EU H2020 MSCA RISE Project FIRST
“virtual Factories:
Interoperation suppoRting buSiness innovation”

Stephan Boese[1], Giacomo Cabri[2], Norbert Eder[1],
Federica Mandreoli[2], Alexander Lazovik[3], Massimo Mecella[4],
Keith Phalp[5], Paul de Vrieze[5], Lai Xu[5], and Hongnian Yu[5]

[1] GK software, Germany
{SBoese,neder}@gk-software.com
[2] Università di Modena e Reggio Emilia, Italy
{gcabri,fmandreoli}@unimore.it
[3] University of Groningen, The Netherlands
a.lazovik@rug.nl
[4] SAPIENZA Università di Roma, Italy
mecella@dis.uniroma1.it
[5] Bournemouth University, U.K.
{kphalp,pdevrieze,lxu,hyu}@bournemouth.ac.uk

Abstract. FIRST – “virtual Factories: Interoperation suppoRting buSiness innovation”, is a European H2020 project, founded by the RESEARCH AND INNOVATION STAFF EXCHANGE (RISE) Work Programme as part of the Marie Skłodowska-Curie actions. The project concerns with Manufacturing 2.0 and aims at providing the new technology and methodology to describe manufacturing assets; to compose and integrate the existing services into collaborative virtual manufacturing processes; and to deal with evolution of changes. This Chapter provides an overview of the state of the art for the research topics related to the project research objectives, and then it presents the progresses the project achieved up to now towards the implementation of the proposed innovations.

1 Introduction

The manufacturing industry is entering a new era in which new ICT technologies and collaboration applications are integrated with traditional manufacturing practices and processes to increase flexibility in manufacturing, mass customization, increase speed, better quality and to improve productivity.

Virtual factories are key building blocks for Manufacturing 2.0, enabling the creation of new business ecosystems. In itself, the concept of virtual factories is a major expansion upon virtual enterprises in the context of manufacturing, which only integrates collaborative business processes from different enterprises to simulate, model and test different design options, to evaluate performance, thus to save time-to-production [1].

Boese, S., Cabri, G., Eder, N., Mandreoli, F., Lazovik, A., Mecella, M., Phalp, K., de Vrieze, P., Xu, L. and Yu, H.
EU H2020 MSCA RISE Project FIRST - âĂIJvirtual Factories: Interoperation suppoRting buSiness innovationâĂİ.
DOI: 10.5220/0008861700030019
In *OPPORTUNITIES AND CHALLENGES for European Projects (EPS Portugal 2017/2018 2017)*, pages 3-19
ISBN: 978-989-758-361-2

Creating virtual factories requires the integration of product design processes, manufacturing processes, and general collaborative business processes across factories and enterprises. An important aspect of this integration is to ensure straightforward compatibility between the machines, products, processes, related products and services, as well as any descriptions of those.

Virtual factory models need to be created before the real factory is implemented to better explore different design options, evaluate their performance and virtual commission the automation systems thus saving time-to-production [2].

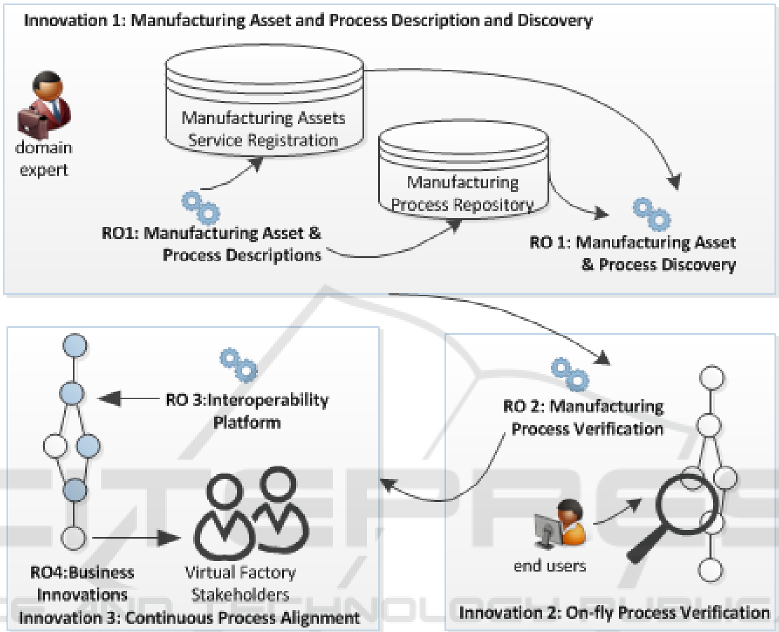

Fig. 1. FIRST research objectives**.**

Providing new technology and methodologies to describe manufacturing assets, to compose and integrate existing services into collaborative virtual manufacturing processes, to deal with evolution of change is the main goal of the FIRST project.

FIRST – "virtual Factories: Interoperation suppoRting buSiness innovation", is a European H2020 project, founded by the RESEARCH AND INNOVATION STAFF EXCHANGE (RISE) Work Programme as part of the Marie Skłodowska-Curie actions. The RISE scheme promotes international and cross-sector collaboration through exchanging research and innovation staff, and sharing knowledge and ideas from research to market (and vice-versa). The project consortium includes five University partners and two industrial partners from Europe and China.

