in the E and the elements in P.

Reasoning Phase: during this phase, the reasoning is performed based on the outputs of the above-mentioned two phases.

Configuration of Reasoning Parameters, in which, based on the customization of semantic annotation schema and the practical situation for different TKRs, corresponding operations that support the configuration of reasoning parameters are performed by annotators or machines, such as the delimitation of semantic block for semantics substitution, the determination of the equivalents of two properties, the comparison of the semantic similarity between two domain semantics of common annotated objects, just to name three possible processes.

Reasoning on Semantic Annotations, in which, the reasoning is performed based on the semantic annotations, the parameters and the reasoning rules to produce inference results. For example, the results of annotation suggestions, annotation inconsistency detection, model conflict identification and so on.

This workflow describes the application of the semantic enrichment solution within one single TKR. In order to deal with the multiple TKRs in a PLM environment we propose a semantic annotation framework to address the issues of semantic interoperability. [IJ4]

2.1.4.2 The Framework Architecture

As shown in the *Figure 10*, on the left side, there is a series of processes to describe a product life cycle. They represent the TKR Creation and Management module. On the right side, there are four main modules of this framework: the OKR Creation and Management module, the Knowledge Cloud module the Semantic Annotation and Processing Agent (SAPA) module and the Reasoning Engine module.

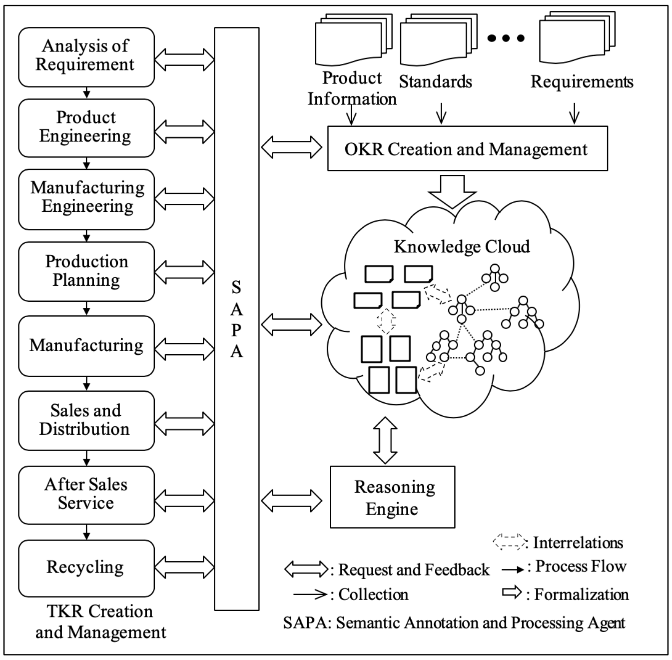


Figure 10 – Semantic annotation framework architecture [DT1]

The TKR Creation and Management module is composed of a number of enterprise systems. Stakeholders in or across enterprises, during a product lifecycle use those systems to create TKRs following the modelling specifications and to manage those TKRs. Those systems need to provide sufficient APIs to enable the communications between themselves and the Semantic Annotation and Processing Agent module.

The OKR Creation and Management module is in charge of capturing, formalizing and managing PLC-related knowledge and model constructs knowledge into a knowledge base, namely, Knowledge Cloud. The OKRs are supposed to be in a platform-independent form, which ensures different kinds of ontologies that are collected from different sources, to be imported, mapped, merged and interrelated with each other.

The Knowledge Cloud module acts as a knowledge repository, which is in charge of storing different kinds of knowledge. As shown in the Figure 11, three main kinds of knowledge are stored in the knowledge cloud:

(1) All the OKRs produced by the OKR Creation and Management module. As shown in the Figure 11 (a), the OKRs are structured as the traditional three-levels structures.

a) The top-level ontology. It contains common terms and specifies the most common terminology that can be used in different domains.

b) The domain level ontology. It is classified into two aspects:
i. The PLC-related ontologies that represent the knowledge related to the product life cycle from different perspectives.
ii. The Meta-model ontologies that represent the knowledge related to model constructs based on different specifications.
c) The application-level ontology. Corresponding to the two aspects of ontologies in the domain level, the related ontologies in this level are responsible for representing the specific terms that are defined and used in an enterprise and for representing the specific implementation of meta-model concepts in different modelling tools respectively.
(2) All the semantic annotations that are created by different stakeholders along the product lifecycle via the Semantic Annotation and Processing Agent module. As shown in the Figure 11 (b), these semantic annotations define the semantic relationships between TKRs and their corresponding OKRs. They also can be used as the bridges to make explicit the interrelations between the annotated elements in the disperse TKRs.
(3) All the reasoning rules, as shown in the Figure 11 (c), which are created based on the concepts and relationships in the OKRs and the customized semantic annotation schema. They are used for supporting the inference in the Reasoning Engine module.

Reasoning Engine module.

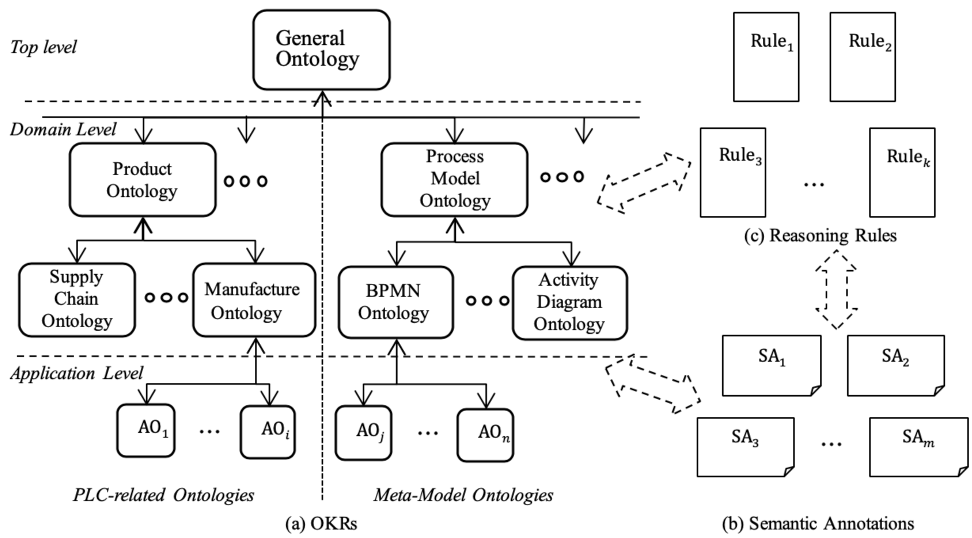


Figure 11 – Reasoning engine module 11 [DT1]

The Reasoning Engine module is an external call pattern-matching search engine, which uses different reasoning algorithms according to the stakeholders' requests. It performs the inferences on the knowledge that is stored in the semantic annotations, in the OKRs and in the reasoning rules. [IJ3]
The Semantic Annotation and Processing Agent (SAPA) model is mainly in charge of the semantic relationships definition process. It also acts as a mediator to support the communications (requests and feedbacks) between various kinds of modelling systems in

different processes of the PLC (TKR Creation and Management module) and the three other modules (Knowledge Cloud module, OKR Creation and Management module and Reasoning Engine module) in the semantic annotation framework:

(1) Between the Knowledge Cloud module and the modelling systems: according to the particular semantic annotation requests from the stakeholders. The SAPA is in charge of querying (with the assistant of the Reasoning Engine) the Knowledge Cloud and provides appropriate OKRs as feedbacks. It also takes care of the management (such as creating, modifying, loading, deleting and so on) of existing semantic annotations and reasoning rules.

(2) Between the OKR Creation and Management module and the modelling system: based on the requests from stakeholders. SAPA is supposed to be able to communicate with OKR Creation and Management module for the manipulation of the OKRs.

(3) Between the Reasoning Engine and the modelling Systems: SAPA submits the inference requests from the stakeholders to the Reasoning Engine for performing the reasoning actions (such as the suggestion of semantic annotations, the detection of annotation inconsistencies, the identification of possible model conflicts and so on) and it sends back the corresponding results to the stakeholders.

This semantic annotation framework makes use of semantic annotations as a bridge between TKRs and their corresponding OKRs: 1) to make explicit the implicit semantics of TKRs; 2) to give a possibility to detect the inconsistencies between semantic annotations and identify the possible conflicts among the annotated elements in TKRs; 3) to make explicit the implicit relationships among all disperse TKRs.

2.2 Silvana Pereira Detro's thesis: A framework for interoperability assessment in E-Health information systems using process semantics mining

Silvana's thesis focused on developing an approach for process model customization, which provides a decision-making support for the user and enable to individualize a process model that respects the user's requirements and the internal and external regulations.

by:

- Applying process mining techniques to build a customizable process model, enabling to identify the process variants and the rules for selecting them;
- Formalizing the relevant knowledge about the business context in ontologies, and through semantic reasoning provide support for process model customization.

2.2.1 Research Questions

In order to achieve the research goal, this research can be divided in two parts. The first part aims to understand the aspects related with the customizable process model, process

mining, ontologies and the relations between them. Based on the results obtained by analysing the aspects in the first part, the aim of the second part is defined.
Based on the objectives from the first part the research questions can be decomposed into sub-questions. Thus, answering these sub-questions lead to answer the related research questions and consequently to achieve the research goal:

- RQ1: How to customize a process model in order to obtain a process variant that correctly represent a business context?
- RQ1.1: What are the aspects that need to be considered when building a customizable process model?
- RQ1.2: How the existing approaches proposes to customize a process model?
- RQ2: What are the theoretical and practical arguments motivating the application of process mining to discover customizable process models?
- RQ2.1: What are the process mining techniques that can be applied to identify the aspects related with the process variants enabling to improve the customizable process model?
- RQ2.2: How to improve the process model from which each process variant is individualized to consider scenarios that are not available in the event log?
- RQ3: What are the theoretical and practical arguments motivating the use of ontologies for process model customization?
- RQ3.1: Can ontologies be applied to provide decision making support during the process model customization? [IJV2]

Based on the results of the first research part, the second part was developed. We point out that a different result of the first part would lead to a different set of research questions and, thus a set of methods to solve the problem. The first part showed that process mining techniques and ontologies can improve the process model customization. This result led to the development of an approach to customize a process model. Based on this, the research question of the second part can be formulate:

- RQ4: How process mining and ontologies can be applied to customize a process model according to all the requirements related to a particular business context?

2.2.2 Framework for configuring process variants through process mining and semantic reasoning

The main requirements to develop a customizable process model is to identify the parts of the model that are common to all process variants, the parts that can change (i.e., the parts that are subject to variation), how they change and the reasons for changing.
These aspects can be identified through process mining which can provide knowledge about how the process is performed, thus helping on making appropriate decisions to improve it. However, despite the benefits that the event log analysis can provide, many enterprises do not use such data appropriately.
Obtaining the configurable process model by means of the event log enables us to improve the process variants by correcting deviations, if they exist, anticipating problems,

discovering if the requirements have been followed, etc. Besides, the implicit knowledge can be captured and made explicit, thus enabling to enrich the process variants.

In each variation point, a decision needs to be made in order to obtain a process variant. Thus, providing guidance and recommendations for the user during the customization ensure that the resulting process variant is correctly customized. The process variants also need to be correctly in a behavioural and structural way.

The recommendations provided by the existing approaches are limited to recommendations about the variation points. However, recommendations can also be related with the application business context, which can improve the customization. Thus, the recommendations may include the information about the variation points (alternatives and rules) and the information about the business context, including internal and external regulations. Considering this need, ontologies can be applied to support decision making during the process model customization.

The framework proposed in this research, intends to provide a decision-making support during the process model customization. The framework focuses in the aspects such as internal and/or external regulations and expert knowledge to provide recommendations about the business context during the process model individualization. Besides, the knowledge about the actual business process executions is captured in order to improve the process model. Figure 12 depicts the proposed framework.

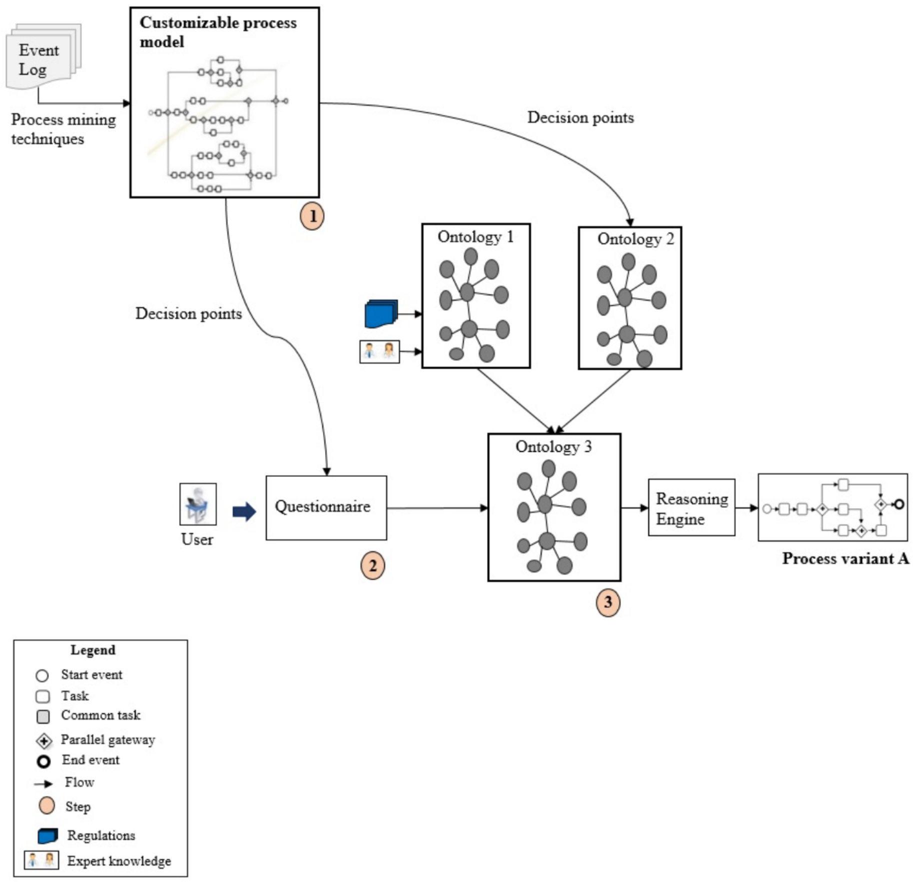


Figure 12 Framework for customize process variants from [IJ5]

2.2.3 Synthesis

The framework proposed in this research aims to support the decision-making during the process model customization by providing recommendations about the business context and the activities in the process model. The recommendations are based on the knowledge obtained from internal and/or external regulations, expert knowledge and the knowledge captured from the process model executions. The framework also provides guidance by means of a questionnaire.

The framework is composed by three steps. In the first step, an event log is analysed by means of process mining techniques. As the event log can be incomplete or contain noise, an approach is proposed to develop a process model based on the event log, the expert knowledge and the internal and external regulations.

In order to analyse the various scenarios that can be extracted from the developed process model, an event log is created to simulate these scenarios. By applying classical process mining techniques, such as α-algorithm or the heuristic miner, the process model is obtained from the generated event log. These approaches enable to identify the variation points; however, they do not provide any information about the rules for choosing the alternatives for each variation point. To discover the data dependencies that affect the routing of a case is applied the decision mining analysis, a process mining technique. In this step, the decision tree concept is used to carry out a decision point analysis, i.e., to find out which properties of a case might lead to taking certain paths in the process.

Based on the decision trees obtained through the decision mining analysis, the questionnaire-model approach is developed in the Step 2. The questionnaire is applied to guide users in providing the information needed for process variant selection. In this approach, each variation point refers to a question, thus the selection of an alternative for a question refers to the selection of the paths available in relation to the respective variation point.

Variations point can be defined as optional or mandatory. The variation points defined as mandatory are related to the selection of process variants, thus they inherit the optional variation points. This definition of the variation points enables to understand the interdependencies between the variation points. In this way, when a selection is made in the mandatory variation points, the related set of optional variation points is enabled. This knowledge is helpful during the definition of the order of dependence on facts and questions in the questionnaire-model approach.

Step 3 refers to the development of the ontologies for process model customization. One ontology formalizes the knowledge related with the variation points. This ontology is developed based on the decision tree obtained through the decision mining analysis. The leaf nodes are defined as concepts in the ontology, which correspond with the alternatives for the variation points. The branches are defined as data properties in the ontology and they correspond with the facts in the questionnaire.

