Semantic CPPS in Industry 4.0

Giuseppe Fenza and Mariacristina Gallo and Vincenzo Loia and Domenico Marino and Francesco Orciuoli and Alberto Volpe

Abstract  Cyber-Physical Systems (CPS) play a crucial role in the era of the 4th Industrial Revolution. Recently, the application of the CPS to industrial manufacturing leads to a specialization of them referred as Cyber-Physical Production Systems (CPPS). Among other challenges, CPS and CPPS should be able to address interoperability issues, since one of their intrinsic requirement is the capability to interface and cooperate with other systems. On the other hand, to fully realize the Industry 4.0 vision, it is required to address horizontal, vertical, and end-to-end integration enabling a complete awareness through the entire supply chain. In this context, Semantic Web standards and technologies may have a promising role to represent manufacturing knowledge in a machine-interpretable way for enabling communications among heterogeneous Industrial assets. This paper proposes an integration of Semantic Web models available at state of the art for implementing a 5C architecture mainly targeted to collect and process semantic data stream in a way that would unlock the potentiality of data yield in a smart manufacturing environment. The analysis of key industrial ontologies and semantic technologies allows us to instantiate an example scenario for monitoring Overall Equipment Effectiveness (OEE). The solution uses the SOSA ontology for representing the semantic data stream. Then, C-SPARQL queries are defined for periodically carrying out useful KPIs to address the proposed aim.\[1]
1 Introduction

Cyber-Physical Systems (CPS) are a combination of physics with cyber components potentially networked and tightly interconnected. The physics components are Plants, Processes, and Systems, while the cyber components are Computation, Software, and Code. These systems are networked and interconnected thanks to interfaces that convert analog to a digital signal and vice-versa. In this way, each physic component has its digital twins that work in symbiosis to perform efficiently, enabling knowledge sharing and fast decision making.

Recently, the 4th Industrial Revolution and the relevance of the CPS inside the manufacturing plan lead to the definition of a niche of CPS applied to smart manufacturing, named Cyber-Physical Production Systems (CPPS).

Due to the heterogeneity of integrated components, the interoperability inside the CPPS begins a crucial requirement. It enables the system to reach and exploit data produced by other systems, and so on. However, a system of systems built with interoperability in mind would be difficult to manage in the long run and will slow the business processes over time. In other words, it becomes an obstacle to shareable knowledge opportunities for the systems and the stakeholders of the business.

Since a continuous flow of time-sensitive information is necessary for the plant activities, another relevant requirement of the CPPs is the ability to manage streams of data points. These data points are characterized by a time dimension that certifies the state of the producing source in a specific time window. It follows that procedures and analytics should be time-aware in terms of the processed stream. They should produce reports that resume the state in a time window.

In this paper, we propose a semantic component for the Cognition level of the usual 5C architecture (see Section 4.1). The semantic component is used to model the entire business, from the production floor to the management level. Each system models specific data points leveraging internal and external knowledge (i.e., produced by other systems). The idea is using the SOSA ontology to represent the different levels of the business to enable reasoning and unlock actionable knowledge of heterogeneous information sources [12]. The use of ontologies constitutes a powerful solution to capture and to share collective knowledge between stakeholders [10][6]. Furthermore, by applying a stream reasoning (i.e., C-SPARQL2), we enable a mechanism of continuous queries in order to retrieve specific knowledge inside the time windows.

The remaining of the document is structured as follows. The next section discusses related works that apply semantic models to CPPS. Section 3 describes the available semantic technologies, while Section 4 details our approach. A sample scenario is presented in Section 5 that proves the usefulness of the proposed component applied to the manufacturing domain. Finally, Section 6 ends the paper.

\[^2\] http://streamreasoning.org/resources/c-sparql
2 Related Works

This work proposes a 5C architecture adopting Semantic Web technologies and standards for unlocking the knowledge usually isolated in silos built and managed by different systems, for instance, like MES (Manufacturing Execution System), ERP (Enterprise Resource Planning), and so forth. The example scenario (see Section 5) focuses on the provisioning of useful KPIs to monitor the Overall Equipment Effectiveness using semantic data stream collected in a smart manufacturing environment.