The main innovations the project aim to introduce are depicted in Fig. 1. The related Research Objectives (RO) are the following:

RO1. The design of a new semantic manufacturing asset/service/process description languages and manufacturing asset and (sub-) process discovery methods to enable on-the-fly service-oriented process verification.

RO2. An on-the-fly manufacturing service-oriented process verification method. Human modellers are not perfect in their modelling, or awareness of physical constraints. On-the-fly service-oriented process verification is thus essential to create valid models and to prevent expensive misconfiguration of machinery.

RO3. The design of an interoperability framework providing support for compatibility and evolution to enable global manufacturing collaboration; improving flexibility of existing product design and manufacturing processes; supporting (new) product-service linkages; and improved management of distributed manufacturing assets.

RO4. Integration for seamless matchmaking of business opportunities, co-creation of product innovation, and creation of novel business models. Innovation requires special expertise, domain knowledge of the industry sector, technical knowledge, business models, finances, and markets.

In the rest of this Chapter, we first provide an overview of the state of the art for the research topics related to the project research objectives then we present the progresses the project achieved up to now towards the implementation of the proposed innovations.

2 Overview of Manufacturing Assets and Services Classification and Ontology

This section describes the state of the art as well as the progress that can be envisaged beyond the state of the art, in relation to two major research topics related to the FIRST project.

2.1 Relevant Technologies, Standards and Frameworks of Product Lifecycle Management

Product lifecycle management is the process of dealing with the creation, modification, and exchange of product information through engineering design and manufacture, to service and disposal of manufactured products. In this section, we review the economic and technical aspects of an interoperation framework for product lifecycle management, related standards, technologies, and projects.

STEP (Standard for the Exchange of Product Model Data) ISO 10303. ISO 10303, also known as STEP (Standard for the Exchange of Product Model Data), is an international standard for industrial automation systems and integration of product data representation and exchange. It is made up of various parts that offer standards for specific topics. Part 242:2014 refers to the application protocol for managing model-based 3D engineering (ISO 2014). The standard will be essential to implementing a digital factory based model.

Open Services for Lifecycle Collaboration (OSLC). Open Services for Lifecycle Collaboration (OSLC) is an open community that creates specifications for the integration of tools, such as lifecycle management tools, to ensure their data and workflows are supported in the end-to-end processes. OSLC is based on the W3C linked data (W3C 2015).

Reference Architecture Model for Industry 4.0. Reference Architecture Model for Industry 4.0 [3] defines three dimensions of enterprise system design and introduces the concept of Industry 4.0 components [3]. The RAMI4.0 is essentially focused on the manufacturing process and production facilities; it tries to focus all essential aspects of Industry 4.0. The participants (a field device, a machine, a system, or a whole factory) can be logically classified in the model and relevant Industry 4.0 concepts described and implemented.

The RAMI4.0 3D model includes hierarchy levels, cycle and value stream, and layers. The layers represent the various perspectives from the assets up to the business process, which is most relevant with our existing manufacturing asset/service classification.

Currently RAMI4.0 does not provide detailed, strict indication for standards related to communication or information models. The devices/assets are provided using *Electronic Device Description* (EDD) (also see section 3.1) [4], which includes the device characteristics specification, the business logic and information defining the user interface elements (UID – User Interface Description).

The optional User Interface Plugin (UIP) that defines programmable components based on the Windows Presentation Foundation specifications, to be used for developing UI able to effectively interact with the device.

The *Functional and Information Layer the Field Device Integration (FDI)* [5] specification as integration technology. The FDI is a new specification that aims at overcoming incompatibilities among some manufacturing devices specifications. Essentially the FDI specification defines the format and content of the so-called *FDI package* as a collection of files providing: the device Electronic Device Description (EDD), the optional User Interface Plugin (UIP), and possible optional elements useful to configure, deploy and use the device, e.g. manual, protocol specific files, etc.

An *FDI package* is therefore an effective mean through which a device manufacturer defines which data, functions and user interface elements are available in/for the device.

Semantics for Product Life-cycle Management (PLM) Repositories. OWL-DL is one of the sublanguages of OWL[1]. OWL-DL is the part of OWL Full that fits in the Description Logic framework and is known to have decidable reasoning. In building product lifecycle management repositories, OWL-DL is used to extract knowledge from PLM-CAD (i.e CATIA) into the background ontology automatically, other non-standard parts, i.e. not from CATIA V5 catalogue, manually into the background ontology. OntoDMU is used to import standard parts into concepts of the ontology.

An ontological knowledge base consists of two parts offering different perspectives on the domain. In Fig. 2, the structural information of a domain is characterized through

[1] https://www.w3.org/TR/owl-guide/

its TBox (the terminology). The TBox consists of a set of inclusions between concepts. The ABox (the assertions) contains knowledge about individuals, e.g. a particular car of a given occurrence of a standard part in a CAD model. It can state either that a given named individual (i.e. 'myCar') belongs to a given concept (e.g., that myCar is, in fact, a car) or that two individuals are related by a given property (e.g. that myCar is owned by me).

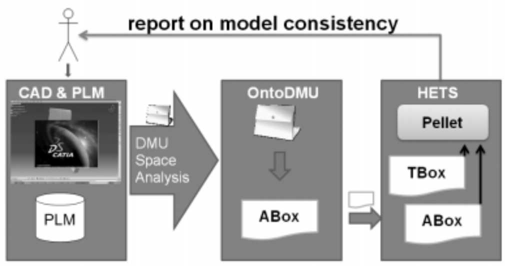

Fig. 2. Ontology based on PLM Repositories [7].