Other ontology formalizes the knowledge about the internal and/or external regulations and expert knowledge. Both ontologies are merged into one ontology. Thus, the resulting ontology contain the knowledge about the business context and the process model. SWRL rules define the relations between the various elements in the ontology. Thus, when a fact is selected, the corresponding data property is enabled, then by reasoning on the ontology, the alternatives related with the business process and the recommendations about the business context are provided for the user.

2.3 Yasamin Eslami's thesis: A Modelling-Based Sustainability Assessment in Manufacturing Organizations

The Eslami's thesis is devoted to a thorough research on introducing a framework that covers a gap between the needs of manufacturing companies to improve their performance in terms of sustainability and the efficiency and capability of the available assessment tools. In addition, literature still lacks a framework that can evaluate sustainable manufacturing as a whole. On the other hand, the lack of systematic view and standardization in the existing assessment methods make them ad hoc and also not capable in recognizing the opportunities to have a sustainable organization. To do so, a 6-step research through the literature was conducted as shown in Figure 13.

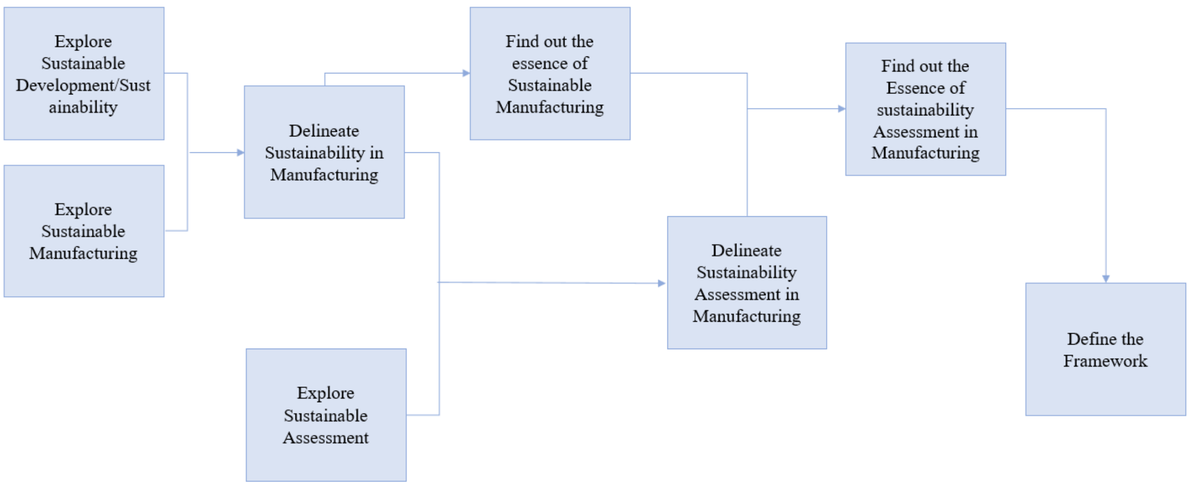


Figure 13 - stream of the logic and the main tasks for the study [DT3]

Based on the figure, prior studies has been done in order to get deep in the concept of sustainability assessment in manufacturing; to serve the purpose the first two research questions were emerged as "How sustainability is defined through its dimensions? and What sub-dimensions can denominate sustainable manufacturing?". The first part of the work is entirely dedicated to finding a respond to these questions. A systematic literature review was conducted on the literature available for the concept of sustainability and sustainable manufacturing to highlight the aspects of sustainability and sustainable manufacturing and scrutinize its dimensions and sub-dimensions. Finally, the observations are analysed through the Formal Concept Analysis approach and the results helped to take one step further in developing the model.

Moving from sustainable manufacturing to sustainability assessment, the arising two other questions were: “How can sustainable manufacturing be achieved” and “How can sustainable manufacturing be assessed?” to find a proper answer for them, an exploration of sustainability assessment was directed on the concepts of sustainability assessments and its tools which resulted in characterization of sustainability assessment in manufacturing and discovering its essence as a consequence; which itself could be followed by proposing the framework.
So, the first part of the thesis delineated the trend toward sustainable manufacturing and sustainability assessment in the scientific domain. Therefore, prior to development of the model, it has been decided to have an analysis of sustainable manufacturing in the manufacturing domain in practice. However, organizations perception of operational sustainability can reveal their strategies on how to be a sustainable organization, endeavouring the three pillars of economics, environmental and social internal assets. The study is centred on the investigation on the role of indicators’ choice and their meaning for the purpose of the sustainability assessment of manufacturing organizations. To this point, an analysis has been conducted on sustainability assessment of 100 manufacturing organizations using GRI indicators for assessing their sustainability state. An FCA study was run to look over the indicators and their interpretations to reach a given degree of sustainability of the organization.

The environmental dimension was the one studied the most alone and alongside others. As the observation showed, dealing with even one sub-dimension from the environmental dimension, was considered as sustainability among manufacturers who practice sustainable manufacturing. Therefore, it was decided to deepen into the dimension and its sub-dimensions while they have been considered for reaching sustainability. Hence, the FCA study scrutinized the usage of the sub-dimensions and to discover the hidden relations between them. The FCA method resulted in two sets of output data: The first set gave a hierarchical relationship of all the established concepts in the form of the concept lattice, while the second one gave a list of all found interdependencies among attributes in the formal context. As mentioned above, FCA helped to display the links between the environmental sub-dimensions through the definition of attributes. Therefore, it was possible to see the combination of the sub-dimensions and their regularity of appearance in the literature. The knowledge extracted and interpreted from the FCA result made evident that three sub-dimensions of “energy”, “waste” and “emission” were the ones used the most alone and alongside the other sub-dimensions.

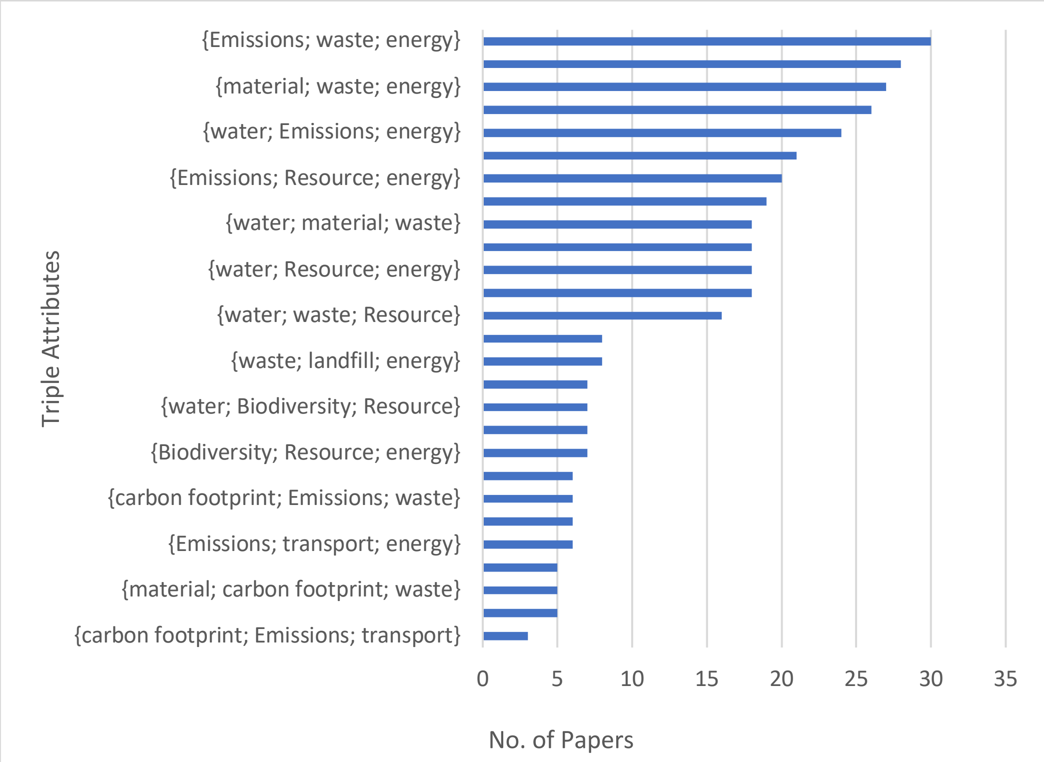


Figure 14 - Triple Combination of Environmental Sub-dimensions [DT3]

This part of the work focused on a systematic literature review on sustainability dimensions and sub-dimensions in order to extract knowledge for manufacturing organizations who want to practice strategies to be more "sustainable" to stay competitive in the market today and also be responsive to the demand of both customers and the government for sustainable products and preservation of natural resources. The main question risen here is to find out "How sustainability is defined through its dimensions? and What sub-dimensions can denominate sustainable manufacturing? Going through the dimensions of sustainability in manufacturing, it was observed that among social, technological, economic, environmental, technology, efficiency and performance management, the traditional three namely: Economic, Environmental and social, also known as the Triple Bottom Line (TBL), were the ones with the most concentration on.

Noticing the result obtained, the focal point of the final part of the work was the final research question: "how we can help manufacturing organizations in terms of assessing sustainability" and "how we can help manufacturing organizations discover opportunities to reach a better state of sustainability". To be able to respond to the question we developed a model-based sustainability assessment tools based on indicators. In addition, the process of aggregating indicators to create a composite sustainability development index has been fully scrutinized.

2.3.1 Model Representation

The reference architecture is aiming at grouping highly divers aspects in a common model to assess sustainability in a holistic way. The special characteristics of the model are therefore its combination of functional level inside a manufacturing organization with the life cycle of the product for the three main dimensions of sustainability. On the other hand, the systematic approach considered to be taken for developing the reference model, permits the maximum traceability of the causes and the effects of sustainability in the whole organization. By means of the present model, the conditions have been created for description, implementation and assessment of the sustainability concept in different dimensions. It prepares a definition for sustainability for each intersection of the three domains with stipulation and requirements. The model enables the manufacturers: to detect a sustainability prevention cause; to know to what functional level it belongs; to discover in which stage of the product life cycle it occurs and to know if the specific problem comes from environmental, social or economic issues.

Based on the abovementioned, three pillars of sustainability must be assessed in all levels of a manufacturing organization throughout the whole life cycle of the product. To make that assessment possible, the three-dimensional model is proposed as shown in Figure 15 to develop and to cover the gap that exists in the literature which is the lack of a model based and a holistic assessment for the manufacturing organizations.
to sum up and to put objectives in a glance, the reference model needs to:

- Comply with standards;
- Be simple and manageable so it can be used by the manufacturers;
- Identify the gaps and loopholes lead to low sustainable performance;
- Identification of overlaps and stipulation of preferred solutions;
- Prepare a holistic assessment for the manufacturing organization.

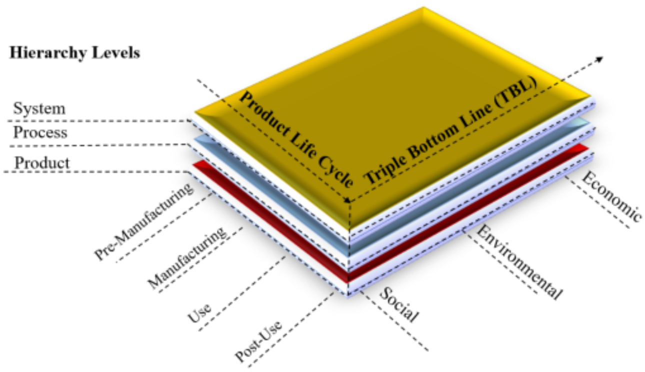


Figure 15 - Three-Dimensional Model for Sustainability Assessment [DT3]

Based on the abovementioned, three dimensions of sustainability must be assessed in all levels of a manufacturing organization throughout the whole life cycle of the product. To make that assessment possible the three-dimensional model is proposed as shown in figure 25. To respond to the need of having a global and not an ad hoc methodology, each cubical of the 3D model (Figure 16) will introduce a standard indicator or measurement based on GRI as it was selected as the indicator source.

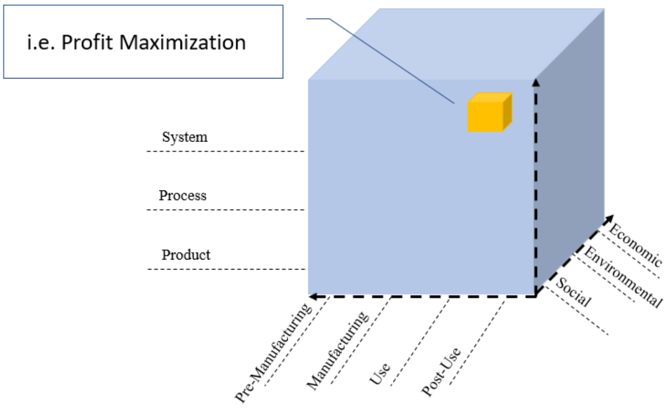


Figure 16 - An example of a sustainability cubical [DT3]

A step-by-step development of indicator-based model was described from selection of the indicators to their normalization, weighting and aggregation of the indicators to achieve a final aggregated composite sustainability development index. However, the study contributes also to the process of selection and weighting of the indicators which has been done based on the association rule mined and FCA lattice respectively. Finally, the effectiveness of the model was validated through its application on a real manufacturing case company.

2.4 Concetta Semeraro's thesis: Contribution to the formalization which is driven by the data of modelling invariants of cyber-physical systems

Manufacturing enterprises are facing the need to align themselves to the new challenges of the market demand: one possible path is to endeavour the new information technologies. These are facing with an array of Industry of the Future challenges: "digital requirements" thus they need to be accurately assessed by sound analysis and deep understanding of the operational and technological criticalities in the manufacturing operations.
The Smart Factory [Zuehlke, 2010] paradigm has born with the so-called "fourth industrial revolution" in manufacturing. It can be synthesized by integrating networks of physical components with control software that contribute to the "smartness" of the overall system, for better managing manufacturing processes. Smart systems typically are well to support decision making based on sound data. This is why smart systems include a variety of technological components, including sensors for signal acquisition, communication units for data transmission between components, control and management units for supporting decision making, and actuators for performing appropriate actions: the coordinated network of sensors is here called the sensing system. The CPSs have amplified the possibility to sense the world through a network of connected devices. The combination of intelligent systems and sensing systems forms a large-scale distributed cyber-physical system, that has enormous potentialities to bring smartness into many application areas. CPSs, on the other hand, suffer from a lack of

modelling techniques for taking into account not only their technological parameters but also their high degree of information and functional inter-relationships.
Another concept associated with the cyber–physical integration is the Digital Twin (DT) [Koulamas, 2018]. The DT refers to a "virtual" image of a system (a twin) synchronized with the real operating scenario, providing as much useful information as possible from this latter, to allow the human decision maker drawing sound decisions. This "digitalization" of the reality is an approach that enables the virtual replication of the factory to forecast and control real time physical processes. It may allow to connect the entire value chain [Ghobakhloo, 2018] by merging sensor data acquired from the physical world into virtual or simulation-based models. Predicting the system behaviour is thus a matter of running these realistic and "up-to-date" models.
As the complexity of these systems continues growing, the challenge of developing integrated intelligent and sensing systems has surpassed the design complexity of their individual components. The main problem in the development of smart and sensing systems lies in the complexity of integrating and managing different components and technologies. A possible solution to this problem is to formalize the shared knowledge and define a modelling approach that helps to design systems regardless of the context. Shared knowledge representation is a branch of artificial intelligence that studies the way human reasoning occurs and defines symbols or languages [Roschelle, 1995] for this purpose.
In this context, the thesis aims firstly at providing an up-to-date picture of the state-of-the-art of main features and challenges about CPS and DT design and use, focussing on the different application domains and their related technologies. To answer the main questions on how a Digital Twin must be designed and implemented, the thesis funds its scientific basis on information, principles and hints derived from a systematic scientific literature review. The outcome is then a multi-perspective picture of the Digital Twin paradigm, forming a framework emerging from scientific literature as well as the direct on-the-field experience of the authors.
The thesis aims to identify an approach to formalize data-driven invariant modelling constructs for improving the smartness of manufacturing processes and products, involving networked components. The idea behind data-driven invariant modelling constructs is to permit the re-use of predefined functional patterns for building digital models based on the specific application. The approach makes shared knowledge more easily reusable and it is the basis of some standardization efforts. The use also of Multi Relational Data Mining techniques, in the specific case of RCA [Valtchev, 2004], allows the extraction of tacit knowledge embedded in the (big) data coming from the analysed processes.
The thesis proposed a series of modelling patterns (data-driven invariant modelling constructs) for the digital transformation of industrial production systems and a prototype for the analysis of a real industrial process on a production line to validate the knowledge extraction approach. The resulting tool can exploit existing knowledge and

information from real systems to identify problems and to propose potential improvements.