Several systems integrating Semantic Web with Web of Things there exist. Some of them are focused on the modeling of CPS’ capabilities with Semantic Web technologies [7, 13, 8], and some others use them to approach CPS’ security aspects [2, 5].

A research work most related to ours is [1], it introduces a model called PriMa that uses a semantic layer to learn and reason about data that comes from stations and sensors of the CPPS and to produce feedback to a Decision Support Dashboard used by the maintenance manager. Shcherbakov et al. [11] exploit the real-time requirements by describing a concept model for a proactive decision support system design specifically for the maintenance of CPS to optimize downtime.

Different from state of the art, our approach proposes a model aiming to support integration at different levels and for different purposes in the CPS context. As also demonstrated by the application scenario, we construct a knowledge base shareable by systems and subsystems.

3 Semantic Technologies for Web of Things

In this section, the description focuses on the identification of semantic technologies for the web of things that are naturally applicable to the smart manufacturing and Industry 4.0.

Devices should be able to describe themselves and the data that they yield in a way that is understandable from both machines and humans. This may be done by equipping the devices and data with Semantic Web standards (or standard de-facto), ontological models, enabling to converge in the meaning of various concepts that describe data and devices, and fully interoperability among them. The formalization of the knowledge graph takes place through the construction of ontologies that allow us to share common meanings applied to a domain of interest and describe concepts semantically, in our case, IoT devices in the industrial field.

Semantic technologies substantially lie on the use of meta-models as: Resource Description Framework (RDF) [3] Resource Description Framework Schema (RDFS) [4].

---

[3] https://www.w3.org/RDF/
[4] https://www.w3.org/TR/2000/CR-rdf-schema-20000327/
Web Ontology Language (OWL). Defined by the World Wide Web Consortium (W3C), they allow to construct ontologies and populate them. Populated ontologies can then be queried by the RDF query language Simple Protocol and RDF Query Language (SPARQL).

During the design and modeling of an ontology for the IoT domain, it is necessary to know relevant aspects about devices such as the function to be performed, the environment in which they operate, the role and privileges within the operating context. Furthermore, it is relevant to reuse existing solutions and, eventually, extend them. In 2009, a W3C’s research selected and reviewed 17 existing ontologies related to the IoT domain. As a result, the consortium creates the Semantic Sensor Network (SSN) ontology.

The following sections will describe SSN and its evolution to model IoT streams and sensor data.

3.1 Semantic Sensor Network Ontology

The Semantic Sensor Networks (SSN) ontology is an OWL ontology that aims to describe the capacities and properties of the sensors, the act of detection, and the following observations. It was introduced by the W3C Semantic Sensor Networks Incubator group (the SSN-XG) held from March 2009 to September 2010.

As a choice of the working group, the ontology contains only the relevant concepts and relationships for the sensors, leaving the concepts related to other domains to be imported from other ontologies, where necessary. It can describe the sensors, their accuracy, observations, and the methods used for detection. The SSN ontology can be viewed from four main perspectives:

- The sensor perspective, with a focus on what to capture and how;
- Data or observation perspective, with a focus on observations and related metadata;
- The system perspective, with a focus on sensor systems and deployment;
- The perspective of features and properties, which focuses on what property to detect, or what observations have been made about a particular property.

3.2 Sensor-Observation-Sampling-Actuator Ontology

During the standardization of the Semantic Sensor Network by the W3C and OGC working groups, Spatial Data had the task of revisioning the SSN by using the ex-

---

5 https://www.w3.org/OWL/
6 https://www.w3.org/TR/rdf-sparql-query/
7 https://www.w3.org/TR/vocab-ssn/
8 https://www.w3.org/2005/Incubator/ssn/ssnx/ssn
periences obtained during the development of the prior version. Due to scope and audience changes, and to overcome the lack of interoperability with other ontologies, SSN was rewritten entirely by using a modular design that enables horizontal and vertical segmentation.