Ontology Mediation for Collaboration of PLM with Product Service Systems (PSS). The PSYMBIOSYS[2] EU Project addresses collisions of design and manufacturing, product and service, knowledge and sentiments, service-oriented and event-driven architectures, as well as business and innovations. Each lifecycle phase covers specific tasks and generates/requires specific information. Ontology mediation is proposed is proposed as a variant of ontology matching since the level of matching can be rather complex.

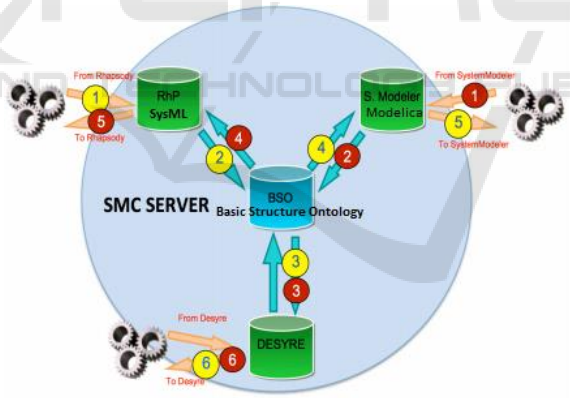

Fig. 3. Ontology Mediation [6].

When matching two different modelling languages, such as Modelica and SysML in, the issue of completeness makes the mapping task impossible. The two languages are significant differences and overlaps. Fig. 3 above presents an ontology mediation approach, which Basic Structure Ontology (BSO) is at the centre, and the mediation

[2] http://www.psymbiosys.eu/

among three different tools was working through three matching sets that connected the common structure ontology which each of the tools: Medelica tool, SysML tool and a 3rd party proprietary tool [6].

Interoperability of Product Lifecycle Management. Integrating among heterogeneous software applications distributed over stakeholders in closed-loop PLM. The capabilities of the Internet of Things are being extended to Cyber-Physical-Systems (CPS), which divide systems into modular and autonomous entities. The systems are able to communicate, to recognize the environment and to make decisions. Different companies with different IT-infrastructures adopt different roles in the product lifecycle.

In order to manage the interoperability of heterogeneous systems throughout the product lifecycle, different approaches could be used [7]:

- Tightly coupled approaches implement federated schema over the systems to be integrated. A single schema is used to define a combined (federated) data model for all involved data sources [7. Any change of the individual system's data models need to be reflected by a corresponding modification of the entire federated schema.
- Object-oriented interoperability approaches are closely related to tightly couple ones. Different types of these approaches are described in (Pitoura, Bukhres, & Elmagarmid, 1995). Object-oriented interoperability approaches use common data models which are a similar problem of dealing with modification of the individual system.
- Loosely coupled interoperability approaches are more suitable to achieving scalable architecture, modular complexity, robust design, supporting outsourcing activities, and integrating third party components. Using Web services for a communication method among different devises, objectives, or databases is one of such loosely coupled interoperability approaches. The semantic meaning of a Web service can be described using OWL (Web Ontology Language). Web services described over third party ontologies [9] are called Semantic Web Services.
- Service Oriented Architecture (SOA) has emerged as the main approach for dealing with the challenge of interoperability of systems in heterogeneous environment [10,11]. SOA offers mechanisms of flexibility and interoperability that allow different technologies to be dynamically integrated, independently of the system's PLM platform in use [12]. Some of standards for PLM using SOA are: *OMG PLM Services* [13] and *OASIS PLCS PLM Web Services* [14].

OMG PLM Services. The current version, PLM Services 2.0 [14], covers a superset of the STEP PDM Schema entities and exposes them as web services. This specification resulted from a project undertaken by an industrial consortium under the umbrella of the ProSTEP iViP Association. Its information model is derived from the latest ISO 10303-214 STEP model (which now includes engineering change management process) by an EXPRESS-X mapping specification and an EXPRESS-to-XMI mapping process. The functional model is derived from the OMG PDM Enablers V1.3. The specification defines a Platform Specific Model (PSM) applicable to the web services implementation defined by a WSDL specification, with a SOAP binding, and an XML

Schema specification. More details on architecting and implementing product information sharing service using the OMG PLM Services can be found in [15].

OASIS PLCS PLM Web Services. Product Life Cycle Support (PLCS) is the phrase used for the STEP standard ISO 10303-239 (Product Life Cycle support (PLCS) Web services V2) [16]. After the initial STEP standard was issued by ISO, a technical committee was formed in the OASIS organization to develop this further. A set of PLCS web services has been developed by a private company (Eurostep) as part of the European Union funded VIVACE project [3]. Eurostep has put this forward on behalf of VIVACE to the OASIS PLCS committee for consideration as the basis for an OASIS PLCS PLM web services standard.

ISA-95/OAGIS SOA in Manufacturing. ISA-95[4] and OAGi are jointly working on standards for manufacturing systems integration. They are actively looking into the suitability of SOA for such integration in manufacturing.

2.2 Manufacturing Assets/Services Classification

Digital Manufacturing Platforms will be fundamental for the development of Industry 4.0 and Connected Smart Factories. They are enabling the provision of services that support manufacturing in a broad sense by aiming at optimising manufacturing from different angles: production efficiency and uptime, quality, speed, flexibility, resource-efficiency, etc. [17]. The available services collect, store, process and deliver data that either describe the manufactured products or are related to the manufacturing processes and assets that make manufacturing happen.

