Semantics for Cyber-Physical Systems: A Cross-Domain Perspective

Marta Sabou a,b,c,*, Stefan Biffl c, Alfred Einfalt a, Lukas Krammer a, Wolfgang Kastner c and Fajar J. Ekaputra b,c
a Siemens Corporate Technologies, Vienna, Austria
E-mail: {first.last}@siemens.com
b Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle
c Faculty of Informatics, Technische Universität Wien, Vienna, Austria
E-mail: {first.last}@tuwien.ac.at

Editors: Pascal Hitzler, Kansas State University, Manhattan, KS, USA; Krzysztof Janowicz, University of California, Santa Barbara, USA
Solicited reviews: Agnieszka Ławrynowicz, Poznan University of Technology, Poland; Stefano Borgo, Laboratory of Applied Ontology, ISTC-CNR, Trento, Italy; Freddy Lécué, CortAIx Thales, Montreal, Canada and INRIA, France

Abstract. Modern life is increasingly made more comfortable, efficient, and sustainable by the smart systems that surround us: smart buildings monitor and adjust temperature levels to achieve occupant comfort while optimizing energy consumption; smart energy grids reconfigure dynamically to make the best use of ad-hoc energy produced by a host of distributed energy producers; smart factories can be reconfigured on the shop-floor to efficiently produce a diverse range of products. These complex systems can only be realized by tightly integrating components in the physical space (sensors, actuators) with advanced software algorithms in the cyber-space, thus creating so-called Cyber-Physical Systems (CPS). Semantic Web technologies (SWT) have seen a natural uptake in several areas based on CPS, given that CPS are data and knowledge intensive while providing advanced functionalities typical of semantics-based intelligent systems. Yet, so far, this uptake has primarily happened within the boundaries of application domains resulting in somewhat disconnected research communities. In this paper, we take a cross-domain perspective by synthesizing our experiences of using SWTs during the engineering and operation of CPSs in smart manufacturing, smart buildings, and smart grids. We discuss use cases that are amenable to the use of SWTs, benefits and challenges of using these technologies in the CPS lifecycle as well as emerging future trends. While non-exhaustive, our paper aims at opening up a dialog between these fields and at putting the foundation for a research area on semantics in CPS.

Keywords: cyber-physical systems, Industrie4.0, smart energy networks, smart buildings, Semantic Web technologies

1. Introduction

Recent years have brought about accelerated developments in embedded networked systems such as the Internet of Things, communication technologies and information processing, as well as, as a side effect of these advances, their convergence to novel, complex systems generically referred to as Cyber-Physical Systems (CPS) [1]. CPS span the physical and cyber-world by linking objects and processes from these spaces. In a typical CPS, data are collected from the physical world via sensors while computation resources from the cyber-space are used to integrate and analyze the information in order to decide on optimal feedback processes which can be put in place by physical actuators. CPS go beyond traditional engineering systems in terms of size, complexity and dynamism. CPS have diffused and play an increasingly important role in a variety of (mission critical) domains and their infrastructures, including public transportation, energy services, and industrial production. Therefore, CPS are at
In terms of the information processing aspect, an emerging concept in CPS research, and beyond, is that of a Digital Twin (DT): the digital representation of a physical system, which reflects the system status thanks to data collected in real-time through sensors across the entire life-cycle of the system. As such, the Digital Twin consists both in a model of the system itself (e.g., its components and their characteristics) as well as real-time data. Therefore, the Digital Twin relies on the combination of several, heterogeneous, often dynamic data sources and should pave the way to analytics that support advanced functionalities – thus providing an excellent context for the use of SWTs.

Not surprisingly, the use of SWTs in settings that bridge into the physical space, have already been investigated in the last decade, for combining sensor networks with the Web [4], augmenting products with semantic descriptions [5], or enabling smart city infrastructures [6]. Since then, the application of SWTs has been steadily increasing, focusing on entire systems (e.g., CPS), even in mission-critical domains. However, research mainly focused within the boundaries of concrete domains and research communities, such as manufacturing [7, 8], electric grids [9], or buildings [10]. As a result, there is a lack of understanding of the commonalities and differences between applying SWTs in the CSP life-cycles of these domains, thus hampering exchange of ideas between communities, comparison of solutions and exchange of data.