The new SSN, published as W3C Recommendation and OGC Standard, is the SOSA framework. This framework is composed of main concepts that enable interoperability with the entities, relationships, and activities involved in sensing, sampling, and actuation. Another relevant aspect of the SOSA ontology is the event-centric approach that enables interoperability with Schema.org and Prov-O models and, most importantly, with standard business processes procedures.

The sensor represents an entity that reacts to changes in the environment (e.g., black-out of a production line, temperature changes in a department). This detection produces the value of a property thanks to the (observation) procedure that was observing the specific property. Subsequent action performed due to the generated value is triggered thanks to an actuator. The last primary entity is the platform that hosts sensors and actuators.

4 Our Approach

A usual smart manufacturing environment is highly automated and networked. Each station is strictly interlinked with a digital twin to collaborate for performing reliably and being aware of the context in which they are placed. In this section, it will be provided a brief description of the 5C architecture that represents the fundamentals stone on which our approach (discussed in the next section) has been developed.

4.1 5C Architecture

In the area of Industrial Internet, 5C architecture provides a step-by-step guide for developing and deploying a CPS. The Industrial Internet, viewed as the application of CPS, is composed of five layers: smart connection level, data-to-information conversion level, cyber level, cognition level, and configuration level (see Figure 1).

The Connection layer has the goal of managing data acquisition from the sensors and controllers based on heterogeneous hardware and software platforms. This complex task can be achieved thanks to the use of the protocols capable of handling different hardware and software.

The Conversion layer uses the previous layer’s output and transforms data into information thanks to the support of external data sources. This component is crucial to realize self-aware physical components.

9 https://www.w3.org/2015/spatial/wiki/SOSA_Ontology
The Cyber layer is the information hub for data processing. It groups the output of the previous layer represented by individual information to support decision making done on the upper layer. It analyzes and compares the current performance of the group of stations and predicts future performances.

The Cognition layer collects and integrates machinery information to provide knowledge of the whole production. In this layer, big data analytic pipelines and knowledge representation are predominant and helps the upper layer (the Configuration one) in managing the entire system and in providing relevant KPI related to the environment.

Then the last layer (i.e., Configuration) supports the managers and shows the decision taken by the Cognition layer with respect to the production floor. It acts as a smart control system to provide self-configuration and self-adaptiveness.

4.2 Semantic Approach and Cognition Layer

Micro, small, medium, and large systems have to be conceptualized with interoperability in mind: the entire ecosystem has to be able to reach data produced by other systems, and so on. This work emphasizes the need for a highly interoperable component enabling integration among different systems in the industrial asset, like Enterprise Resource Planning (ERP) and Manufacturing Execution System (MES).

The solution proposes the adoption of semantic technologies inside the Cognition Layer of 5C architecture. The idea is to use the SOSA Ontology to model both static RDF data that represents the described companies’ assets and streaming part that is the on-line information that crosses the shop floor. Using the SOSA Ontology allows
us to: (1) exploiting an ontology that already models fundamental concepts about sensors, samplings, procedures, and so on. (2) Reusing existing ontology models. One of the main benefits of this abstraction is the integration of, for instance, ERP and MES data allowing to support decision-making processes.

Through SOSA’s classes and properties, we represent the entire environment and annotate raw data that came from the Cyber Layer. Then, by adopting a reasoner, it is also possible to infer new facts based on the forwarded rules. In that way, the component creates knowledge that is the sum of facts, forward rules, backward rules, and goals. Annotated data are stored in a triplestore (e.g., RDF4J, AllegroGraph, and so on.), offering a SPARQL endpoint for submitting queries. Such knowledge, inside the Configuration Layer, supports managers in supervising the business processes.

Sensors produce a rapidly evolving huge amount of data whose annotated version grows in turn. So, the semantic component must develop reliable production status in an online and proactive manner. To satisfy this requirement, the semantic component leverages C-SPARQL to do continuous queries over sensors’ streams. By designing suitable queries, the component can support managers’ decisions giving them focused knowledge for each specific time-frame.