As pointed out in [17], pre-requisites for digital platforms to thrive in a manufacturing environment include the need for agreements on industrial communication interfaces and protocols, common data models and the semantic interoperability of data, and thus on a larger scale, platform inter-communication and inter-operability. The achievement of these objective will allow a boundaryless information flow among the single product lifecycle phases [18] thus enabling an effective, whole-of-life product lifecycle management (PLM). Indeed, the most significant obstacle is that valuable information is not readily shared with other interested parties across the Beginning-of-Life (BoL), Middle-of-Life (MoL), and End-of-Life (EoL) lifecycle phases but it is all too often locked into vertical applications, sometimes called *silos*. Moreover, these objectives are strictly related to the need of achieving the full potential of the Internet of Things in the manufacturing industry. Indeed, without a trusted and secure, open, and unified infrastructure for true interoperability, the parallel development of disparate solutions, technologies, and standards will lead the Internet of Things to become an ever-increasing web of organization and domain-specific intranets.

The EU PROMISE project[5] developed the foundation of the Quantum Lifecycle Management (QML) Technical Architecture to support and encourage the flow of

---

[3] https://cordis.europa.eu/project/rcn/72825_en.html

[4] https://isa-95.com/

[5] The PROMISE Project (2004-2008): A European Union research project funded under the 6th Framework Program (FP6) which focused on information systems for whole-of-life product lifecycle management.

lifecycle data between multiple enterprises throughout the life of an entity and its components. QML was further developed by the Quantum Lifecycle Management (QLM)[6], a Work Group of The Open Group whose members work to establish open, vendor-neutral IT standards and certifications in a variety of subject areas critical to the enterprise. The three main components of QML are the Messaging Interface (MI), the Data Model (DM), and the Data Format (DF) [19]. The Message Interface provides a flexible interface for making and responding to requests for instance-specific information. The Data Model, instead, enables detailed information about each instance of a product to be enriched with "field data"; i.e., detailed information about the usage and changes to each instance during its life. Finally, the Data Format represents, through an XML schema, the structure of the message exchanged between many products and/or systems. The structure of the message is similar to the Data Model schema so that it could be easily recognize by a system QLM DM compatible, thereby automating the data collection.

Various works adopt QLM for manufacturing assets representation and classification. For instance, the paper [20] proposes data synchronization models based upon QLM standards to enable the synchronization of product-related information among various systems, networks, and organizations involved throughout the product lifecycle. These models are implemented and assessed based on two distinct platforms defined in the healthcare and home automation sectors. Främling, Kubler, & Buda [21] describe two implemented applications using QLM messaging, respectively, defined in BoL and between MoL-BoL.

The former is a real case study from the LinkedDesign EU FP7 project, in which different actors work on a production line of car chassis. This process segment involved two robots to transfer the chassis part from machine to machine. The actors involved in the manufacturing plan expressed, on the one hand, the need to check each chassis part throughout the hot stamping process and, on the other hand, the need to define communication strategies adapted to their own needs. Accordingly, scanners are added between each operation for the verification procedure, and QLM messaging is adopted to provide the types of interfaces required by each actor. The latter, instead, involves actors from two distinct PLC phases: 1) In MoL: A user bought a smart fridge and a TV supporting QLM messaging; 2) In BoL: The fridge designer agreed with the user to collect specific fridge information over a certain period of the year (June, July, August) using QLM messaging. Also in this case, the appropriate QLM interfaces regarding each actor have been set up in such a way that the involved actors can get the required information about the smart objects.

In most applications scenarios, taxonomies are usually adopted as common ground for semantic interoperability. Classifying products and services with a common coding scheme facilitates commerce between buyers and sellers and is becoming mandatory in the new era of electronic commerce. Large companies are beginning to code purchases in order to analyse their spending. Nonetheless, most company coding systems today have been very expensive to develop. The effort to implement and maintain these systems usually requires extensive utilization of resources, over an extended period of

[6] http://www.opengroup.org/subjectareas/qlm-work-group

time. Additionally, maintenance is an on-going, and expensive, process. Another problem is that company's suppliers usually don't adhere to the coding schemes of their customers, if any.

Samples of taxonomy including the description and classification of manufacturing assets and services are: eCl@ss, UNSPSC, and MSDL. eCl@ss[7] is an international product classification and description standard for information exchange between customers and their suppliers. It provides classes and properties that can be exploited to standardise procurement, storage, production, and distribution activities, both intra-companies and inter-companies. It is not bound to a specific application field and can be used in different languages. It is compliant to ISO/IEC. It adopts an open architecture that allows the classification system to be adapted to an enterprise's own internal classification scheme, so granting flexibility and standardization at the same time. Thanks to its nature, it can be exploited in the Internet of Things field in order to enable interoperability among devices of different vendors. As of October 2017, there are about 41,000 product classes and 17,000 uniquely described properties which are categorized with only four levels of classification; this enables every product and service to be described with an eight-digit code. One of the aims of eCl@ass is to decrease inefficiencies, so that packaging and distribution take place automatically, relying on the classes and identifier available by the standard. The nature of eCl@ss enables the definition of several aspects in virtual factories.