The thesis defines the following research questions:
RQ1: What is the existing research work on digital twin: 'Which technologies need to be adopted or new ones explored?' 'How to design a Digital Twin?' 'How to implement a Digital Twin?'

One of the first steps in our research is to investigate and to define the digital twin paradigm in the context of Industry of the Future. We conducted a systematic literature review to identify the context, the applications, the functions in product life cycle, the possible architectures and the components of a DT. Next, we intend to combine the different aspects of a Digital Twin and some related issues for developing our proposed digital twin paradigm. Hence, we need to link the different approaches related to the digital twin modelling for developing a common approach to detect data-driven invariants modelling constructs. [IJV4]

RQ2: How to develop data-driven invariant modelling constructs?
To answer to this question, we conducted a second literature review on model-based, data-driven and hybrid approaches to define the contribution positions. The idea is to identify and to formalize invariant modelling based on data analysis. The idea is to use, and especially re-use, predefined data-driven constructs for building digital models for different applications.

The contribution positioning of the thesis is to make the modelling more structured and reliable. The key issue to be addressed is the invariance. The idea is to detect automatically from data invariant modelling constructs. Invariant modelling constructs need to be developed to describe/emulate a system independently from its context of application. Data-driven invariant modelling constructs can be used and in particular re-used to create digital models for different systems or processes [C30].

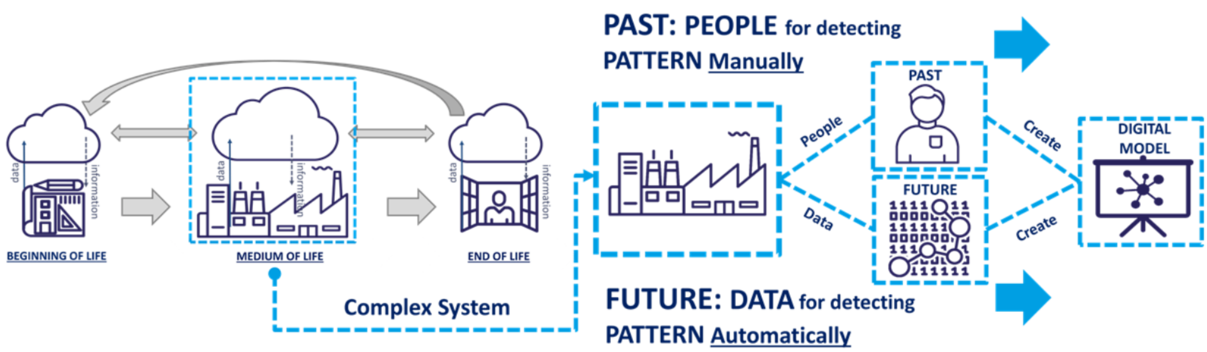


Figure 17 – Research context [DT4]

The process to detect invariant modelling constructs is shown in Figure 18. The first aspect is to identify the systems to analyse (1), as well as the choice of products and the related manufacturing process. A model-based approach (2) is required to model the system describing the function, the structure and the behaviour. The model-based approach draws a detailed representation of system under consideration and it enables the selection of data to analyse (3). When the system and the data are selected, the data-driven approach is used (4) to detect and to discover automatically associations and relationships among data. The associations can describe recurrent behaviours of the system and it can codify tacit knowledge that can be used to better understand the behaviour of the system. The discovered association are analysed based on the model developed to extract knowledge (5) from data and to define the physical meaning of the associations. The extracted knowledge represents a data-driven invariant modelling construct. Data-driven invariant modelling constructs can be formalized (6) to be easier to understand and to analyse where and when these can be applied. Data-driven invariant modelling constructs enable to design the virtual model of a system for realising its digital twin (7). Data-driven invariant modelling constructs can be reused among systems or processes operating in a similar condition (8).
The next chapter presents the approach to detect data-driven invariant modelling constructs.

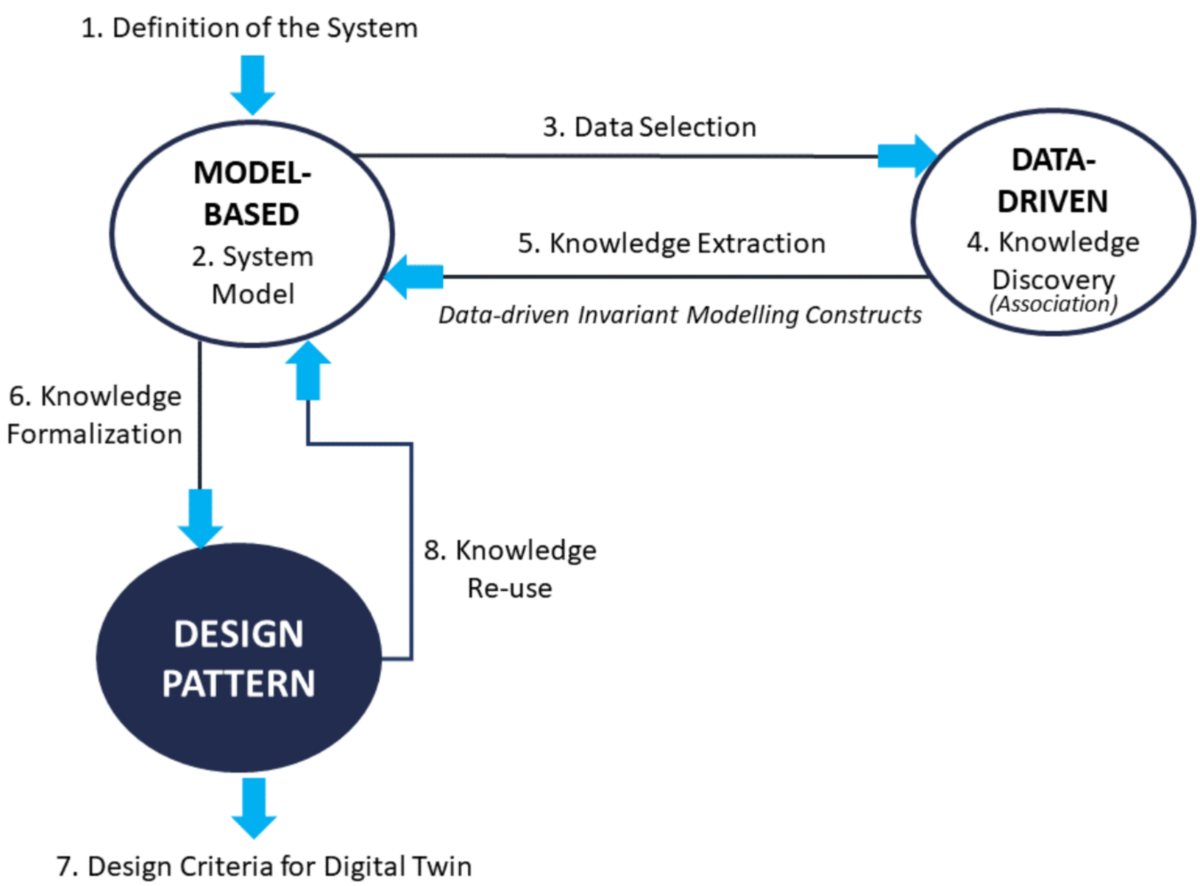


Figure 18 - The Process for Detecting Data-driven Invariant Modelling Constructs [DT4]

2.4.1 Definition of the approach to extract data-driven invariant modelling constructs

The approach is articulated in eight different stages as shown in *Figure 19*. The stages are: 1) Definition of the system; 2) System model; 3) Data selection; 4) Knowledge discovery; 5) Knowledge extraction; 6) Knowledge formalization; 7) Design criteria for building a digital twin; 8) Knowledge re-use.
The follows paragraphs explain each single stage in more detail.

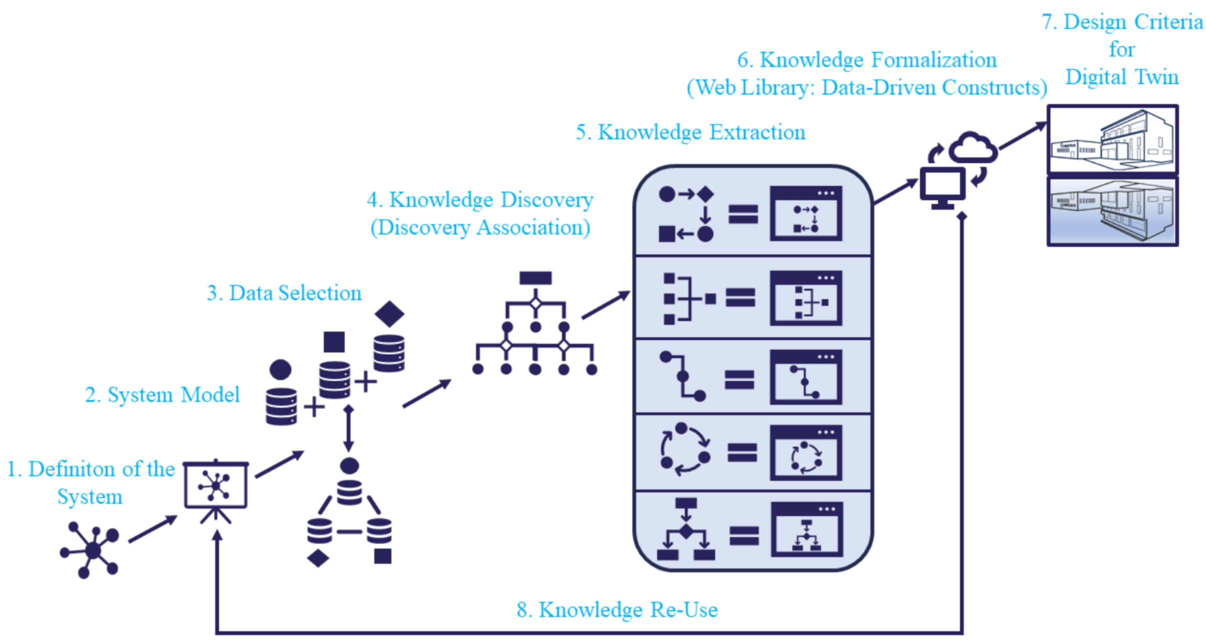


Figure 19 - The Approach to Extract and to Formalize Data-driven Modelling Construct [DT4]

2.4.2 Knowledge discovery

Multi relation data mining (MRDM) and specifically the Relational Concept Analysis approach is the data-driven approach selected for discovering association. RCA extends formal concept analysis by enabling multi relational information [Rouane-Hacene, 2013]. Given a set of objects, a set of attributes, and defined the relations between objects and attributes, a formal concept represents a subset of objects sharing the same sub-set of attributes. A concept is constituted by two parts: its extension which consists of all objects belonging to the concept, and its intension which comprises all attributes 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 attributes [Williams, 2006].

Data modelling

Data modelling is the process of creating a data model for the data stored in a Database. The data model illustrates the logical structure between databases by defining the entities (objects), their attributes, and showing the relationships between them.

The data selected in the previous step need to be organized in a data table as shown in. The data table presents the objects (G) on the rows and the attributes on the columns (M). The cross indicates that exists a relation between an object and an attribute (I). The relation between an object and an attribute is binary (exist/ does not exist).

Discovery association

RCA is applied to discover automatically association rules in data. RCA, using lattice creation algorithms, in the specific context using an improved version of "Inclose V" algorithm [Andrews, 2009], converts automatically the data table into a concept lattice [Venter, 1997]. Concept lattice in *Figure 20*b graphically portrays the underlying relationships between the objects and attributes for extracting the useful information [Wille, 2002].

2.4.3 Knowledge extraction

The concept lattice, as mathematical abstraction of concept systems, support knowledge extraction. The concept lattice needs to be analysed to extract data-driven invariant modelling constructs (patterns). The data-driven constructs are evaluated with the SysML models as shown in *Figure 20*c to understand the meaning of their content. It aims to validate if a certain model (or hypothesis) is consistent with the available data or if it is necessary to collect new data. It means that it is possible to verify the comprehensibility, the usefulness and the robustness of the model. At the same time, the data-driven constructs can detect tacit knowledge or discover associations to improve the model.

A data-driven invariant modelling construct contains the following sections:

- ID Pattern: unique name that helps in identifying the pattern.
- Pattern Name.
- Description: the goal behind the pattern and the reason for using it.
- Example of association rule: an example of a concept for the pattern in analysis.
- Data View: a representation of the data for the pattern in analysis.

- SysML Model View: a graphical representation of the association between data. Block definition diagram or internal definition diagram or parametric or state chart diagram may be used for this purpose.
- Applicability: situations in which this pattern is usable.

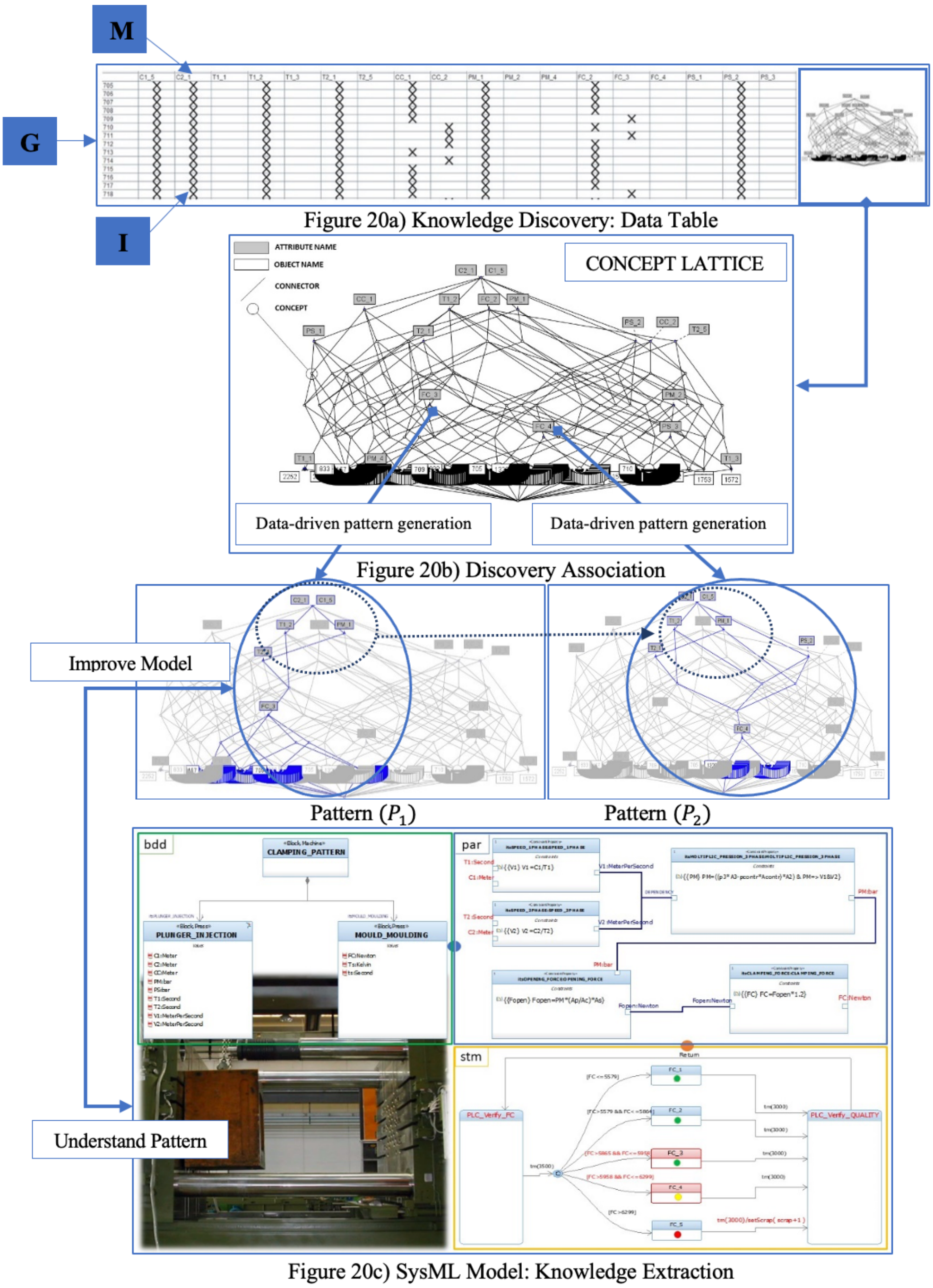


Figure 20 - From Knowledge Discovery to Knowledge Extraction [C30]

2.5 Mickael Wajnberg's thesis: Relational concept analysis: a versatile method for knowledge extraction

This thesis focuses on the knowledge extraction discipline. It was developed to meet the needs for synthesis and analysis of data sets too large to be humanly gripped. The Wajnberg's work begins by considering the definition of extracting knowledge from a dataset. It continues with a semantics definition of [Fayyad, 1996] that is a key concept for the development of the thesis. Data are considered as inputs, and information as the set of possible aggregations of them. Knowledge extraction is then the discipline of designing an information-based knowledge model. The scientific work highlights the key importance of the association rules, a form of information presenting the cooccurrences between characteristics of objects. The proposed method extracts the association rules, while preventing the production of references to the pre-missing in the conclusion that hinder their interpretation. This new form of association rules, if formalized in a univocal way, allows the discovery of non-trivial relational associations of the concepts generated by the RCA. These association rules are of the form if an object has the characteristics "a" and "b" then it also has "c", "d" and "e". The most classical measures of interest for these rules are support, the number of objects concerned by a rule, and trust, the truthfulness of a rule. The contributes of this scientific work are the development of the RCA to allow the extraction of association rules. In particular how to bypass the presence of cycles in the object descriptions on any binarized multi-relational dataset. In addition, it reinforces the theory of RCA by redefining the notion of formal concept, mathematically demonstrating the interest of RCA in relation to FCA, and by introducing certain improvements aimed at eliminating redundancy in the rules that could arise due to the inherent relationship between relational characteristics.