5 Sample Scenario: Overall Equipment Effectiveness

This section aims to describe an application scenario in which the defined semantic component could result useful. The objective is to support managers by generating up-to-date KPIs related to the production line showed in Figure 2.

![Fig. 2 Basic production line with five stations.](image)

The semantic component intends to provide a fast and accurate *Synthetic Production Rate* that measures the efficiency of a facility. In particular, we will define the C-SPARQL queries able to evaluate: Availability, Efficiency of Performance, and Product Quality Rate of the factory. The starting point is the subset of SOSA ontology entities shown in Fig. 3 that presents a basic model related to a production plan, Blue boxes indicate SOSA classes, and white boxes indicate class instances, while solid lines represent RDF:type relations, and dashed lines, properties between entities. In particular, Fig. 5 shows the five stations that are instances of the SOSA class.

10 [https://rdf4j.org/](https://rdf4j.org/)
11 [https://franz.com/agraph/allegrograph/](https://franz.com/agraph/allegrograph/)
Fig. 3 Basic production line modeled with Sosa ontology.

Platform, with an additional Platform instance for the production line. These platforms host sensors observing features of interest. Following we will discuss some sampling C-SPARQL queries related to this ontology schema.

5.1 Availability

Availability is one of the three OEE factors, and its score takes into consideration any event (planned and unplanned) that stops the production. Planned stops refer to changes over time, while unplanned stops refer to equipment failures and material shortage. A lower availability score indicates problems or faults during the production; it leads to higher cycle-time that can influence the ability to satisfy demand. It is computed by recording the planned and unplanned stops. In particular, it is defined by Equation 1, where \(TotalTime\) represents the full-time window (e.g., one day) while \(DownTime\) is the sum of the planned and unplanned stops.

\[
Availability = \frac{TotalTime - DownTime}{TotalTime}
\]  

In the following code blocks, two streams that register and compute the \(DownTime\) and \(Availability\) values, respectively, are defined.

Listing 1 DownTime evaluation stream

```sql
REGISTER STREAM DownTime COMPUTED EVERY 24h AS
PREFIX sosa: <http://www.w3.org/ns/sosa/>
SELECT ?downTime
FROM STREAM <http://../production> [RANGE 24h STEP 1m]
WHERE {?sensor sosa:observes ?voltage.
  ?voltage rdf:type sosa:FeatureOfInterest.
  ?productionLine sosa:hosts ?sensor}
```
In Listing 1, we read voltage value from the specific sensor. We assume if such value is less than 5 V, the production line is down. So, the query counts the “down” triples and returns the total value. Successively, as expressed in Listing 2, assuming that the TotalTime is 1440 (i.e., 24 * 60), the query returns the resulting availability.

5.2 Efficiency Of Performance

The Efficiency of Performance (or simply performance) is a KPI that indicates how fast the process is running. It takes into consideration the ideal cycle time, the total production, and the operating time and is calculated as expressed in Equation 2.

\[
EfficiencyOfPerformance = \frac{TheoreticalCycleTime \times TotalProduction}{OperatingTime}
\]  

(2)

The TheoreticalCycleTime is specific for each production. It indicates the type of process, task sequence, technological dependencies, and duration of each task. The TotalProduction regards the number of final products for the production line while the OperatingTime is the inverse of the DownTime defined previously.

The evaluation of this KPI needs two different queries. The first one (see Listing 3) computes TotalProduction by counting products passing through the assembly sensor. The second one evaluates the EfficiencyOfPerformance by exploiting both TotalProduction and DownTime determined previously. For simplicity, the TheoreticalCycleTime is supposed to be a constant (i.e., 25 minutes). The corresponding query is described in Listing 4.