The United Nations Standard Products and Services Code (UNSPSC)[8] provides an open, global multi-sector standard for efficient, accurate classification of products and services. The UNSPSC was jointly developed by the United Nations Development Programme (UNDP) and Dun & Bradstreet Corporation (D & B) in 1998. It has been managed by GS1 US since 2003. UNSPSC is an efficient, accurate and flexible classification system for achieving company-wide visibility of spend analysis, as well as, enabling procurement to deliver on cost-effectiveness demands and allowing full exploitation of electronic commerce capabilities. Encompassing a five-level hierarchical classification codeset, UNSPSC enables expenditure analysis at grouping levels relevant to the company needs. The codeset can be drilled down or up to see more or less detail as is necessary for business analysis. The UNSPCS classification can be exploited to perform analysis about company spending aspects, to optimize cost-effective procurement, and to exploit electronic commerce capabilities.

The Manufacturing Service Description Language (MSDL) [22] is a formal ontology for describing manufacturing capabilities at various levels of abstraction including the supplier-level, process-level, and machine-level. It covers different concepts like actors, materials, like ceramic and metal, physical resources, tools, and services. Description Logic is used as the knowledge representation formalism of MSDL in order to make it amenable to automatic reasoning. MSDL can be considered an "upper" ontology, in the sense that it provides the basic building blocks required for modeling domain objects and allows ontology users to customize ontology concepts based on their specific needs; this grants flexibility and standardization at the same time. MSDL is composed of two main parts: 1) MSDL core and 2) MSDL extension. MSDL core is

[7] http://www.eclasscontent.com/index.php?language=en&version=7.1
[8] http://www.unspsc.org/

the static and universal part of MSDL that is composed of basic classes for manufacturing service description; MSDL extension is dynamic in nature and includes a collection of taxonomies, sub-classes and instances built by users from different communities based on their specific needs; MSDL extensions drive evolution of MSDL over time.

3 On-the-fly Manufacturing Service-oriented Process Verification

The increasing digital interconnection of people and things, anytime and anywhere, is referred to as hyperconnectivity. Advances in connectivity are already leading to strong development and enhancement of networked service-oriented collaborative organisational structures. Hyperconnectivity, including developments in the areas of Internet of Things and Cyber-Physical Systems, provides new opportunities to manufacturing, value-added services, e-Health and care, crisis/disaster management, logistics, etc. Collaboration and networking are critical to hyperconnected world and intelligent autonomous systems. BU research is thus focused on design effective collaborative processes for Internet scale and verifying correctness of distributed collaborative processes at runtime, as well as deals with interoperability of multi-party collaborative processes at a change environment, which is key enablement of intelligent autonomous systems, such as virtual factory.

To be able to verify distributed collaborative business processes [29], the process description modelling language goes beyond traditional activity level, which into invoked services. Therefore, avoiding conflicts of conditions of invoked service could be checked at control flow level, which is beyond state-of-the-art. Taking recent advantages in Big Data, our research demonstrates the ability to collect large amounts of data and metrics on a large variety of processes, including distributed collaborative processes [30]. To make maximum use of this data, the new runtime verification tools are created and designed to take this data into account, which is the key issues of WP4 of the FIRST project.

Compliance constrains business processes to adhere to rules, standards, laws and regulations. Non-compliance subjects enterprises to litigation and financial fines. Collaborative business processes cross organizational and regional borders implying that internal and cross regional regulations must be complied with. To protect customs' data, European enterprises must comply with the EU data privacy regulation (general data protection regulation - GDPR) and each member state's data protection laws. Compliance verification is thus essential to deploy and implement collaborative business process systems. It ensures that processes are checked for conformance to compliance requirements throughout their life cycle. BU research also looks at checking authorisation compliance among collaborative business processes in the context of virtual factories/enterprises. Security is an important issue in collaborative business processes, in particular for applications that handle sensitive personal information and checking compliance of collaborative business processes in the virtual factory/environment context as well as providing traceable commercial sensitive data distribution for using new technologies, such as blockchain, etc.

The design of an interoperability framework provides support for compatibility and evolution, which is essential for designing intelligent autonomous systems. Building upon the results of semantic asset/service/process description languages and service-oriented distributed collaborative process verification methods, a common schema and schema evolution framework is used to facilitate interoperability on data/information, services and processes respectively. It enables: global (manufacturing) process collaboration; improving flexibility of existing product design and manufacturing processes; supporting (new) product-service linkage; and improved management of distributed manufacturing assets. BU research particularly focus on identifying related concepts of factories of the future and research challenges of interoperability of virtual factory are addressed. Following BU previous research on resilience of SOA collaborative process systems, paper [23] and architecture design of collaborative processes for managing short term, low frequency used collaborative processes; the results of our research provide the good foundation for identifying research requirements for D1.1 of the FIRST project. More specific explanations of each scientific papers are followed.

The internet and pervasive technology like the Internet of Things (i.e. sensors and smart devices) have exponentially increased the scale of data collection and availability. This big data not only challenges the structure of existing enterprise analytics systems but also offer new opportunities to create new knowledge and competitive advantage. Businesses have been exploiting these opportunities by implementing and operating big data analytics capabilities. Social network companies such as Facebook, LinkedIn, Twitter and Video streaming company like Netflix have implemented big data analytics and subsequently published related literatures. However, these use cases did not provide a simplified and coherent big data analytics reference architecture as well as currently, there still remains limited reference architecture of big data analytics. [30] aims to simplify big data analytics by providing reference architecture based on existing four use cases and subsequently verified the reference architecture with Amazon and Google analytics services.