2.5.1 Formal Concept Analysis

The formal concepts analysis, introduced in [Wille, 1982], analyses the unary data sets. They are composed of objects, and attributes that characterise them. It is an algebraic method which aims to discover the abstractions, called formal concepts, of such data sets.

2.5.1.1 Formal context

A formal context is a triplet K = (O, A, I) where
- O is called the set of objects,
- A is called the set of attributes and
- $I \subseteq O \times A$ is a binary relationship, called the incidence relationship.

A formal context is a representation of a set of unary data. It can be represented by a unary table. Following an idea developed in Mickael's thesis it will be presented an example that will illustrate all the definitions and its representations.

Let us consider four cats: Demetra (dm), Lea (le), Pepita (pe) and Talpi (tp). Among the different characteristics of those cats, we highlight the following properties:

- always being hungry (sa);
- always wanting cuddles (co);
- having a unique colour coat (mc);
- and always protecting their territory (te).

Having this information, we can define a formal context.
The formal concept K = (O, A, I) where O = { de, le, pe, tp } is the set of objects, and A = { sa, co, mc, te } is the set of attributes. Such a context can be represented in the form of a unary table, where the elements of O correspond to the rows, the elements of A to the columns, and the incidence relationship I is identified by the crosses. A representation is given in the Table 6. A cross in the K [o_i, a_j] box indicates that the object o_i has the attribute a_j.

Table 6 – Example of a Formal Context

K	sa	co	mc	te
de	x			x
le	x		x	
pe			x	x
tp		x		x

2.5.1.2 Derivations operators

The FCA aims to extract sets of objects with common attributes. Before defining the formal concept notion, it will be presented the basic operation of the FCA, called derivation that is crucial to the understanding of the formal concept.
The derivation operation on objects is an application from P(O) to P(A). For a set of objects X ∊ P(O), its derivative, denoted X', is given by:

$$X' = \{a \in A \mid \forall o \in X, (o, a) \in I\} = \cap_{o \in X} \{a \in A \mid (o, a) \in I\} \qquad (2.4.1)$$

Symmetrically, the attribute derivation operation is an application from P(A) to P(O). For a set Y ∊ P(A), its derivation, denoted Y', is given by

$$Y' = \{o \in O \mid \forall a \in Y, (o, a) \in I\} = \cap_{a \in Y} \{o \in O \mid (o, a) \in I\} \qquad (2.4.2)$$

Thus, the derivation of a set of X objects is the set of attributes of A carried jointly by all the X objects. In a dual way, the derivation of a set of attributes Y is the set of objects of O carrying jointly all the attributes of Y. The following example shows those definitions on the formal concept defined in the Table 6.

- {le, de}′ = {sa}
- {pe, de}′ = {te}
- {te}′ = {tp, pe, de}
- {mc, te}′ = {pe}

2.5.1.3 Formal concepts

A pair C=(X,Y) ∈ P(O) x P(A) such that Y =X′ and X = Y′ is called a formal concept. X is called extent, and Y is called intent, of the concept C. The concepts are the fundamental abstractions that the FCA aims to extract from a formal context.
A formal context can be represented in different ways, with a two-part graph or can be visualised as a maximum rectangle of the cross after swapping the rows and columns of the table.

However, regardless of the preferred graphic representation, the intuition behind the formal concept is always the following: a concept is a pair of sets of objects and attributes, such as

- the objects carry all the attributes and are the only ones in the context to jointly carry all the designated attributes.
- each of the objects carries all the attributes, and no other attribute of the context is common to all the objects of the concept.

Formal concepts can be characterised as closures and they are intrinsically linked to the notion of equivalence class. The single maximum of an equivalence class is called closure [Ganter, 1999]. It contains one or more minimum elements. An element U ∈ [Y′′] is said to be minimal if for any V ⊂ U, V ∉ [Y′′]. These minima are called class generators. Finally, since the formal concepts are described by the set {(Y', Y'') | Y ∈ P(A)}, the set of closures of the equivalence classes corresponds exactly to the set of intentions of the formal concepts.

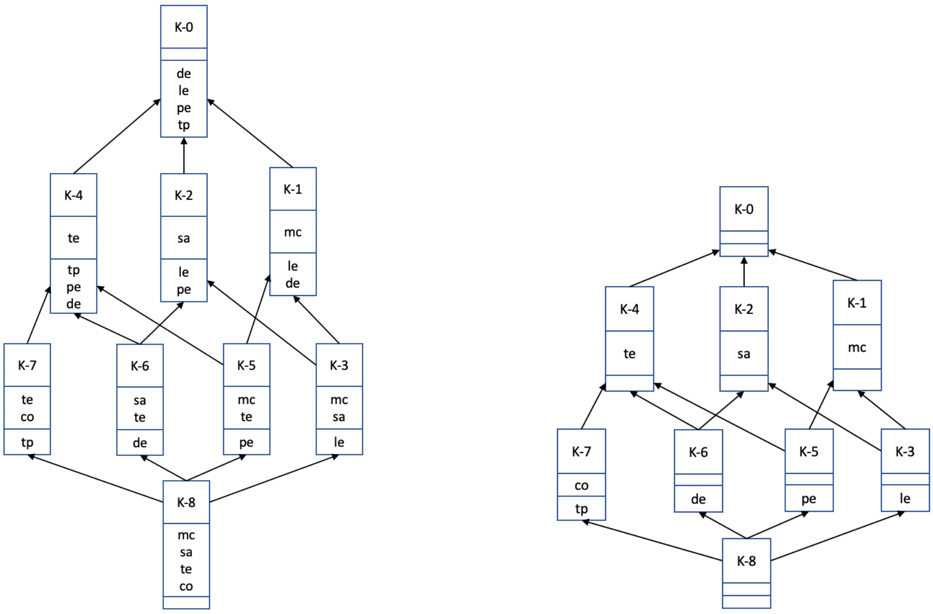
2.5.1.4 Concept lattices

Let K = (O, A, I) a formal context. Let us note:

- C_K the set of all the formal concepts of P(O) x P(A) and
- $\leq_K$ the relation of inclusion on the extensions of the concepts.

The partial ordered set (poset) $\mathcal{L}_K$ = (C_K, $\leq_K$) forms a complete finished lattice. It is called the **concept lattice** of context K [Ganter, 1999]. Equally, selecting the inverse inclusion relation on concept intensities produces the same concept lattice [Ganter, 1999].

The representation of the Hasse diagram in Table 6 is shown exhaustively in Figure 21a. Each concept is represented by a box of three sections which, from top to bottom, are the concept identification number, its intention, its extension.



(a) Complete (b) Compressed

Figure 21 – Hass diagram of Table 6

The Figure 21 (a) and (b) both present exactly the same understanding of the contextual concepts presented in Table 6. The Figure 21(b) is a compressed representation of the Figure 21(a). Each identically numbered vertex on the two grids presents exactly the same concept. The reading of a concept in compressed representation is therefore interpreted as follows:

- the intention of a concept is the set of attributes of the represented intentions.

of the concept under consideration and of any super-concept.

- the extension of a concept is the set of objects of the represented extensions of the considered concept and any sub-concept.

The organisation of the concepts, in the form of a lattice in Hasse diagram representation, allows the quick extraction of all the existing closures and generators on a context.

2.5.2 Relational Concept Analysis

The formal concept analysis, as described, aims at extract knowledge, in the form of association rules, which is distilled on a homogeneous dataset. The extension of this method to the multi-relational paradigm is named relational concept analysis (RCA), and it is presented in [Rouane-Hacene, 2013].

The RCA, introduced in [Huchard, 2002], aims to extend the FCA to relational data compatible with the entity-association model [Chen, 1976]. Such a model considers binary relations between objects. The RCA enriches the description of objects by integrating so-called relational attributes, which reveal the relationships between objects. The RCA is a method of multi-relational data mining (MRDM). It therefore exploits different types of objects as well as links between them.

2.5.2.1 *Relational family of contexts*

Let us consider two formal contexts $K_i = (O_i,A_i,I_i)$ and $K_j = (O_j,A_j,I_j)$. A relation having for domain K_i (also called source) and for co-domain K_j (also called target), is a binary relation subset of $O_i \times O_j$. We note $R_{i,j,k}$ the k-th relation of domain K_i and co-domain K_j.

The FCA does not allow direct use of the information included in such a relationship. Through the integration of a scaling system, the RCA makes it possible to extend the FCA to integrate this information and thus characterise the objects of a source context according to the relationships maintained with other objects.
A relational family of contexts (RFC) is a pair (K,R) such that:
- K is a set of formal contexts $K_i = (O_i, A_i, I_i)$
- R is a set of relations $R_{i,j,k} \subseteq O_i \times O_j$ for $i, j \in \{1, ..., |K|\}$

2.5.2.2 *Propositionalisation operator and Graduation*

Lachiche defines propositionalisation as a process explicitly transforming a relational dataset into a propositional dataset [Lachiche, 2010], that is, a set of arity 1 literals. Among the main interests of propositionalisation are the construction of features that can be combined into hypotheses [Srinivasan, 1996] and the reduction of research space [Lachiche, 2010]. It should be noted that this language bias (the choice of operators) does not guarantee that it is possible to reconstitute the complete information of the data set with all the statements made by propositionalisation [Krogel, 2005]. A propositionalization operator makes it possible to transform a class, via a relation into an attribute.
The graduation consists in extending a context $K_i = (O_i, A_i, I_i)$ of a relational family of contexts by integrating relational information, in the form of attributes. The relational attributes translate the links between objects in the two contexts into a relationship.

The resulting RCA lattices are difficult to interpret because of these relational attributes. The result extraction methods are based either on fixed-point descriptions which may be recursive and not very easy to interpret, or on intrinsically circuit-free models which do not cover the general case.

2.5.3 Improvement of the Relational Concept Analysis

The impact of relational attributes on RCA association rules and information redundancy is one of the studies and results of the Wajnberg's work.

2.5.3.1 Association Rules

The ultimate goal of our lattice construction procedure is the knowledge extraction in the form of association rules. Moreover, the concept lattice serves as a basis for extracting a compact representation of the set of association rules. Whether a lattice is produced in a FCA or RCA process the way to extract the rules is the same: for each concept C we extract all the rules g → Y \g where g is a generator of C, and Y is either the intention of C or that of a concept that is the immediate predecessor of C.
However, the RCA introduces relational attributes, which, like non-relational attributes, are incorporated into the rules. For a rule to be usable, however, it must be possible to interpret the rules, including by resolving references to the concepts of these attributes. In effect, a rule of the form a_1, a_2 → a_3, pr: $C_{j,k}$ does not bring any information if we cannot characterize $C_{j,k}$ Several options are thus available to us to clarify these relational attributes:

Extension. Since the extension of concepts is preserved throughout the iterations, we can replace the reference to the concept by this one. In such a case, a relational attribute then characterizes the direct links between objects. However, this kind of modelling denotes the very nature of association rules, which is to identify tendencies, by freeing oneself from instances, through abstractions; moreover, it seems rather unpromising when the list of objects of a relational attribute is large.

Intention. Conversely, if one decides to replace the reference by the concept's intention, several ambiguities emerge. First of all, a choice must be made about the definition of a concept, since it presents several options. By choosing the evolving concept option several stages of intention are possible. In purely informative terms, for a concept, the most accurate for describing its objects is fixed-point intent. However, within this framework, a problem inherent to RCA emerges cyclical dependencies. In effect, the relational paths of the multigraph can form circuits. The interpretation of such circuits is a difficult case and requires the use of the theory of smallest fixed points, as pointed out [Baader, 2003].
Finally, considering the canonical generators to represent a concept in a relational attribute without a cycle, it allows the extraction of rules.

2.5.3.2 Reasoning

Another notable limitation of the RCA to which we propose a solution is what can be called relational redundancy. In effect, the idea behind the FCA's production of association rules is to extract a minimum set covering all information. RCA, of course, uses the FCA method of maximum rule extraction whose premises are generators and, in fact, extracts a set of rules without attribute redundancy. However, an association rule A → B where A (intersects) B = (void set) having relational attributes, may have redundancy that comes from the definition of the scaling operators

The Wajnberg's work presents a modified version of the RCA. Among the proposed modifications there is the possibility of considering several propositionalization operators per relationship. For this, it is introduced an associative table which, in a relation $R_{i,j,k} \rightarrow R$ associates a set of operators to be applied for graduation.

The PhD thesis formally demonstrates the interest of RCA in relation to the application of FCA on a relational dataset, after aggregation of contexts. It presented the limits of formalism, in particular the definition of a formal concept within a framework of contexts extended in an iterative way. It demonstrated that with a more exhaustive formalism, the rules of association could be extracted and defined without any attribute to be solved recursively.

3 Research and teaching project

Dijo que... necesitaba terminar el poema, porque en un rincón... había un Aleph.

Explicó que un Aleph es uno de los puntos del espacio que contiene todos los puntos. […]

Sí, el lugar donde están todos los lugares de la tierra, sin confundirse, visto desde todos los rincones. No revelé mi descubrimiento a nadie, pero volví a él de nuevo...

He said that... he needed to finish the poem, because in one corner... there was an Aleph.

He explained that an Aleph is one of the points in space that contains all the points. […]

Yes, the place where all the places on earth are, without being confused, seen from every corner. I did not reveal my discovery to anyone, but I returned to it again...

Jorge Luis Borges – Aleph

The introductory chapter of this document presented the issues related to knowledge management in systems in factories of the future contexts. In order to make these systems more effective and allow them a more complete interoperability, the implicit knowledge of the systems studied must be made available.
The research work I made during the last eight years were in line with the above issues: the semantic interoperability of systems and the formalization of knowledge. The presented solutions, through my published papers and supervised thesis, were achieved through the semantic annotations approach, the ontologies, the methods of synchronization of information and products through CPS and approaches based on multi relational data mining techniques.
Despite this, the approach presented must be extended in order to increase its effectiveness and generality.

As presented in the general introduction part the smart products and "data-driven technologies" enable the possibility to acquire useful data from heterogenous sources within the enterprises as well as from customers and suppliers. Therefore, Industry 4.0 stresses the huge potentialities of data that can be used in real time, enriching contextual knowledge or generating new one in the way products can be produced and used.

3.1 Research project: Formal methods for extracting and reusing knowledge from heterogeneous sources for semantic interoperability of distributed architectures in a Factories of the future context.

From 2014 I decided to evolve my research activities towards the study of the problems of formalization and knowledge extraction in systems characterized by heterogeneous data sources in an industrial environment. My research project obviously relies heavily on the results presented during the last 8 years. The next section presents the path from data recovery to knowledge extraction.

3.1.1 From rough data to knowledge extraction: a long path

As well highlighted by a young researcher I worked with, Mickael Wajnberg, in his PhD thesis, to let understand the importance of knowledge extraction we need to make a path from rough data towards the knowledge extraction taking in account various steps inside the knowledge engineering discipline.