Listing 3 TotalProduction evaluation stream

REGISTER STREAM TotalProduction COMPUTED EVERY 24h AS
PREFIX sosa: <http://www.w3.org/ns/sosa/>
SELECT ?total
FROM STREAM <http://../production>
WHERE {
  ?assemblySensor sosa:observes ?product.
  ?product rdf:type sosa:FeatureOfInterest.
  ?platform sosa:hosts ?assemblySensor
}

Listing 4 EfficiencyOfPerformance evaluation stream

REGISTER STREAM EfficiencyOfPerformance COMPUTED EVERY 24h AS
PREFIX sosa: <http://www.w3.org/ns/sosa/>
SELECT (1440 - ?downTime)/1440 AS ?availability
FROM STREAM <http://../DownTime>
WHERE {
  ?assemblySensor sosa:observes ?product.
  ?product rdf:type sosa:FeatureOfInterest.
  ?platform sosa:hosts ?assemblySensor
}
AGGREGATE {(?total, COUNT, {?product})
FILTER (?platform = "http://.../ASSEMBLY/AP1A")}

Listing 4 Performance evaluation stream

REGISTER STREAM Performance COMPUTED EVERY 24h AS
PREFIX sosa: "http://www.w3.org/ns/sosa/"
SELECT (25 * ?total)/(1440 - ?downTime) AS ?performance
FROM STREAM <http://../TotalProduction> [RANGE 24h STEP 1m]
FROM STREAM <http://../DownTime> [RANGE 24h STEP 1m]

5.3 Product Quality Rate

The quality score indicates how many products (or components) have been produced without flaws, and it is computed by Equation 3. The TotalProduction is the same as the previous equation, while the DefectedProduction is the number of defective pieces.

\[
\text{Product Quality Rate} = \frac{\text{TotalProduction} - \text{DefectedProduction}}{\text{TotalProduction}} \tag{3}
\]

In this case, as expressed in Listing 5, it is necessary to read the TotalProduction value from the previously generated stream (i.e., "TotalProduction") and count the number of defective products passing through the integrity sensor.

Listing 5 Quality evaluation stream

REGISTER STREAM Quality COMPUTED EVERY 24h AS
PREFIX sosa: "http://www.w3.org/ns/sosa/"
SELECT ((?total - ?defectTotal) / ?total) AS ?quality
FROM STREAM <http://../TotalProduction> [RANGE 24h STEP 1m]
FROM STREAM <http://../production> [RANGE 24h STEP 1m]
WHERE { ?integritySensor sosa:observes ?defect. ?defect rdf:type sosa:FeatureOfInterest. ?platform sosa:hosts ?integritySensor }
AGGREGATE {(?defectTotal, COUNT, {?defect})
FILTER (?platform = "http://.../INTEGRITY/IT1A")}

5.4 Overall Equipment Effectiveness (OEE)

The Overall Equipment Effectiveness (OEE) resumes the overall production state by considering the three evaluated factors: Availability, Performance, and Quality,
through Equation 4. A 100% score indicates that the plant is manufacturing perfect parts and in the fastest way possible without stops. To compute the final OEE value, it is necessary to read from the three streams and compute the product of the values in the specific time-window (see Listing 6).

$$OEE = \text{Availability} \times \text{Performance} \times \text{Quality}$$ \hspace{1cm} (4)

Listing 6 OEE evaluation query

```sql
REGISTER QUERY OEE COMPUTED EVERY 24h AS
    SELECT (?availability * ?performance * ?quality) AS ?oee
    FROM STREAM <\protect\url{http://../Availability}> [RANGE 24h STEP 1m]
    FROM STREAM <\protect\url{http://../Performance}> [RANGE 24h STEP 1m]
    FROM STREAM <\protect\url{http://../Quality}> [RANGE 24h STEP 1m]
```

6 Conclusion

CPPS enables high efficiency and dynamics that are transversal through the manufacturing plants and the business units of a firm. To unlock this potentiality is required higher interoperability that sometimes can occur in a complex system in resource and maintenance needs. By taking into consideration these issues, this work proposes a semantic-based approach. We demonstrate how a semantic component could be beneficial to CPPS by providing interoperability and shareable knowledge. In particular, an analysis of the key industrial ontologies and semantic technologies is provided, and a common scenario is analyzed to show how powerful can be the proposed solution. Examples of queries relative to OEE show how effective is the use of a semantic component inside the Cognitive layer. In fact, the model supports the business and manufacturing processes and improves awareness about the overall manufacturing plant. In conclusion, the semantic component has resulted useful for:

- **Internal interoperability**: integration between the static knowledge represented by the industrial assets and the dynamic knowledge produced by the IoT sensors;
- **External interoperability**: integration between multiple business and industrial assets;
- **Time-sensitive KPIs**: computation of relevant and on-time KPIs useful during the decision-making process.