The users of virtual factory are not experts in business process modelling to guarantee the correct collaborative business processes for realising business process execution. To enable automatic execution of business processes, verification is an important step at the business process design stage to avoid crashes or other errors at runtime. Research in business process model verification has yielded a plethora of approaches in form of methods and tools that are based on different technologies like Petri nets family and temporal logic among others. From the literature no report specifically targets and presents a comparative assessment of these approaches based on criteria as one we propose. [29] therefore presents an assessment of the most common verification approaches based on their expressibility, flexibility, suitability and limitations. Furthermore, we look at how big data impacts the business process verification approach in a data-rich world.

[31] proposes architecture of collaborative processes for managing short term, low frequency collaborative processes. A real world case of collaborative processes is used to explain the design and implementation of the cloud-based solution for supporting collaborative business processes. Service improvement of the new solution and computing power costs are also analysed accordingly.

In paper [23], we have proposed resilience analysis perspectives of SOA collaborative process systems, i.e., overall system perspective, individual process model perspective, individual process instance perspective, service perspective, and resource perspective. A real world collaborative process is reviewed for illustrating our resilience analysis. This research contributes to extend SOA collaborative business process management systems with resilience support, not only looking at quantification and identification of resilience factors, but also considering ways of improving the resilience of SOA collaborative process systems through measures at design and runtime.

Concepts and research challenges of interoperability of virtual factory are addressed [32]. We present a comprehensive review on basic concepts of factories of the future, i.e. smart factory, digital factory and virtual factory. The relationships among smart factory, digital and virtual factory are studied. Interoperability of virtual factories is defined. Challenges of interoperability of virtual factories are identified: lack of standards of virtual factories; managing traceability of sensitive data, protected resources and applications or services are critical for forming and using virtual factories; handling multilateral solutions and managing variability of different solutions/virtual factory models are also impact to the usability of the virtual factory. In short, the interoperability of virtual factory related to many newly developed ICT of the hardware and software innovation. An interoperation framework allows evolutional and handling changes, which is crucial for generating and maintaining virtual factories among different industrial sectors.

4 A Reliable Interoperability Architecture for Virtual Factories

One of the key issues in digital factories is to provide, manage and use the different services and data that are connected to the production processes. Manufacturing machines typically provide data about their status and services. These services are usually exploited at the digital factory level together with data and services coming from other departments, such as purchasing and marketing. We face heterogeneous situations: from the one hand, machines are from different vendors and, even if not proprietary, they are likely to adopt different standards and vocabulary, and data are managed by different systems as well; from the other hand, services can be provided at different levels of granularity, from very fine grained one (in terms of functionalities) to very coarse. The role of the digital factory is to *integrate* the different services and data and to combine them in order to make the whole process as efficient and competitive as possible in the achievement of the specific goals.

Another importation issue to be faced is the fact that the process can cover a space wider than the single factory (it supports a supply chain): usually a factory gets the raw material from suppliers and provide products or semi-finished products to customers, through delivery agents, requiring the corresponding services and data to integrate to each other or at least to be able to interact in a *scalable* and *flexible* way.

We propose to achieve this through a general three-layer interoperability framework, i.e., based on *processes, services, and data,* and assists users in the achievement of their objectives through the discovery of service and data flows that best fit the expressed requirements. In the following the three layers are detailed.

Process Space Layer - Goal-oriented Process Specification

The top layer of the proposed architecture deals with the goals and the processes able to achieve such goals. Notably, companies would like to define, on the basis of such goals, specific KPIs – Key Performance Indicators, which qualify the QoS of the production process. Clearly goals and KPIs are defined over many aspects, including the interactions with external companies being part of the process.

Service Space Layer - Dynamic Service Discovery and Composition

Starting from the goals and processes defined in the process layer, services must be dynamically composed to achieve the goal(s). OpenAPIs are exposed by such services in order to control, discover, and compose them in a dynamic way. Rich semantic descriptions of the services should be available in the interoperability platform, in order to support both the discovery of the services and their execution/invocation. The descriptions should include some keywords that identify the context of the service (e.g., “food”, “cooking”), the equipment (e.g., “oven”, “mixer”), the performed operation (e.g., “turn-on”, “speedup”), and the parameters (e.g., “temperature”, “speed”).

With regard to the discovery phase, the semantic description is exploited to search for specific services without knowing their exact name and their syntax a priori. Semantic techniques can be exploited to find synonyms and keywords related to the words searched for in this phase. Searches can be performed either automatically by the process layer, in particular by the orchestration engine enacting processes, or by a human operator acting in the factory, which may be involved when needed (e.g., the adaptation techniques realized in the process layer fail, and a human intervention is needed in order to make the process progress) [24].

But the semantic descriptions can be exploited also in the composition phase. Being the composition dynamic, the platform must not only find but also exploit the needed service in an automatic way or providing an effective support to the human operator. To this purpose, the semantic description of the service parameters is needed in order exploit the meta-services of the data layer to adapt the client service invocation to the server syntax (see next subsection). Some proposals and examples of semantic service descriptions exist, such as in the SAPERE project [25] mentioned later.