Knowledge engineering is the discipline of designing the knowledge base. It consists of seven steps [Russell, 2002]:

1. Identifying the task and defining the boundaries of the knowledge to be represented.
2. Acquire the knowledge.
3. Formalize the collected knowledge.

4. Encode the formalized knowledge.
5. Encode the description of the problems to be solved.
6. Apply the reasoner.
7. Correct the knowledge base (back to step 1).

Two of those steps are between the core of my research interests, the task of knowledge extraction (step 2) for knowledge bases compatible with the OWL2 format, the ontology implementation language recommended by the W3C [W3C, 2012] and the formalization of the collected knowledge.

3.1.1.1 The DIKW pyramid

In information science, the terms data, information and knowledge are often used. In this field, these terms follow the definitions of the Data-Information-Knowledge-Wisdom (DIKW) model, often represented by a pyramid. One example can be presented in the Figure 22.

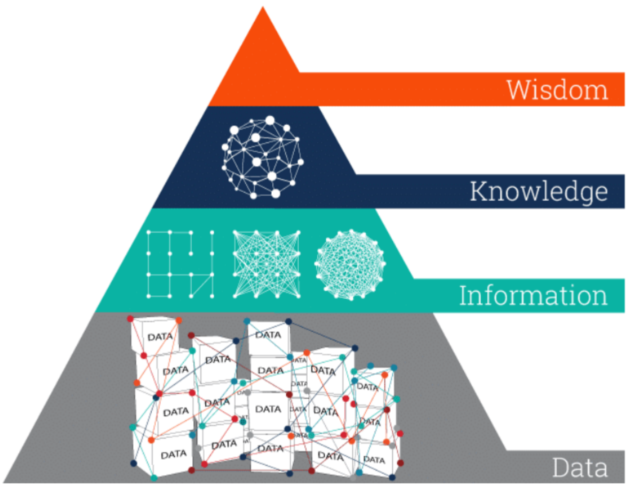


Figure 22 - The progression from data to information, knowledge, and wisdom [Ackoff, 1989]

The terms data, information and knowledge as defined in the article [Fayyad, 1996], serves as a reference in our vocabulary. In this article, we read "*Historically, the notion of finding useful patterns in data has been given a variety of names, including data mining, knowledge extraction, information discovery, information harvesting, data archaeology, and data pattern processing*" and "*Given these notions, we can consider a pattern to be knowledge if it exceeds some interestingness threshold*". Thus, starting from the postulate that knowledge extraction consists in discovering a set of information, in the form of regularities. This set must be organized so as to present only the most relevant information, for an optimal integration in the knowledge system.

3.1.1.2 Knowledge representation

In the literature, there are multiple logical formalism and knowledge representation formalisms. The modal logics [Schild, 1991], restrictions of first-order logic to decidable fragments [Andréka, 1998], database models, as well as knowledge representations used in the field of artificial intelligence, such as semantic networks [Brachman, 1979], the frame system [Hayes, 1981] and conceptual graphs [Sowa, 1983].
In the last years with the imposition of the web of data, the W3C, the international standardization body for web technologies, recommends the development of ontologies in the OWL2 format for reuse and interoperability purposes [W3C, 2012]. OWL2 is a language based on the *SROIQ* description logic. Thus, it is important before presenting the definition of ontologies for a knowledge base, to present the ontology formalization languages, notably the *SROIQ* description logic [Horrocks, 2006].

First-order logic is composed of functions, variables, constants, and predicates that are used to express the properties and relationships of these variables and elements. A term is a variable, a constant, or the composition of one of these elements in a function. The predicates, added to their terms, form atoms, the fundamental brick of this formalism. The abstraction above the atom that encompasses its binary valuation (true or false) is called a literal. Predicates, like functions, can link several terms. The number of links is called the arity of the predicate.
First-order logic is particularly expressive. However, it presents a major flaw in the knowledge that can be formalized from the data: it is not a decidable logic. That is, it is possible to design formulas, for which it is impossible to decide whether they are true for all data. The undecidability of this logic is a major problem when implemented in a computer system. One could, among other things, ask to verify a formula that would trigger an infinite processing, without having a way to know if the processing is indeed infinite, or simply very long.
However, some fragments (or restrictions) of first order logic are decidable. Decidable fragments include the bivariate restriction, where all predicates and functions have only arities 1 or 2. Description logics are logics based on first-order logic accepting some of its connectors and quantifiers. They are (usually) restricted to two variables and decidable [Horrocks, 2006].
There are multiple description logics. I focused my interest on the *SROIQ* logic [Horrocks, 2006]. This choice is motivated by the fact that *SROIQ* has been selected by the W3C for the specifications of OWL2, the recommended definition language for ontologies.
The interest of such a formalism is, on the one hand, its expressivity, and on the other hand, the fact that be decidable [Horrocks, 2006].

How to structure the knowledge
A knowledge base is a set of sentences expressed in a knowledge representation language. Sentences that are immediately considered true, without being inferred, are called axioms

[Russell, 2002]. These sentences can either describe the schema of the knowledge base, i.e., present a set of concepts and the relationships between them, or describe the instances, either the links between them or their links to the elements of the schema. The schema of such a base is called knowledge base ontology [Russell, 2002].

Ontology has its origins in philosophy and refers to the study of being as such [Gašević, 2009]. Artificial intelligence (AI) borrowed the word and changed its meaning. In AI, the main question is what an AI system has to reason about to be able to perform a useful task [Borst, 1997]. Now, its importance is being recognized in research fields as diverse as knowledge acquisition [Tudorache, 2013], medicine [Arsene, 2015], knowledge representation [Yao, 2013], language engineering [Guizzardi, 2015], among others.

There are many definitions of the concept of ontology. The most accepted one is from [Gruber, 1995], which states that an 'Ontology is an explicit specification of a conceptualization', meaning that ontology is a description of the concepts and relationships that exist in a domain.

An ontology on a certain domain aims to capture, represent, share, (re)use and exchange the common understanding about the concepts in the domain, their taxonomies, classification, their relationships and the domain axioms [Gašević, 2009]. Ontologies enhance knowledge sharing and reuse across different applications [Neches, 1991]. Ontology is also used to unify Databases, Data Warehouses, and knowledge bases vocabularies, in order to overcome the obstacles of knowledge integration, which basically consists of merging past and new knowledge [Djellali, 2013].

Depending on the level of the knowledge that an ontology aims to represent, ontologies can be generally categorized into three levels as follows [Roche, 2003]:

(1) Top level ontology, which specifies only general concepts and relationships (such as time and space) and can be used in different domains.

(2) Domain level ontology, which captures the knowledge that is dedicated to a specific domain (such as production domain) and can be use and reused for different tasks in the same domain.

(3) Application-level ontology, which represents the specific knowledge that is dedicated to a task in an application and normally is not reusable for other applications.

Knowledge discovery

"Knowledge discovery is the non-trivial process of identifying patterns in data that are valid, new, potentially useful and ultimately understandable" [Piatetsky-Shapiro, 1996]. There are primarily three high-level objectives for which such a process should be implemented: to test a hypothesis, to create a prediction model, or to create a description model [Fayyad, 1996]. A prediction model aims to estimate certain values for new data, or the evolution of certain values for existing data [Fayyad, 1996]. A descriptive model aims to detect and describe patterns (clusters and trends) in the data [Fayyad, 1996]. To achieve these objectives, several phases are required: data selection, pre-processing, transformation, data mining, interpretation and evaluation of the information [Fayyad, 1996].

Knowledge extraction

Extracting knowledge from a data set requires analysing the available data to extract the regularities that make up the information. The discipline of automated analysis of computerized data sets is called data mining [Fayyad, 1996]. The discipline of data mining brings together a set of methods that study the regularities present in a homogeneous data set [Ye, 2003].

Multi-Relational Data Mining (MRDM) [Džeroski, 2003] is a more structured sub-domain of data mining that aims to observe data by considering several tables, each representing a different type of entity, as well as the relationships between data from different tables. This last alternative is of interest here. Multiple MRDM methods have been proposed, responding to extended versions of the problems of classical data mining [Džeroski, 2003]. Among these applications are regression [Cai, 2005], classification [Frank, 2007], association discovery [Galárraga, 2013] and clustering [Nergiz, 2008].

3.1.1.3 *Formalised Knowledge extraction (from data to Ontology)*

Ontology extraction aims at automatically designing an ontology from data. Whether the method is manual, automatic or semi-automatic, the procedure must take into account the format of the data: unstructured (e.g., text in natural language), semi-structured (e.g., text in naturel language with words and sections marked by labels), or structured (e.g., database). The report by Hazman discusses various techniques for extracting unstructured and semi-structured data, statistical approaches, automatic language processing techniques, and various data mining techniques [Hazman, 2011]. Other researchers have been interested in structured data. Many tools are available and referenced [Unbehauen, 2012]. For example, the RDB-to-Onto tool [Cerbah, 2008] aims to extract an ontology from a relational database using both schema and data to discover and enrich concepts.
Several of these tools are notably based on a process of association rules. The association rules are a formalism which makes it possible to express regularities of the form "X% of the objects which have the properties of A also have the properties of B" [Agrawal, 1996]. They have been used both on unstructured data to study correlations between words in a text and their categories in a taxonomy and on structured data [Maedche, 2000], both for ontology fusion and for the discovery of subsumption relationships between classes [David, 2007].

3.1.1.4 *Summary of the presented process of knowledge extraction*

First, we start by defining the notion of knowledge as a set of interesting regularities in a set of data. Then, we present various formalisms allowing to express this knowledge, in particular the logics of description. Finally, methods for extracting the information, in the form of a rule of association, serving as an intermediary step between data and knowledge, are presented. However, it can be noted that the methods for extracting these rules on a relational dataset are limited either by format constraints (rule conclusion can be limited to a literal of arity 2) or by data annotation constraints (positive and negative examples to be provided as input).

3.1.2 Research project description

The research project focuses on the creation of mathematical models and the implementation of intelligent sensors, Cyber-Physical Systems (CPS) to enrich the layer of data that comes up from the field. One of the most relevant scientific challenges is the lack of formalization (in other words mathematical) of models of systems and the information systems that result from them, as well as the definition of the semantics of the concepts and relationships they apply, in order to ensure their common understanding and to facilitate their interoperability by minimizing semantic losses.

In order to make precise and concrete scientific contributions to this project an approach to the engineering of interoperable systems [Ramos, 2011] and [Morel, 2003] will be used that consists of relying on different types and levels of abstraction or models. These models must express and formalize not only the "structural" aspect of the system components, but also their behavior [Maier, 1998], which may be limited by the specific requirements of the system domain. Another type of constraint may be induced by the interoperability protocol(s), which may impose strict rules to endow interoperable systems with properties such as autonomy, confidentiality and transparency [Zdravković, 2016].

The objective of this research project is twofold: on the one hand, to model data from heterogeneous sources and create knowledge from them and, on the other hand, to study the problems posed by model-driven engineering in cooperative systems. Involving cooperation concerning "systems of actors" willing to interoperate. Collaborative systems are now organized in networks, or as complex systems [Camarinha-Matos, 2014].
The complex systems envisaged will be composed of networks of CSPs, intelligent sensors, which will retrieve data by inserting the context and thus form information networks.
Collaborative objects can be likened to an open object reticular system of object systems [Brownsword, 1998]. Specifying such a network involves moving from a single integration paradigm to an interoperation paradigm [Fisher, 2006].
One of the requirements of this need for collaboration concerns the ability of these components to interoperate, i.e. their interoperability, which is more or less total.

The scientific challenge is thus to make available languages and modeling tools adapted to each project of systems with distributed architecture, despite the heterogeneity of business skills and the multidisciplinary structure of the domains. This challenge has two dimensions: on the one hand, the capacity of modeling to provide tools for business processes, which requires the definition and formalization of their invariants; on the other hand, the study of the conditions of use of the models in practice, which is always evolving and uncertain.

The analysis of formal concepts [Priss, 2006] is a useful and powerful tool for formally describing the links between any objects (which form a context), in particular between objects conveying knowledge. This method is based on lattice theory [Wille, 2009], which can be used to solve problems of interoperability assessment between information systems within companies.
As presented the Relational Concepts Analysis, an extension of the FCA mechanisms, has been introduced in [Rouane-Hacene, 2013] where the focus is on data sets compatible with Relationship Entity Models (RE) [Chen, 1976] or, alternatively, with the Resource Description Framework (RDF) [Miller, 1998]. Linked Open Data has been recognized as a valuable source for general information on data mining and knowledge graphs are a method for formalizing this knowledge [Ristoski, 2016]. This is a method for extracting conceptual knowledge from multi-relational data. Information extraction is part of the field of study called data mining [Manning, 2008], information that can be related to each other can be studied through multi-relational data mining (MRDM) methods [Džeroski, 2003] that deal with multi-contextual data.
The RCA method is not limited to extracting knowledge from separate contexts: it aims to express knowledge by interoperating the semantics of different contexts, i.e., in addition to extracting knowledge from a specific context, data contained in other contexts are used to enrich knowledge extraction.

Faced with this challenge, scientific locks concern:
- The **lack of formalization** (in other mathematical terms) of the agglomeration of information in the models of systems and information systems that emerge from them, as well as the definition of the semantics of the concepts and relationships they implement, to ensure their common understanding, and to facilitate their interoperation by minimizing semantic losses.
- Using tools of an algebraic and/or geometric nature (lattice theory, category theory, homological algebra) **in the context of relational concept analysis**, which is a recent approach that has not yet been fully developed (even from a mathematical point of view).

3.1.3 Research project justification

The purpose of this section is to justify my project by a positioning in relation to the laboratory and then to the national and international communities.

3.1.3.1 Local context

The project I wish to defend is in line with the project defended by the ISET department of CRAN and its S&O2I project which aims to deal with Industry 4.0 issues.

- the formalization and extraction of knowledge produced from databases generated by intelligent products and objects and their interactions with production systems.

The proposed research project can easily find more points of collaboration with another project of the ISET department: Operational safety, PHM and maintenance of complex systems (SdF-PHM2).
The project is complementary with some research actions of some researchers who have recently passed their qualification to lead research (HDR).

Hind Bril El Hauzi [Bril, 2017] works on intelligent piloting architectures of production systems through the use of simulation models. With whom I will work on the ANR PRCE ISOBIM.

Alexis Aubry [Aubry, 2018] works on the definition of models that allow pro-active and reactive programming in order to take into account the internal and external disturbances of the system.

William Derigent [Derigent, 2018] works on aspects of data models and the management of this data for holonic systems. Its objective is to define and implement the intelligent product from the point of view of information transmission and technology.

Philippe Thomas [Thomas, 2014] works on the use of data for model extraction to drive production systems using the formalisms of neural networks and Petri dishes.

3.1.3.2 Regional context

The Grand Est region presents its regional "Industry of the Future" plan to support local businesses in their efforts to push them towards digitization. It proposes a fully funded diagnostic to SMEs wishing to integrate the concepts of industry 4.0 within their businesses.
I have been collaborating for two years with the Epinal Chamber of Commerce and Industry and its incubator the Quai Alpha. In June 2020, with great teamwork coordinated by me, the incubator became the first Digital Innovation Hub - Fully Operational of the Grand Est. This has allowed a wider diffusion of scientific culture towards regional enterprises.

At the level of the University of Lorraine, the "Lorraine Université d'Excellence" project (in response to the call for a national IDEX-ISITE project) proposes a set of programs to finance research work.

3.1.3.3 National context

The issue of Industry 4.0 is at the heart of the interests of the French scientific community as can be seen from the outcome of the CNIS (Cercle de confiance Numérique des Industries Stratégiques) workshop on Industry 4.0 (Industry 4.0 and IoT: vision and prospects). Among the issues identified in the Montaigne Institute's report on the Industry of the Future published at the end of 2018 are the following:

- the set of physical sensors placed on equipment or products, enabling data to be collected from factories at affordable costs and production parameters to be shared in digital format for optimization purposes.

- The use of advanced algorithmic techniques, some of which are based on artificial intelligence, making it possible to exploit very large quantities of data rapidly. They optimize machine performance by implementing predictive maintenance, and improve production quality through process control, by identifying correlations between multiple production parameters.

My research project is completely in line with these objectives.

In the ANRT 2020 calls for proposals, my project is completely in line with the "Factory of the future: People, organization, technologies" and with the axe 5.2 - CE23 - Artificial Intelligence. The CE23 (evaluation committee definition) expects research in artificial intelligence, in the broadest sense of the term, and knowledge extraction, formalization and management.