With the amplified interest in CPS, this is therefore a good time to go beyond the boundaries of domain-focused research communities, and to reflect on commonalities across them, such as:

- What are domain-overarching CPS use cases amenable for the use of SWTs?
- Which SWT capabilities can support CPS use cases best?
- What challenges were observed when applying SWTs in CPS life-cycle so far?
- What are future trends that will influence semantic research in CPS in the next decade(s)?

In this paper, we aim to answer these questions by taking a cross-domain view on the applications of Semantic Web research for CPS. To do so, we build on an extensive study of the use of SWTs in smart manufacturing [7] and extend it with experiences in the areas of smart grids and buildings gathered in Austria’s largest Smart City Living Lab, the Aspern Smart City Initiative³. Admittedly, we aim at focusing on a few key aspects of the topic and do not have the ambition to be exhaustive, but rather to establish a first dialog across researchers applying SWTs in different CPS domains.

We start with a brief introduction of the three CPS-based application domains that informed this paper in Section 2, then discuss topics regarding semantics-amenable use cases, benefits and challenges of SWTs as well as emerging future trends in Sections 3 to 5.

2. CPS in mission-critical domains

For each mission-critical domain, we discuss the notion of CPS and why SWTs are promising.

2.1. Production Systems (Industry4.0)

The manufacturing sector is facing challenges such as shorter time to market, increased product diversification and customization, highly flexible (mass-)production while ensuring high product quality and improved production efficiency. Several initiatives aim to address these challenges by modernizing industrial production: Industrie 4.0 [11] in Germany, the Factory of the Future initiative in France, and the UK [12] or the Industrial Internet Consortium in the US.

Core to these initiatives is the focus on increased digitization of production systems in factories and of production processes. These digitization efforts lead to the upgrade of traditional factories to cyber-physical production systems (CPPS) and the digital representation of the CPPS through their digital twins.

Industrial production has several characteristics that make it an attractive application area of Semantic Web research. First, it is a knowledge and data intensive domain: the engineering of products and of the factories that produce them rely on complex engineering knowledge; large data sets are handled both during the engineering (e.g., a factory may be described by tens of thousands of signals and components) and operation of CPPS (e.g., logs of the production process). Second, the engineering of complex mechatronic objects, especially production systems, is increasingly driven

---

³Aspern Smart City Research: https://www.ascr.at/en/
by information models that enable representing different aspects of the produced system [13]. To that end, a range of IEC/ISA standard information models are adopted during the engineering of factories. However, these standard information models lack a formal semantics that would make them amenable to automated processing. Third, data exchange standards, such as SysML and AutomationML [14], provide standardized schemas to represent engineering information and as such address syntactic heterogeneity across engineering disciplines, but again they do not address semantic heterogeneity of the data encoded with them. Therefore, challenging tasks according to [13] include: model representation, model transformation, model integration, model consistency management, and flexible comparison of components, as detailed in Section 3.

2.2. Energy Systems (Smart Grids)

Smart Grids, also referred to as cyber-physical energy systems (CPES) [15], are the next evolution step of the traditional power grid and are characterized by a bidirectional flow of information and energy [16]. A key driver of this trend is the shift towards more sustainable energy supply by using renewable energy resources. This fundamental change influences the whole value creation chain in electric power systems as well as the operation of the underlying infrastructure. In order to manage the volatile nature of renewables, smart grid solutions highly depend on advanced automation and control concepts as well as elaborate information and communication technologies. The complexity of applications offering new services, such as demand response, load shedding and shifting [9] is steadily increasing. Various controllers, actuators, sensors, and measurement units connected to devices from different stakeholders must work together with supervisory control and management (SCADA) systems, often in heterogeneous environments [17]. Furthermore, the energy markets are switching from a consumption- to a production-oriented paradigm with the ability for dynamic pricing and offering flexibility as a service [18] to improve the economics and reliability of the grid.