The dynamism is useful to handle unexpected situations, often notified by a human operator. Clearly, the platform must also consider *failure* situations, such as an equipment out of work, and so on. These issues require the frequent involvement of humans in the loop in order to deal with them in an effective way.

Data Space Layer - Service-oriented Mapping Discovery and Dynamic Dataspace Alignment

Data are managed and accessed in a data space. The data space must be able to deal with a huge volume of heterogeneous data by autonomous sources and support the different information access needs of the service level. In particular, a large variety of data types should be managed at the dataspace level. According to the level of dynamicity, data can be static such as data available in traditional DBMSs but also highly dynamic like sensor data that are continuously generated. Moreover, it should accommodate data that exhibit various degrees of structures, from tabular data like relational data and CSV data to fully unstructured data like textual data. Finally, it should cope with the very diversified data access modalities sources offer, from low level streaming access to high level data analytics.

To this extent, the data modelling abstraction we adopt to represent the data space is fully decentralized, thereby bridging, on the one hand, existing dataspace models that usually rely on a single mediated view and, on the other hand, P2P approaches for data sharing [26]. The dataspace is therefore a collection of heterogeneous data sources that can be involved in the processes, both in-factory and out-factory. Those data are either describing the manufactured products or the manufacturing processes and assets (material, machine, enterprises, value networks and factory workers) [27]. Each data source has its data access model that describes the kind of managed data, e.g., streaming data vs. static data, and the supported operators. As an example, sensed parameters such as temperature in an oven, temperature in a packing station, etc. are all streaming data needed in the dataspace that can be accessed only through simple windowing operators on the latest values. On the other hand, supplier data can be recorded in a DBMS that offers a rich access model both for On Line Transaction Process (OLTP) operations and On Line Analytical Process (OLAP) operations.

Data representation relies on the graph modelling abstraction. This model is usually adopted to represent information in rich contexts. It employs nodes and labelled edges to represent real world entities, attribute values and relationships among entities.

The main problem the interoperability platform must cope with when dealing with data is data heterogeneity. Indeed, the various services gather data, information and knowledge from sources distributed over different stakeholders and external sources, e.g., the delivery agents and the Web. All these sources are independent, and we argue that a-priori agreements among the distributed sources on data representation and terminology is unlikely in large digital supply chains over several digital factories.

Data heterogeneity can concern different aspects: *(1)* different data sources can represent the same domain using different data structures; *(2)* different data sources can represent the same real-world entity through different data values; *(3)* different sources can provide conflicting data. The first issue is known as *schema heterogeneity* and is usually dealt with through the introduction of mappings. Mappings are declarative specifications describing the relationship between a target data instance and possibly more than one source data instance. The second problem is called *entity resolution* (a.k.a. record linkage or duplicate detection) and consists in identifying (or linking or grouping) different records referring to the same real-world entity. Finally, conflicts can arise because of incomplete data, erroneous data, and out-of-date data. Returning incorrect data in a query result can be misleading and even harmful. This challenge is usually addressed by means of *data fusion* techniques that are able to fuse records on the same real-world entity into a single record and resolve possible conflicts from different data sources.

Traditional approaches that address data heterogeneity propose to first solve schema heterogeneity by setting up a data integration application that offers a uniform interface to the set of data sources. This requires the specification of schema mappings that is a really time- and resource-consuming task entrusted to data curation specialists. This solution has been recognized as a critical bottleneck in large scale deeply heterogeneous and dynamic integration scenarios, as digital factories are. A novel approach is the one where mapping creation and refinement are interactively driven by the information access needs of service flows and the exclusive role of mappings is to contribute to execute service compositions [28]. Hence, we start from a chain of services together with their information needs expressed as inputs and outputs which we attempt to satisfy in the dataspace. We may need to discover new mappings and refine existing mappings induced by composition requirements, to expose the user to the inputs and outputs thereby discovered for their feedback and possibly continued adjustments. Therefore, the service composition induces a data space orchestration that aims at aligning the data space to the specific service goals through the interactive execution of three steps: mapping discovery and selection, service composition simulation, feedback analysis. Mappings that are the outcome of this process can be stored and reused when solving similar service composition tasks.

5 Concluding Remarks

The FIRST project will continue till the end of 2020. In this paper we have presented the initial outcomes of the project, which will be improved and refined over the next years.

References

1. EFFRA. ‘’Factories of the Future 2020 Roadmap Multiannual roadmap for the contractual PPP under Horizon 2020. (2013)
2. Debevec, M., M. Simic, and N. Herakovic. "Virtual factory as an advanced approach for production process optimization" International journal of simulation modelling 13, no. 1 (2014): 66-78.
3. Reference Architecture Model Industrie 4.0 (RAMI4.0). https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2016/januar/GMA_Status_Report__Reference_Archtitecture_Model_Industrie_4.0__RAMI_4.0_/GMA-Status-Report-RAMI-40-July-2015.pdf .
4. Naumann, F., & Riedl, M. (2011). EDDL - Electronic Device Description Language. Oldenbourg Industrieverl.
5. Field Device Integration Technology. http://www.fdi-cooperation.com/tl_files/images/content/Publications/FDI-White_Paper.pdf.
6. Shani, U., Franke, M., Hribernik, K. A., & K. D. Thoben. (2017). Ontology mediation to rule them all: Managing the plurality in product service systems. 2017 Annual IEEE International Systems Conference (SysCon), (pp. 1-7). Montreal, QC: IEEE.
7. Franke, M., Klein, K., Hribernik, K., Lappe, D., Veigt, M., & Thoben, K. D. (2014). Semantic web service wrappers as a foundation for interoperability in closed-loop product lifecycle management. . Procedia CIRP, 22, 225-230.
8. Pitoura, E., Bukhres, O., & Elmagarmid, A. (1995). Object orientation in multidatabase systems. ACM Computing Surveys (CSUR), 27(2), 141-195.