3.1.3.4 International context

At the international level, my research project is fully in line with the issues considered by TC5.3 of IFAC and WG12.1 and WG12.6 of IFIP of which I am a member.
The issue of Factories of the future and Industry 4.0 at European level is a cornerstone of the next funding program Horizon Europe 2021-2027.

3.2 Teaching project

Since my arrival at the University of Lorraine in 2010 as Post-doc and later as ATER (Temporary Teaching and Research member) I immediately came into contact with the French teaching system. The first two academic years, 2011 and 2012, I spent them in

Telecom Nancy School of Engineering teaching many disciplines related to Databases. Since 2012 I started my activities in the QLIO department of IUT Hubert Curien d'Epinal managing all the modules related to IT and data management in the company.
In 2017 I made the accreditation to the HCERES of the E-Commerce and Digital Marketing Professional Licence in order to manage it in the following years.
Since 2012 I have never abandoned teaching at the Telecom Nancy school of computer engineering as my interest in being able to stay in relation with students who specialize in computer science is very high.

3.2.1 Three axes of my teaching project

Having the possibility to teach at all levels of the university cycle (DUT, Licence and Master) and in addition to being able to range from a department with a rather industrial vocation to a school of computer engineering, I can structure my future teaching project focusing on industrial and theoretical aspects at the same time.

3.2.1.1 Industry 4.0 aspects:

- The introduction of Industry 4.0 concepts into the QLIO department's university degree. Together with the head of the department we asked for, and obtained, funding of 500,000 euros for the construction of a lean 4.0 atelier. This is allowing us, with the support of the whole pedagogical team, to structure an innovative didactic path. It will be for the students, at the same time, a concrete approach that will solidly lay the foundations on all the theoretical aspects of enterprise management. As far as my contribution is concerned, I have structured, in the 4 modules that I teach during the two years of the diploma, a path that starts from the understanding of basic computer science with the generation of data to the structuring of information and finally to the formalisation of knowledge related to the enterprise systems.

3.2.1.2 Knowledge creation and formalisation aspects:

- The deepening of the methods of knowledge creation and formalization, especially the teaching of methods of ontologies creation, in the Licence E-Commerce and Digital Marketing. The objective of the training course is the creation of future superior technicians who will have specific skills in the management of complex projects related to e-commerce. They must be able to synthesise, automatically, various sources of information and link them to the existing knowledge within the projects they are dealing with.

- The deepening of issues related both to the formalisation of knowledge and Industry 4.0 in the teachings I manage in the Telecom Nancy Engineering School specifically:

    - Enterprise 4.0: where I teach CPS programming and how to manage the data flow generated by smart sensors to create new knowledge formalized in taxonomies that can be used in the future. In addition, in the same module I teach how to

    - Artificial Intelligence: where, in collaboration with several teachers, we help students to understand the basic algorithms to understand the discipline itself and push them to search for new solutions related to the world of research.

    - Integrated Enterprise Management: in this module I introduce students to the knowledge of methods of model creation and model validation against specific software such as ERP.

3.3 General conclusion:

To date, the editing of this document gives me the opportunity to make an explicit point of my work, the concepts and models I have been able to use. My research experience, enriched by the various collaborations I have had in the supervision of PhD theses, has allowed me to develop a knowledge base in knowledge management and industrial organisation. This base, which has generic potential, has been put at the service of the paradigm of production control systems. From my training at the CRAN laboratory alongside people such as Professors Hervé Panetto, Benoit Jung and Hind Bril El Haouzi, I have maintained a systemic culture with its different arguments.
Thus, since 2012, the triumvirate composed of knowledge, paradigm and culture has been the guiding thread of my work. The observation of problems arising from industrial reality has led me to explore classes of problems in the literature and to use modelling paradigms of both computer and production engineering such as IMS (Intelligent Manufacturing System) and computer engineering such as MAS (Multi-Agent System), IDM (Model Driven Engineering) and IA (Ontologies, Machine learning, FCA, RCA...). The tools of the digital factory (simulation and 3D modelling) have allowed me to identify solutions to the industrial problems posed. The new frontiers, both technological and scientific, have led me to reconsider the role of knowledge in all the systems considered and to try to make it as explicit as possible. Therefore, the main scientific challenge of my research project will be to define methodologies for the clarification and formalisation of knowledge based on intelligent/connected objects in the industry of the future. The original idea of this work lies in the use of deductive methods based on explicit knowledge of algorithms applied to data from CPPS (cybernetic-physical production systems) and on the knowledge of domain experts to identify behaviour patterns.

List of the author's publications

[Cestari et al, 2014a] José Marcelo Cestari, Eduardo Rocha Loures, Eduardo Santos, Yongxin Liao, Hervé Panetto, Mario Lezoche. An overview of attributes characterization for interoperability assessment from the public administration perspective. Third Workshop on Industrial and Business Applications of Semantic Technologies and Knowledge-based information systems, INBAST 2014, Oct 2014, Amantea, Italy. pp.329-338, ⟨10.1007/978-3-662-45550-0_33⟩. ⟨hal-01074447⟩

[Cestari et al, 2014b] José Marcelo Cestari, Mario Lezoche, Eduardo Rocha Loures, Hervé Panetto, Eduardo Portela Santos. A method for e-government concepts interoperability assessment. 4th International Conference on Information Society and Technology, ICIST 2014, Mar 2014, Kopaonik, Serbia. pp.79-85. ⟨hal-00993431⟩

[Dassisti et al, 2017] Michele Dassisti, Hervé Panetto, Mario Lezoche, Pasquale Merla, Concetta Semeraro, et al.. Industry 4.0 paradigm: The viewpoint of the small and medium enterprises. 7th International Conference on Information Society and Technology, ICIST 2017, Mar 2017, Kopaonik, Serbia. pp.50-54. ⟨hal-01526397⟩

[De Nicola et al, 2007] Antonio de Nicola, Mario Lezoche, Michele Missikoff. An Ontological Approach to Business Process Modeling. 3th Indian International Conference on Artificial Intelligence 2007, Dec 2007, India. pp.ISBN 978-0-9727412-2-4. ⟨hal-00656686⟩

[Detro et al, 2016] Silvana Pereira Detro, Dmitry Morozov, Mario Lezoche, Hervé Panetto, Eduardo Portela Santos, et al.. Enhancing semantic interoperability in healthcare using semantic process mining. 6th International Conference on Information Society and Techology, ICIST 2016, Feb 2016, Kopaonik, Serbia. pp.80-85. ⟨hal-01298125⟩

[Detro et al, 2017a] Silvana Pereira Detro, Eduardo Portela Santos, Hervé Panetto, Eduardo Rocha Loures, Mario Lezoche. Configuring process variants through semantic reasoning in systems engineering. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2017, 20 (4), pp.36-39. ⟨10.1002/inst.12179⟩. ⟨hal-01673070⟩

[Detro et al, 2017b] Silvana Detro, Eduardo Santos, Hervé Panetto, Eduardo Freitas Rocha Loures, Mario Lezoche. Managing business process variability through process mining and semantic reasoning: an application in healthcare. 18th Working

Conference on Virtual Enterprises, PRO-VE 2017, Sep 2017, Vicenza, Italy. pp.333-340, ⟨10.1007/978-3-319-65151-4_31⟩. ⟨hal-01674899⟩

[Detro et al, 2018] Silvana Pereira Detro, Eduardo Portela Santos, Hervé Panetto, Eduardo Loures de Freitas, Mario Lezoche. Applying process mining and semantic reasoning for process model customization in healthcare. Enterprise Information Systems, Taylor & Francis, 2018, ⟨10.1080/17517575.2019.1632382⟩. ⟨hal-02155320⟩

[Eslami et al, 2018] Yasamin Eslami, Michele Dassisti, Hervé Panetto, Mario Lezoche. Sustainability assessment of manufacturing organizations based on indicator sets: A formal concept analysis. OTM/IFAC/IFIP International Workshop on Enterprise Integration, Interoperability and Networking, EI2N 2018, Oct 2018, Valletta, Malta. pp.36-44, ⟨10.1007/978-3-030-11683-5_4⟩. ⟨hal-02047367⟩

[Eslami et al, 2019] Yasamin Eslami, Michele Dassisti, Mario Lezoche, Hervé Panetto. A survey on sustainability in manufacturing organisations: dimensions and future insights. International Journal of Production Research, Taylor & Francis, 2019, 57 (15-16), pp.5194-5214. ⟨10.1080/00207543.2018.1544723⟩. ⟨hal-01911366⟩

[Eslami et al, 2020] Yasamin Eslami, Mario Lezoche, Hervé Panetto, Michele Dassisti. On analysing sustainability assessment in manufacturing organizations: A survey. International Journal of Production Research, Taylor & Francis, 2020, ⟨10.1080/00207543.2020.1755066⟩. ⟨hal-02524117⟩

[Leal et al, 2016a] Gabriel Leal, Wided Guédria, Hervé Panetto, Erik Proper, Mario Lezoche. Using formal measures to improve maturity model assessment for conceptual interoperability. 11th OTM/IFAC Enterprise Integration, Interoperability and Networking Workshop, EI2N, OM, Oct 2016, Kallithea, Rhodes, Greece. pp.47-56, ⟨10.1007/978-3-319-55961-2_5⟩. ⟨hal-01387780⟩

[Leal et al, 2016b] Gabriel Leal, Wided Guedria, Hervé Panetto, Mario Lezoche. Towards a comparative analysis of interoperability assessment approaches for collaborative enterprise systems. 23rd IPSE International Conference on Transdisciplinary Engineering, Oct 2016, Curitiba, Brazil. pp.45-54, ⟨10.3233/978-1-61499-703-0-45⟩. ⟨hal-01376442⟩

[Lezoche et al, 2008a] Mario Lezoche, Michele Missikoff, Leonardo Tininini. Business Process Evolution: a Rule-based Approach. 9th Workshop on Business Process Modeling, Development, and Support (BPMDS'08), Jun 2008, Montpellier, France. pp.ISSN 1613-0073 vol-335. ⟨hal-00656687⟩

[Lezoche et al, 2008b] Mario Lezoche, Antonio de Nicola, Tania Di Mascio, Francesco Taglino. Semantic Lifting of Business Process Models. 12th Enterprise Distributed Object Computing Conference Workshops, Sep 2008, Munich, Germany. pp.ISBN: 978-0-7695-3720-7. ⟨hal-00656690⟩

[Lezoche et al, 2009] Mario Lezoche, Taglino Francesco. Business Process Driven Solutions for Innovative Enterprise Information Systems. Information Systems: People, Organizations, Institutions, and Technologies, Physica Verlag Heidelberg, pp.Part 7, 407-414, 2009, ⟨10.1007/978-3-7908-2148-2_47⟩. ⟨hal-00656681⟩

[Lezoche et al, 2011] Mario Lezoche, Hervé Panetto, Alexis Aubry. Conceptualisation approach for cooperative information systems interoperability. 13th International Conference on Enterprise Information Systems, ICEIS 2011, Jun 2011, Beijing, China. pp.101-110, ⟨10.5220/0003508401010110⟩. ⟨hal-00583866⟩

[Lezoche et al, 2012a] Mario Lezoche, Alexis Aubry, Hervé Panetto. Formal Fact-Oriented model transformations for Cooperative Information Systems semantic conceptualisation. Enterprise Information Systems. Lecture Notes in Business Information Processing, 2012, LNBIP 102, pp.117-131. ⟨10.1007/978-3-642-29958-2_8⟩. ⟨hal-00660047⟩

[Lezoche et al, 2012b] Mario Lezoche, Esma Yahia, Alexis Aubry, Hervé Panetto, Milan Zdravković. Conceptualising and structuring semantics in Cooperative Enterprise Information Systems Models. Computers in Industry, Elsevier, 2012, 63 (8), pp.775-787. ⟨10.1016/j.compind.2012.08.006⟩. ⟨hal-00722419⟩

[Lezoche et al, 2017] Mario Lezoche, Hervé Panetto. New challenges and Advances in MBSE in French Universities. France. INSIGHT - International Council on Systems Engineering (INCOSE), 20 (4), Willey, pp.8-10, 2017, INSIGHT, 2156-4868. ⟨10.1002/inst.12172⟩. ⟨hal-01673066⟩

[Lezoche et al, 2019] Mario Lezoche, Hervé Panetto. Cyber-Physical Systems, a new formal paradigm to model redundancy and resiliency. Enterprise Information Systems, Taylor & Francis, 2019, ⟨10.1080/17517575.2018.1536807⟩. ⟨hal-01895093⟩

[Lezoche et al., 2020] Mario Lezoche, Jorge Hernandez, Maria del Mar Alemany Diaz, Hervé Panetto, Janusz Kacprzyk. Agri-food 4.0: a survey of the supply chains and technologies for the future agriculture. Computers in Industry, Elsevier, 2020, 117:103187, ⟨10.1016/j.compind.2020.103187⟩. ⟨hal-02395411⟩

[Liao et al, 2011a] Yongxin Liao, Mario Lezoche, Hervé Panetto, Nacer Boudjlida. Why, Where and How to use Semantic Annotation for Systems Interoperability. 1st

UNITE Doctoral Symposium, Jun 2011, Bucarest, Romania. pp.71-78. ⟨hal-00597903⟩

[Liao et al, 2011b] Yongxin Liao, Mario Lezoche, Hervé Panetto, Nacer Boudjlida. Semantic Annotation Model Definition for Systems Interoperability. OTM 2011 Workshops 2011 - 6th International Workshop on Enterprise Integration, Interoperability and Networking (EI2N), Oct 2011, Hersonissos, Crete, Greece. pp.61-70, ⟨10.1007/978-3-642-25126-9_14⟩. ⟨hal-00627909⟩

[Liao et al, 2012] Yongxin Liao, Mario Lezoche, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. Formalization of Semantic Annotation for Systems Interoperability in a PLM environment. OTM Federated conferences and worlshops, 2nd Workshop on Industrial and Business Applications of Semantic Web Technologies (INBAST), Sep 2012, Rome, Italy. pp.207-218, ⟨10.1007/978-3-642-33618-8-207⟩. ⟨hal-00722740⟩

[Liao et al, 2013] Yongxin Liao, Mario Lezoche, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. Semantic enrichment of models to assist knowledge management in a PLM environment. 21st International Conference on Cooperative Information Systems CoopIS 2013, Sep 2013, Graz, Austria. pp.267-274. ⟨hal-00850145⟩

[Liao et al, 2014a] Yongxin Liao, Mario Lezoche, Eduardo Rocha Loures, Hervé Panetto, Nacer Boudjlida. A semantic annotation framework to assist the knowledge interoperability along a product life cycle. Advanced Materials Research, Trans Tech Publications, 2014, 945-949, pp.424-429. ⟨10.4028/www.scientific.net/AMR.945-949.424⟩. ⟨hal-01026591⟩

[Liao et al, 2014b] Yongxin Liao, Mario Lezoche, Hervé Panetto, Nacer Boudjlida, Eduardo Rocha Loures. Formal semantic annotations for models interoperability in a PLM environment. 19th IFAC World Congress, IFAC'14, International Federation of Automatic Control, Aug 2014, Cape Town, South Africa. pp.2382-2393, ⟨10.3182/20140824-6-ZA-1003.02551⟩. ⟨hal-01082576⟩

[Liao et al, 2015] Yongxin Liao, Mario Lezoche, Hervé Panetto, Nacer Boudjlida, Eduardo Rocha Loures. Semantic annotation for knowledge explicitation in a product lifecycle management context: a survey. Computers in Industry, Elsevier, 2015, 71, pp.24-34. ⟨10.1016/j.compind.2015.03.005⟩. ⟨hal-01123854⟩

[Liao et al, 2016] Yongxin Liao, Mario Lezoche, Hervé Panetto, Nacer Boudjlida. Semantic annotations for semantic interoperability in a product lifecycle management context. International Journal of Production Research, Taylor & Francis, 2016, 54 (18), pp.5534-5553. ⟨10.1080/00207543.2016.1165875⟩. ⟨hal-01286475⟩

[Liu et al, 2019] Shaofeng Liu, Guoqing Zhao, Huilan Chen, Alejandro Fernandez, Diego Torres, Mario Lezoche, et al.. Knowledge mobilisation crossing boundaries: a multi-perspective framework for agri-food value chains. 6th Model-IT International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain, Jun 2019, Molfetta, Italy. ⟨hal-02191439⟩

[Monnin et al, 2017] Pierre Monnin, Mario Lezoche, Amedeo Napoli, Adrien Coulet. Using formal concept analysis for checking the structure of an ontology in LOD: the example of DBpedia. 23rd International Symposium on Methodologies for Intelligent Systems, ISMIS 2017, Warsaw university of technology, Jun 2017, Warsaw, Poland. pp.674-683, ⟨10.1007/978-3-319-60438-1_66⟩. ⟨hal-01511909⟩