Acknowledgements This research was partially supported by the ECSEL-JU under the program ECSEL-Innovation Actions-2018 (ECSEL-IA) for research project CPS4EU (ID-826276) research project in the area Cyber-Physical Systems.
References

1. Fazel Ansari, Robert Glawar, and Tanja Nemeth. Prima: a prescriptive maintenance model for cyber-physical production systems. *International Journal of Computer Integrated Manufacturing*, pages 1–22, 2019.

2. Elias Bou-Harb, Walter Lucia, Nicola Forti, Sean Weerakkody, Nasir Ghani, and Bruno Sinopoli. Cyber meets control: A novel federated approach for resilient cps leveraging real cyber threat intelligence. *IEEE Communications Magazine*, 55(5):198–204, 2017.

3. Michael Compton, Payam Barnaghi, Luis Bermudez, Raúl García-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, et al. The ssn ontology of the w3c semantic sensor network incubator group. *Web semantics: science, services and agents on the World Wide Web*, 17:25–32, 2012.

4. Rainer Drath and Alexander Horch. Industrie 4.0: Hit or hype?[industry forum]. *IEEE industrial electronics magazine*, 8(2):56–58, 2014.

5. George Hatzivasilis, Ioannis Papaeftathiou, and Charalampos Manifavas. Real-time management of railway cps secure administration of iot and cps infrastructure. In *2017 6th Mediterranean Conference on Embedded Computing (MECO)*, pages 1–4. IEEE, 2017.

6. Tobias Hoppe, Harald Eisenmann, Alexander Viehl, and Oliver Bringmann. Shifting from data handling to knowledge engineering in aerospace industry. In *2017 IEEE International Systems Engineering Symposium (ISSE)*, pages 1–6. IEEE, 2017.

7. Veera Ragavan Sampath Kumar, Alaa Khamis, Sandro Fiorini, Joel Luis Carbonera, Alberto Olivares Alarcos, Maki Habib, Paulo Gonçalves, Howard Li, and Joanna Isabelle Olaszewska. Ontologies for industry 4.0. *The Knowledge Engineering Review*, 34, 2019.

8. Peter Gorm Larsen, John Fitzgerald, Jim Woodcock, Peter Fritzson, Jörg Brauer, Christian Kleijn, Thierry Lecomte, Markus Pfeil, Ole Green, Stylianos Basagiannis, et al. Integrated tool chain for model-based design of cyber-physical systems: The into-cps project. In *2016 2nd International Workshop on Modelling, Analysis, and Control of Complex CPS (CPS Data)*, pages 1–6. IEEE, 2016.

9. Jian-Qiang Li, F Richard Yu, Genqiang Deng, Chengwen Luo, Zhong Ming, and Qiao Yan. Industrial internet: A survey on the enabling technologies, applications, and challenges. *IEEE Communications Surveys & Tutorials*, 19(3):1504–1526, 2017.

10. Zhou Liandong and Wang Qifeng. Knowledge discovery and modeling approach for manufacturing enterprises. *Intelligent Information Technology Applications, 2007 Workshop on*, 1:291–294, 01 2009.

11. Maxim V Shcherbakov, Artem V Glotov, and Sergey V Cheremisinov. Proactive and predictive maintenance of cyber-physical systems. In *Cyber-Physical Systems: Advances in Design & Modelling*, pages 263–278. Springer, 2020.

12. Ovidiu Vermesan, Peter Friess, et al. *Internet of things-from research and innovation to market deployment*, volume 29. River publishers Aalborg, 2014.

13. Alexander Willner, Christian Diedrich, Raâd Ben Younes, Stephan Hohmann, and Andreas Kraft. Semantic communication between components for smart factories based on onem2m. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–8. IEEE, 2017.