9. Martin, D., Domingue, J., Sheth, A., Battle, S., Sycara, K., & Fensel, D. (2007). Semantic web services, part 2. IEEE Intelligent Systems, 22(6), 8-15.
10. Srinivasan, V., Lämmer, L., & Vettermann, S. (2008). On architecting and implementing a product information sharing service. Journal of Computing and Information Science in Engineering, 1-11.
11. Wang, X., & Xu, X. (2013). An interoperable solution for Cloud manufacturing. Robotics and Computer-Integrated Manufacturing, 29(4), pp.232-247.
12. Jardim-Goncalves, R., Grilo, A., & Steiger-Garcao, A. (2006). Challenging the interoperability between computers in industry with MDA and SOA. Computers in Industry, 57(8), 679-689.
13. PLM Services 2.1 . (2011, May). Object Management Group.
14. Product Life Cycle support (PLCS) Web services V2. (n.d.). OASIS . http://www.plcs-resources.org/plcs_ws/v2/ .
15. Srinivasan, V., Lämmer, L., & Vettermann, S. (2008). On architecting and implementing a product information sharing service. Journal of Computing and Information Science in Engineering, 1-11.
16. ISO 10303-239 2005. Industrial automation systems and integration–Product data representation and xchange–Part 239: Application protocol: Product life cycle support. International Organization for Standardization, Geneva (Switzerland).
17. EFFRA, H. 2. (2016). Factories 4.0 and Beyond Recommendations for the work programme 18-19-20 of the FoF PPP. Retrieved from Horizon 2020 EFFRA: http://effra.eu/sites/default/files/factories40_beyond_v31_public.pdf
18. QLM. (2012). An introduction to Quantum Lifecycle Management (QLM) by The Open Group QLM work Group. Retrieved from http://docs.media.bitpipe.com/io_10x/io_102267/item_632585/Quantum%20Lifecycle%20Management.pdf
19. Parrotta, S., Cassina, J., Terzi, S., Taisch, M., Potter, D., & Främling, K. (2013). Proposal of an interoperability standard supporting PLM and knowledge sharing. In IFIP International Conference on Advances in Production Management Systems (pp. 286-293). Springer.
20. Kubler, S., Främling, K., & Derigent, W. (2015). P2P Data synchronization for product lifecycle management. Computers in Industry, 66, 82-98.
21. Främling, K., Kubler, S., & Buda, A. (2014). Universal messaging standards for the IoT from a lifecycle management perspective. IEEE Internet of things journal, 1(4), , 319-327.
22. Ameri, F., & Dutta, D. (2006). An upper ontology for manufacturing service description. In ASME Conference, (pp. 651-661).
23. de Vrieze P., Xu L. (2018) Resilience Analysis of Service Oriented Collaboration Process Management systems. Service Oriented Computing and Applications. Springer. 2018.
24. Marrella, A., Mecella, M., & Sardiña, S. (2017). Intelligent Process Adaptation in the SmartPM System. ACM Transaction on intelligent systems, 8(2).
25. Castelli, G., Mamei, M. Rosi, A., & Zambonelli F. (2015). Engineering pervasive service ecosystems: The SAPERE approach. ACM Transactions on autonomous and adaptive systems, 10(1).
26. Penzo, W., Lodi, S., Mandreoli, F., Martoglia, R., Sassatelli, S. (2008). Semantic peer, here are the neighbors you want!. Proceedings of ACM EDBT.
27. EFFRA Factories 4.0 and Beyond Recommendations for the work programme 18 - 19 - 20 of the FoF PPP under Horizon 2020, Online: http://www.effra.eu/sites/default/files/factories40_beyond_v31_public.pdf
28. Mandreoli, F. (2017). A Framework for User-Driven Mapping Discovery in Rich Spaces of Heterogeneous Data. Proc. 16th OTM Conferences (2), pp. 399-417.

29. Kasse, J.P., Xu, L,, de Vrieze, P, (2017) A Comparative Assessment of Collaborative Business Process Verification Approaches. In:18th IFIP Working Conference on Virtual Enterprises (PRO-VE 2017), Vicenza, Italy, 18 Sep 2017 - 20 Sep 2017. Springer
30. Sang, G., Xu, L., de Vrieze, P. (2017) Simplifying Big Data Analytics System with Reference Architecture. In: 18th IFIP Working Conference on Virtual Enterprises (PRO-VE 2017), Vicenza, Italy, 18 Sep 2017 - 20 Sep 2017. Springer.
31. Xu, L, de Vrieze, P. (2018) Supporting Collaborative Business Processes: a BPaaS Approach. International Journal of Simulation and Process Modelling. Inderscience.
32. Xu, L., de Vrieze, P., Yu, H., Phalp, K., and Bai, Y. (2018) Interoperability of Virtual Factory: an Overview of Concepts and Research Challenges. International Journal of Mechatronics and Manufacturing Systems.