[Morozov et al, 2015] Dmitry Morozov, Sergei O. Kuznetsov, Mario Lezoche, Hervé Panetto. Formal methods for process knowledge extraction. 10ème Colloque sur la Modélisation des Systèmes Réactifs, MSR 2015, Nov 2015, Nancy, France. ⟨hal-01224489⟩

[Morozov et al, 2017] Dmitry Morozov, Mario Lezoche, Hervé Panetto. FCA modelling for CPS interoperability optimization in Industry 4.0. 7th International Conference on Information Society and Technology, ICIST 2017, Mar 2017, Kopaonik, Serbia. pp.55-59. ⟨hal-01526383⟩

[Morozov et al, 2018] Dmitry Morozov, Mario Lezoche, Hervé Panetto. Multi-paradigm modelling of Cyber-Physical Systems. 16th IFAC Symposium on Information Control Problems in Manufacturing, INCOM 2018, Jun 2018, Bergamo, Italy. pp.1385-1390, ⟨10.1016/j.ifacol.2018.08.334⟩. ⟨hal-01813386⟩

[Panetto et al, 2020] Hervé Panetto, Mario Lezoche, Jorge Hernandez, Maria del Mar Eva Alemany Diaz, Janusz Kacprzyk. Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. Computers in Industry, Elsevier, 2020, 116:103188, ⟨10.1016/j.compind.2020.103188⟩. ⟨hal-02450378⟩

[Semeraro et al, 2019] Concetta Semeraro, Hervé Panetto, Mario Lezoche, Michele Dassisti, Stefano Cafagna. A monitoring strategy for industry 4.0: Master italy s.r.l case study. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2019, Systems engineering research at French Universities, 22 (4), pp.20-22. ⟨10.1002/inst.12269⟩. ⟨hal-02423272⟩

[Semeraro et al, 2019] 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⟩

[Taglino et al, 2008] Francesco Taglino, Mario Lezoche. Business Process driven solutions for innovative enterprise information systems. itAIS 2008 - 5th Conference of the Italian Chapter of AIS, Dec 2008, Paris, France. pp.ISBN:978-88-6105-076-1. ⟨hal-00656678⟩

[Uva et al, 2016] Giuseppina Uva, Michele Dassisti, Francesco Iannone, Franco Maddalena, Michele Ruta, Mario Lezoche, et al.. Modelling framework for sustainable co-management of multi-purpose exhibition systems: the “Fiera del Levante” case. International High-Performance Built Environment Conference, iHBE 2016 (A Sustainable Built Environment Conference 2016 Series - SBE16), Nov 2016, Sydney, Australia. pp.812-821, ⟨10.1016/j.proeng.2017.04.242⟩. ⟨hal-01637828⟩

[Wajnberg et al, 2017] Mickael Wajnberg, Mario Lezoche, Blondin Alexandre Massé, Petko Valtchev, Hervé Panetto. Complex system tacit knowledge extraction trough a formal method. INSIGHT - International Council on Systems Engineering (INCOSE), Wiley, 2017, 20 (4), pp.23-26. ⟨10.1002/inst.12176⟩. ⟨hal-01673069⟩

[Wajnberg et al, 2018] Mickael Wajnberg, Mario Lezoche, Alexandre Blondin Masse, Petko Valtchev, Hervé Panetto, et al.. Semantic interoperability of large systems through a formal method: Relational Concept Analysis. 16th IFAC Symposium on Information Control Problems in Manufacturing, INCOM 2018, Jun 2018, Bergamo, Italy. pp.1397-1402, ⟨10.1016/j.ifacol.2018.08.330⟩. ⟨hal-01813398⟩

[Wajnberg et al, 2019a] Mickael Wajnberg, Petko Valtchev, Mario Lezoche, Hervé Panetto, Alexandre Blondin Masse. Mining process factor causality links with multi-relational associations. 10th International Conference on Knowledge Capture, K-CAP'19, Nov 2019, Marina Del Rey, CA, United States. pp.263-266, ⟨10.1145/3360901.3364446⟩. ⟨hal-02377662v2⟩

[Wajnberg et al, 2019b] Mickael Wajnberg, Petko Valtchev, Mario Lezoche, Alexandre Blondin Masse, Hervé Panetto. Concept analysis-based association mining from linked data: A case in industrial decision making. 2nd International Workshop on Data meets Applied Ontologies in Open Science and Innovation, DAO-SI 2019, Sep 2019, Gratz, Austria. ⟨hal-02455243⟩

[Yahia et al, 2011a] Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Semantics enactment in Enterprise Information Systems. 18th IFAC World Congress, IFAC WC'2011, Aug 2011, Milan, Italy. pp.13064-13073, ⟨10.3182/20110828-6-IT-1002.03597⟩. ⟨hal-00582507⟩

[Yahia et al, 2011b] Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Extraction de la sémantique dans les modèles de systèmes d'information d'entreprises collaboratifs. Journée Nationale du GT Easy-DIM, Apr 2011, Lyon, France. ⟨hal-00586779⟩

[Yahia et al, 2012a] Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Semantics enactment for interoperability assessment in Enterprise Information Systems. Annual Reviews in Control, Elsevier, 2012, 36 (1), pp.101-117. ⟨10.1016/j.arcontrol.2012.03.008⟩. ⟨hal-00671856⟩

[Yahia et al, 2012b] Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Extraction de la structure de la sémantique dans les modèles de système d'information d'entreprises collaboratives. Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information, Lavoisier, 2012, 17 (4), pp.49-77. ⟨10.3166/ISI.17.4.49-77⟩. ⟨hal-00722247⟩

[Zdravković et al, 2011] Milan Zdravković, Miroslav Trajanović, Hervé Panetto, Alexis Aubry, Mario Lezoche. Ontology-based supply chain process configuration. 34th International Conference on Production Engineering, ICPE 2011, Sep 2011, Nis, Serbia. pp.399-402. ⟨hal-00627531⟩

[Zdravković et al, 2016] Milan Zdravković, Miroslav Trajanović, João Sarraipa, Ricardo Jardim-Gonçalves, Mario Lezoche, Alexis Aubry, Hervé Panetto. Survey of Internet-of-Things platforms. 6th International Conference on Information Society and Techology, ICIST 2016, Feb 2016, Kopaonik, Serbia. pp.216-220. ⟨hal-01298141⟩

Bibliography

[Abubakar, 2019] Abubakar, A. M., Elrehail, H., Alatailat, M. A., & Elçi, A. (2019). Knowledge management, decision-making style and organizational performance. Journal of Innovation & Knowledge, 4(2), 104–114.

[Ackoff, 1989] Ackoff, Russell L. 1989. 'From Data to Wisdom'. Journal of Applied Systems Analysis 16 (1): 3–9.

[Adamson, 2017] Adamson, G., Wang, L., & Moore, P. (2017). Feature-based control and information framework for adaptive and distributed manufacturing in cyber physical systems. Journal of Manufacturing Systems, 43, 305–315.

[Agrawal, 1996] Agrawal, R. et al. (1996). Fast discovery of association rules. Advances in knowledge discovery and data mining, 12(1), 307–328.

[Ameri, 2005] Ameri, F., Dutta, D.: Product lifecycle management: closing the knowledge loops. Comput.-Aided Des. Appl. 2, 577–590 (2005).

[Anderson, 2012] Anderson, C. (2012). Makers: The new industrial revolution. New York: Crown Business Books.

[Andréka, 1998] Andréka, H., Németi, I. et van Benthem, J. (1998). Modal languages and bounded fragments of predicate logic. Journal of philosophical logic, 27(3), 217–274.

[Andrews, 2009] Andrews, Simon. 2009. 'In-Close, a Fast Algorithm for Computing Formal Concepts'.

[Arsene, 2015] Arsene, O., Dumitrache, I. and Mihu, I., Expert system for medicine diagnosis using software agents. Expert Systems with Applications, 42(4), pp.1825-1834, 2015.

[Aspray, 1985] Aspray, W. F. 1985. The Scientific Conceptualization of Information: A Survey. Annals of the History of Computing, 7: 117-40. IEEE

[Aubry, 2018] Aubry A. Systèmes industriels soumis à de la variabilité : vers un pilotage sous perturbations pour les systèmes de production holoniques. Automatique / Robotique. Université de Lorraine, 2018. ⟨tel-01871356⟩

[Baader, 2007] Baader F. et al. Completing description logic knowledge bases using formal concept analysis. In IJCAI 2007, pages 230–235, 2007.

[Barker, 1990] Barker, R., CASE* Method: Entity Relationship Modelling, Addison-Wesley, Wokingham, England, 1990

[Bititci, 2004] Bititci, U.S., V. Martinez, P. Albores, et J. Parung. «Creating and managing value in collaborative networks.» Édité par Emerald Group Publishing Limited. International Journal of Physical Distribution and Logistics Management 34, n° 3/4 (2004): 251--268.

[Borst, 1997] Borst, W.N. Construction of engineering ontologies for knowledge sharing and reuse. Universiteit Twente, 1997.

[Boudjlida, 2008] Boudjlida, N., Panetto, H.: Annotation of enterprise models for interoperability purposes. In: IEEE (ed.) IEEE International Workshop on Advanced Information Systems for Enterprises, IWAISE'2008. pp. 11–17. , Constantine, Algeria (2008).

[Brachman, 1979] Brachman, R. J. (1979). On the epistemological status of semantic networks. In Associative networks 3–50. Elsevier.

[Bril, 2017] Bril El-Haouzi H. Contribution à la conception et à l'évaluation des architectures de pilotage des systèmes de production adaptables : vers une approche anthropocentrée pour la simulation et le pilotage. Ingénierie assistée par ordinateur. Université de lorraine, 2017. ⟨tel-01875639⟩

[Brownsword, 1998] Brownsword, L., Carney, D., & Oberndorf, T. (1998). The opportunities and complexities of applying commercial-off-the-shelf components. Crosstalk, 11(4), 4-6.

[Buitelaar, 2005] Buitelaar P. et al. Ontology Learning from Text: Methods, Evaluation and Applications, volume 123. IOS Press, 2005.

[Cai, 2005] Cai, D., Shao, Z., He, X., Yan, X. et Han, J. (2005). Community mining from multi-relational networks. Dans European Conference on Principles of Data Mining and Knowledge Discovery, 445–452. Springer.

[Camarinha, 2008] Camarinha-Matos, L., et H. Afsarmanesh. Collaborative Networks: Reference Modeling. Édité par L. Camarinha-Matos and H. Afsarmanesh. Springer-Verlag New York Inc, 2008.

[Camarinha-Matos, 2014] Camarinha-Matos, L. M., & Afsarmanesh, H. (2014, October). Collaborative systems for smart environments: trends and challenges. In Working Conference on Virtual Enterprises (pp. 3-15). Springer, Berlin, Heidelberg.

[Cerbah, 2008] Cerbah, F. (2008). Learning highly structured semantic repositories from relational databases. Dans European Semantic Web Conference, 777–781. Springer.

[Chen, 1976] Chen, P. P. S. (1976). The entity-relationship model—toward a unified view of data. ACM transactions on database systems (TODS), 1(1), 9-36.

[Chen, 2006] Chen, D., Dassisti, M., Elvesaeter, B., Panetto, H., et al., 2006. In DI.2: Enterprise Interoperability Framework and knowledge corpus, Interoperability Research for Networked Enterprises Applications and Software Network of Excellence, n° IST 508-011.

[CIMdata, 2003] CIMdata: PDM to PLM: Growth of An Industry. (2003).

[Czejdo, 1990] Czejdo, B., Elmasri, R., Rusinkiewicz, M. & Embley, D.W. 1990, "A graphical data manipulation language for an extended entity-relationship model", IEEE Computer, March 1990, pp. 26-37.

[D'Aquin, 2011] d'Aquin M. and Motta E. Extracting relevant questions to an rdf dataset using formal concept analysis. In Proceedings of the sixth international conference on Knowledge capture, pages 121–128. ACM, 2011.

[David, 2007] David, J. (2007). Association rule ontology matching approach. International Journal on Semantic Web and Information Systems (IJSWIS), 3(2), 27–49.

[De Sousa, 2018] De Sousa Jabbour, A. B. L., Jabbour, C. J. C., Foropon, C., & Filho, M. G. (2018). When titans meet – Can industry 4.0 revolutionise the environmentally-sustainable manufacturing wave? The role of critical success factors. Technological Forecasting and Social Change, 132 (January), 18–25. https://doi.org/10.1016/j.techfore.2018.01.017.

[Dehaspe, 1999] Dehaspe and L., Toivonen H. Discovery of frequent datalog patterns. Data Mining and knowledge discovery, 3(1):7–36, 1999.

[Dehaspe, 2001] Dehaspe L. and Toivonen H. Discovery of relational association rules. In Relational data mining, pages 189–212. Springer, 2001.

[Derigent, 2018] Derigent W. Modèles de données et mécanismes de gestion de l'information distribuée dans des systèmes holoniques : Application à la matière communicante. Informatique ubiquitaire. Université de Lorraine (Nancy), 2018. ⟨tel-02098190⟩

[Derler, 2013] Derler, P., Lee, E. A., Tripakis, S., & Törngren, M., 2013, Cyber-physical system design contracts. In Proceedings of the ACM/IEEE 4th International Conference on Cyber-Physical Systems (pp. 109-118). ACM.

[Djellali, 2013] Djellali, C. A new data mining system for ontology learning using dynamic time warping alignment as a case. Procedia Computer Science, 21, pp.75-82, 2013.

[Ducq, 2008] Ducq, Y., et D. Chen. «How to measure interoperability: Concept and Approach.» Proceedings of the 14th International Conference on Concurrent Enterprising: ICE2008A- New Wave of Innovation in Collaboration Networks. Lisbon, Portugal, 2008. 583-590.

[Džeroski, 2003] Džeroski, S. (2003). Multi-relational data mining: an introduction. ACM SIGKDD Explorations Newsletter, 5(1), 1–16.

[Elgueder, 2010] Elgueder, J., Cochennec, F., Roucoules, L., Rouhaud, E.: Product–process interface for manufacturing data management as a support for DFM and virtual manufacturing. Int. J. Interact. Des. Manuf. IJIDeM. 4, 251–258 (2010).

[Euzenat, 2001] Euzenat, J., 2001. Towards a principled approach to semantic interoperability. In CEUR Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing, Seattle, USA, August 4-5, , ISSN 1613-0073, Vol. 47., 19-25.

[Fayyad, 1996] Fayyad, U., Piatetsky-Shapiro, G. et Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), 37–37.

[Ferré, 2015] Ferré S. A proposal for extending formal concept analysis to knowledge graphs. In ICFCA 2015, pages 271–286. Springer, 2015.

[Fisher, 2006] Fisher, D. A. (2006). An emergent perspective on interoperation in systems of systems (No. CMU/SEI-2006-TR-003). Carnegie-Mellon Univ Pittsburgh Pa Software Engineering Inst.

[Frank, 2007] Frank, R., Moser, F. et Ester, M. (2007). A method for multi-relational classification using single and multi-feature aggregation functions. Dans European Conference on Principles of Data Mining and Knowledge Discovery, 430–437. Springer.

[Frankel, 2003] Frankel D. S. 2003. Model Driven Architecture: Applying MDA to Enterprise Computing. John Wiley & Sons.

[Galárraga, 2013] Galárraga, L. A., Teflioudi, C., Hose, K. et Suchanek, F. (2013). Amie : association rule mining under incomplete evidence in ontological knowledge bases. Dans Proceedings of the 22nd international conference on World Wide Web, 413–422. ACM.

[Ganter, 1999] Ganter, B. et Wille, R. (1999). Formal concept analysis: mathematical foundations.

[Gašević, 2009] Gašević, D., Djuric, D. and Devedžic, V. Model driven engineering and ontology development. Springer Science & Business Media, 2009.

[Ghobakhloo, 2018] Ghobakhloo, Morteza. 2018. 'The Future of Manufacturing Industry: A Strategic Roadmap toward Industry 4.0'. Journal of Manufacturing Technology Management 29 (6): 910–936.

[Goethals, 2002] Goethals B. and Van den Bussche J. Relational association rules: getting warmer. In Pattern Detection and Discovery, pages 125–139. Springer, 2002.

[Golatowski, 2003] Golatowski F., et al., "Service-Oriented Software Architecture for Sensor Networks", in International Workshop on Mobile Computing (IMC), pp. 93-98. 2003.

[Gruber, 1995] Gruber, T.R., Toward principles for the design of ontologies used for knowledge sharing? International Journal of Human-Computer Studies, vol. 43, no. 5, pp.907-928, 1995.

[Guizzardi, 2015] Guizzardi, G., Wagner, G., Almeida, J.P.A. and Guizzardi, R.S. Towards ontological foundations for conceptual modeling: the unified foundational ontology (UFO) story. Applied ontology, 10(3-4), pp.259-271, 2015.

[Gunes, 2014] Gunes V., Peter S., Givargis T., and Vahid F., A Survey on Concepts, Applications, and Challenges in Cyber-Physical Systems, KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS, December 2014

[Hayes, 1981] Hayes, P. J. (1981). The logic of frames. In Readings in artificial intelligence 451–458. Elsevier.

[Hazman, 2011] Hazman, M., El-Beltagy, S. R. et Rafea, A. (2011). A survey of ontology learning approaches. International Journal of Computer Applications, 22(9), 36–43.

[Hohenstein, 1991] Hohenstein, U. and Engels, G., "Formal semantics of an entity-relationship-based query language, Entity-Relationship Approach: the core of conceptual modelling", Proc. 9th ER conf., ed. H. Kangassalo, Elsevier Science Pub., Amsterdam, 1991

[Horrocks, 2006] Horrocks, I., Kutz, O. et Sattler, U. (2006). The even more irresistible sroiq. Kr, 6, 57–67.

[Huchard, 2002] Huchard, M., Roume, C. et Valtchev, P. (2002). When concepts point at other concepts: the case of uml diagram reconstruction. Dans Proceedings of the 2nd Workshop on Advances in Formal Concept Analysis for Knowledge Discovery in Databases (FCAKDD), 32–43.

[Izza, 2009] Izza, S., "Integration of industrial information systems: from syntactic to semantic integration approaches". Enterprise Information Systems. 3(1), pp. 1-57, Taylor & Francis, 2009

[Jeon, 2012] Jeon J., Chun I., Kim W., 2012, Metamodel-Based CPS Modeling Tool, in Embedded and Multimedia Computing Technology and Service, Springer Netherlands, pp. 285-291

[Józefowska, 2008] Józefowska J. et al. On reducing redundancy in mining relational association rules from the semantic web. In International Conference on Web Reasoning and Rule Systems, pages 205–213. Springer, 2008.

[Klimeš, 2014] Klimeš J., 2014, Using Formal Concept Analysis for Control in Cyber-Physical Systems, 24th DAAAM International Symposium on Intelligent Manufacturing and Automation, Procedia Engineering 69, pp. 1518–1522

[Klingenberg, 2019] Klingenberg, C. O., Viana Borges, M. A., & Valle Antunes Jr., J. A. (2019). Industry 4.0 as a data- driven paradigm: a systematic literature review on technologies. Journal of Manufacturing Technology Management. https://doi.org/10.1108/JMTM-09-2018-0325.

[Koulamas, 2018] Koulamas, Christos, and Athanasios Kalogeras. 2018. 'Cyber-Physical Systems and Digital Twins in the Industrial Internet of Things [Cyber-Physical Systems]'. Computer 51 (11): 95–98.

[Kramer, 2001] Kramer S. et al. Propositionalization approaches to relational data mining. In S. Džeroski and N. Lavracˇ, editors, Relational Data Mining, pages 262–291. Springer, 2001.

[Krogel, 2005] Krogel, M.-A. (2005). On propositionalization for knowledge discovery in relational database

[Lachiche, 2010] Lachiche, N. (2010). Propositionalization, Dans C. Sammut et G. I. Webb (dir.). Encyclopedia of Machine Learning, (p. 812–817). Springer US : Boston, MA

[Liao, 2013] Liao, Y.: Semantic annotations for systems interoperability in a PLM environment. Other. Université de Lorraine, 2013. English. ⟨NNT: 2013LORR0135⟩. ⟨tel-01750128⟩

[Leibniz, 1989] Leibniz, G.W.: The Monadology. In: Loemker, L.E. (ed.) Philosophical Papers and Letters. pp. 643–653. Springer Netherlands, Dordrecht (1989).

[Maedche, 2000] Maedche, A. et Staab, S. (2000). The text-to-onto ontology learning environment. Dans Software Demonstration at ICCS-2000-Eight International Conference on Conceptual Structures, volume 38, 890–930. sn.

[Maier, 1998] Maier, M. W. (1998). Architecting principles for systems-of-systems. Systems Engineering: The Journal of the International Council on Systems Engineering, 1(4), 267-284.

[Manning, 2008] Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. Cambridge university press.

[Mezhuyev, 2013] Mezhuyev V., Samet R., 2013, Geometrical Meta-Metamodel for Cyber-Physical Modelling, Cyberworlds (CW), 2013 International Conference

[Miller, 1998] Miller, E. (1998). An introduction to the resource description framework. Bulletin of the American Society for Information Science and Technology, 25(1), 15-19.

[Min, 2018] Min X., David J.M., and Kim S.H. "The fourth industrial revolution: opportunities and challenges." International journal of financial research 9.2 (2018): 90-95.

[Monostori, 2014] Monostori, L., 2014, Cyber-physical production systems: roots, expectations and R&D challenges. Procedia Cirp, 17: pp.9-13.

[Morel, 2003] Morel, G., Panetto, H., Zaremba, M., & Mayer, F. (2003). Manufacturing enterprise control and management system engineering: paradigms and open issues. Annual reviews in Control, 27(2), 199-209.

[Neches, 1991] Neches, R., Fikes, R. E., Finin, T., Gruber, T., Patil, R., Senator, T., & Swartout, W.R. Enabling technology for knowledge sharing. AI magazine, 12(3), 36-56, 1991.

[Nergiz, 2008] Nergiz, M. E., Clifton, C. et Nergiz, A. E. (2008). Multirelational k-anonymity. IEEE Transactions on Knowledge and Data Engineering, 21(8), 1104–1117.

[O'Leary, 1998] O'Leary, D.E.: Enterprise knowledge management. Computer. 31, 54–61 (1998).

[Oberndorf, 1998] Oberndorf P. COTS and Open Systems. Technical report, PA 15213-3890, Pittsburgh, Pennsylvania, USA: Software Engineering Institute. Carnegie Mellon University, 1998.

[Obrst, 2003] Obrst, L., "Ontologies for semantically interoperable systems". In Proceedings of the 12th International Conference on Information and Knowledge Management. New Orleans, USA, 2003

[Ocampo-Martinez, 2017] Ocampo-Martinez C., Quijano N. Game-theoretical methods in control of engineering systems. IEEE Control Systems Magazine, 37(1):30-32, 2017.

[Oren, 2006] Oren, E., Hinnerk Möller, K., Scerri, S., Handschuh, S., Sintek, M.: What are Semantic Annotations? (2006).

[Panetto, 2007] Panetto H. Towards a Classification Framework for Interoperability of Enterprise Applications International Journal of Computer Integrated Manufacturing, Taylor & Francis, 2007, 20 (8), pp.727-740.

[Panetto, 2008] Panetto H., Molina A. Enterprise Integration and Interoperability in Manufacturing Systems: trends and issues Computers in Industry, Elsevier, 2008, 59 (7), pp.641-646.

[Piatetsky-Shapiro, 1996] Piatetsky-Shapiro, G., Fayyad, U. et Smith, P. (1996). From data mining to knowledge discovery: An overview. Advances in knowledge discovery and data mining, 1, 35.

[Porter, 2015] Porter, M., & Heppelmann, J. (2015). How smart, connected products are transforming companies. Harvard Business Review, 93(10), 1–19.

[Prisecaru, 2016] Prisecaru, P. (2016). Challenges of the Fourth Industrial Revolution. Knowledge Horizons. Economics, 8(1), 57-62.

[Priss, 2006] Priss, U. (2006). Formal concept analysis in information science. Annual review of information science and technology, 40(1), 521-543.

[Ramos, 2011] Ramos, A. L., Ferreira, J. V., & Barceló, J. (2011). Model-based systems engineering: An emerging approach for modern systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(1), 101-111.

[Ristoski, 2016] Ristoski, P., & Paulheim, H. (2016, October). Rdf2vec: Rdf graph embeddings for data mining. In International Semantic Web Conference (pp. 498-514). Springer, Cham.

[Roblek, 2016] Roblek, V., Meško, M., & Krapež, A. (2016). A complex view of Industry 4.0. SAGE Open, 6(2), 1–11.

[Roche, 2003] Roche, C.: ONTOLOGY: A SURVEY. Presented at the 8th Symposium on Automated Systems Based on Human Skill and Knowledge IFAC (2003).

[Roschelle, 1995] Roschelle, Jeremy, and Stephanie D. Teasley. 1995. 'The Construction of Shared Knowledge in Collaborative Problem Solving'. In Computer Supported Collaborative Learning, 69–97. Springer.

[Rouane-Hacene, 2013] Rouane-Hacene, M., Huchard, M., Napoli, A., & Valtchev, P. (2013). Relational concept analysis: mining concept lattices from multi-relational data. Annals of Mathematics and Artificial Intelligence, 67(1), 81-108.

[Rumbaugh, 1991] Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorensen W., Object Oriented Modeling and Design. Prentice Hall, Book Distribution Center, New York, USA, 1991

[Russell, 2002] Russell, S. et Norvig, P. (2002). Artificial intelligence: a modern approach.

[Sabou, 2006] Sabou, M., Lopez, V., Motta, E., Uren, V.: Ontology selection: Ontology evaluation on the real semantic web. (2006).

[Schild, 1991] Schild, K. (1991). A correspondence theory for terminological logics: Preliminary report. Citeseer.

[Schneider, 2018] Schneider, P. (2018). Managerial challenges of Industry 4.0: An empirically backed research agenda for a nascent field. Review of Managerial Science, 12(3), 803–848.

[Sheth, 1998] Sheth, A., 1998. Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics. In M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, editors. In Interoperating Geographic Information Systems, pp. 5– 30. Kluwer.

[Sheth, 2013] Sheth A., Anantharam P., and Henson C., "Physical-Cyber-Social Computing: An Early 21st Century Approach", in IEEE Intelligent Systems, vol. 28, no. 1, pp. 78-82, 2013.

[Son, 2012] Son H.S., Kim W.Y., Kim R.,Y., Min H., 2012, Metamodel Design for Model Transformation from Simulink to ECML in Cyber Physical Systems, in Computer Applications for Graphics, Grid Computing, and Industrial Environment, Springer Berlin Heidelberg, pp. 56-60

[Sowa, 1983] Sowa, J. F. (1983). Conceptual structures : information processing in mind and machine.

[Srinivasan, 1996] Srinivasan, A., Muggleton, S. H., Sternberg, M. J. et King, R. D. (1996). Theories for mutagenicity : A study in first-order and feature-based induction. Artificial Intelligence, 85(1-2), 277–299.

[Tao, 2018] Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. Journal of Manufacturing Systems, 48, 157–169.

[Thomas, 2014] Thomas P., « Exploitation de données pour la modélisation. : De l'évaluation des systèmes de transport aux chaînes logistiques... » Réseau de neurones [cs.NE]. Université de Lorraine, 2014. ⟨tel-01096878⟩

[Tudorache, 2013] Tudorache, T., Nyulas, C., Noy, N.F. and Musen, M.A. WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the web. Semantic web, 4(1), pp.89-99, 2013.

[Unbehauen, 2012] Unbehauen, J., Hellmann, S., Auer, S. et Stadler, C. (2012). Knowledge extraction from structured sources. In Search computing 34–52. Springer.

[Valtchev, 2004] Valtchev, Petko, Rokia Missaoui, and Robert Godin. 2004. 'Formal Concept Analysis for Knowledge Discovery and Data Mining: The New Challenges'. In International Conference on Formal Concept Analysis, 352–371. Springer.

[Venter, 1997] Venter, F. J., G. D. Oosthuizen, and J. D. Roos. 1997. 'Knowledge Discovery in Databases Using Lattices'. Expert Systems with Applications, Selected Papers from the PACES/SPICIS'97 Conference, 13 (4): 259–64. https://doi.org/10.1016/S0957-4174(97)00047-X.

[Vrandevcić, 2009] Vrandevcić, D.: Ontology evaluation. Springer (2009).

[Vyvyan, 2006] Vyvyan, E., 2006. Lexical Concepts, Cognitive Models and Meaning-Construction. In Cognitive Linguistics 17 (4): 491-534.

[W3C, 2012] W3C (2012). Owl 2 web ontology language document overview (second edition). https://www.w3.org/TR/owl2-overview/. Accessed: 2020-02-15.

[Wille, 1982] Wille, Rudolf. 1982. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts, Dans I. Rival (dir.). Ordered Sets: Proceedings of the NATO Advanced Study Institute held at Ban, Canada, August 28 to September 12, 1981, (p. 445–470). Springer Netherlands: Dordrecht

[Wille, 2002] Wille, Rudolf. 2002. 'Why Can Concept Lattices Support Knowledge Discovery in Databases?' Journal of Experimental & Theoretical Artificial Intelligence 14 (2–3): 81–92. https://doi.org/10.1080/09528130210164161.

[Wille, 2009] Wille, R. (2009, May). Restructuring lattice theory: an approach based on hierarchies of concepts. In International Conference on Formal Concept Analysis (pp. 314-339). Springer, Berlin, Heidelberg.

[Williams, 2006] Williams, Graham J., and Simeon J. Simoff. 2006. Data Mining: Theory, Methodology, Techniques, and Applications. Vol. 3755. Springer.

[Yao, 2013] Yao, J. and Gu, M. Conceptology: Using Concept Map for Knowledge Representation and Ontology Construction. JNW, 8(8), pp.1708-1712, 2013.

[Ye, 2003] Ye, N. (2003). The handbook of data mining. CRC Press.

[Zdravković, 2016] Zdravković, M., Trajanović, M., Sarraipa, J., Jardim-Gonçalves, R., Lezoche, M., Aubry, A., & Panetto, H. (2016, February). Survey of Internet-of-Things platforms.

[Zeleny, 1980] Zeleny, M.: Management support systems: towards integrated knowledge management. Management. 1, 7 (1980).

[Zuehlke, 2010] Zuehlke, Detlef. 2010. 'SmartFactory—Towards a Factory-of-Things'. Annual Reviews in Control 34 (1): 129–138.

Résumé

Ce texte présente mes activités de recherche et d'enseignement au laboratoire CRAN et à l'IUT Hubert Curien à Epinal entre 2012 et 2020. Celles-ci s'inscrivent globalement dans la formalisation et l'extraction de connaissances dans les systèmes liés à l'environnement de production intelligent. L'introduction présente ce que sont les systèmes de production et leur évolution jusqu'à l'industrie 4.0. Dans ce contexte, nous montrons comment la gestion des connaissances par des méthodes telles que l'interopérabilité sémantique, la formalisation des connaissances, leur extraction et leur réutilisation sont des actions nécessaires pour le meilleur fonctionnement des systèmes considérés. Enfin, le dernier chapitre présente le projet de recherche et d'enseignement dont l'objectif est la définition de modèles de formalisation et d'extraction des connaissances appliqués à des sources de données hétérogènes dans un contexte d'industrie du futur.

Mots-clés : Industrie du futur, Industrie 4.0, Formalisation de la connaissance, Extraction de la connaissance, Systèmes cyberphysiques.

Abstract

This text presents my research and teaching activities in the CRAN laboratory and at the IUT Hubert Curien in Epinal between 2012 and 2020. These are globally inscribed in the formalisation and extraction of knowledge in systems related to the intelligent production environment. The introduction presents what production systems are and their evolution up to industry 4.0. In this context we show how knowledge management through methods such as semantic interoperability, formalisation of knowledge, its extraction and reuse are necessary actions for the best functioning of the systems under consideration. Finally, the last chapter presents the research and teaching project whose objective is the definition of knowledge formalisation and extraction models applied to heterogeneous data sources in a factory of the future context.

Keywords: Factory of the future, Industry 4.0, Formalisation of knowledge, Knowledge extraction, Cyber physical systems.

I felt, having reached the last page, that my narrative was a symbol of the man I was when I was writing it, and that, in order to write it, I had to be that man, and that, in order to be that man, I had to write that story, and so on and so forth.

Sentí, al llegar a la última página, que mi narración era un símbolo del hombre que era cuando la estaba escribiendo, y que, para escribirla, tenía que ser ese hombre, y que, para ser ese hombre, tenía que escribir esa historia, y así sucesivamente.

Jorge Luis Borges – Aleph