This increased complexity, heterogeneity and automation of the grid requires adequate digitization, i.e., through digital twins. Electric digital twins could enable planning, operation, and maintenance of grids based on a set of information models. For instance, it will be possible to plan how to integrate new components and controls in the daily network operation business considering different steps of operation, e.g., the planning process or the daily field work. Digital twins could also offer a solution to dealing with sensor data—created as a side effect of decentralization and renewables—which is complex to manage and exchange.

Electric digital twins enabled by SWTs are already available for high-voltage grids because they contain a relatively small and static number of devices. It is therefore feasible to semantically describe these devices and manage the resulting static digital twin4. There is also an abundance of domain vocabularies for modeling power grid information. For example, the Common Information Model (CIM) allows describing power system resources such as energy management systems, SCADA systems and power system topology. It can therefore act as a domain ontology for digital twins in the energy sector.

In low-voltage grids, however, both the number and diversity of devices is much higher than in high-voltage grids thus making the application of SWTs more desirable, but at the same time also more challenging. Therefore, medium-sized Distribution System Operators, in charge of low-voltage grids, are still at the beginning of mapping their infrastructure to a coherent electrical digital twin, and provide a promising application area for SWTs.

2.3. Building Management (Smart Buildings)

Residential and commercial buildings are the third main sector of final energy consumption besides industry and transportation. Key to sustainability in this area are cyber-physical systems in the form of networked automation systems, also known as building automation systems (BAS). While BAS provide technological means to increase energy efficiency and preserving comfort, a recent trend is the evolution of buildings towards so-called smart spaces, in which humans and technology-enabled systems interact in open, connected, coordinated, and intelligent ecosystems [19].

The need for information modelling and integration of heterogeneous data is present in several aspects of smart buildings. The development of BAS over the years led to a heterogeneous landscape of networking standards, technologies and proprietary BAS solutions. Deployed BAS solutions are often specialized for a distinct field of application (i.e., trade) in a building. Therefore, it is necessary to deploy more than one

---

technology within a single building. These subsystems must be able to exchange information by relying on shared data models and interfaces (i.e., digital twins). To that end, SWTs, in particular ontologies, are extensively used for information modeling for building automation [10].

Another active area of research is the interoperability of different BAS standards. For example, **OPC Unified Architecture (OPC UA, IEC 62541)** supports the modeling of engineering and runtime data through an object-oriented approach. For technical equipment in buildings, also more specific models and taxonomies are used, for example **brick schema** [20] or the **project haystack** [21], some of which already use RDF to facilitate basic integration across standards.

The advent of the **digital building and Building Information Modeling (BIM)** further increased the importance of modelling engineering and runtime data as well as the relation between them in a digital twin. To that end, several industry standard data models were developed, such as **Industry Foundation Class (IFC)** or BIM. SWTs can facilitate this integrated use of the models through techniques such as ontology-based data integration or ontology matching.

### 3. Where can Semantic Web technologies help?

Based on the experiences of the authors in the three domains, we present a (non-exhaustive) list of use cases valid across the three domains, where the application of SWTs is promising. Fig. 1 captures a simplified view of the CPS life-cycle across two stages, namely (1) engineering and (2) operation and maintenance. In each stage, we distinguish between the physical and the digital space. The physical space covers the concrete, material system both during CPS construction (engineering) and, after commissioning, CPS operation and maintenance. The digital space depicts the main information models at each stage and the use cases related to these data models.

#### 3.1. Engineering

The engineering of a complex CPS (a smart factory, smart building, or smart grid), is typically performed in multi-disciplinary settings, where (engineering) experts with different expertise collaborate towards creating a **Digital Model** of the CPS according to which the real-world CPS is built as part of the **deployment** process. Such multi-disciplinary settings are characterized by the need to support this collaborative effort towards (1) integrating heterogeneous engineering data models into a single, complete and consistent digital model; (2) ensuring the consistency of this model; and (3) supporting modifications of this model through artefact reuse. Accordingly, there is an opportunity to use SWTs to support these tasks, as follows:

**Engineering model integration** aims to bridge semantic gaps in engineering environments between
project participants (and their tools), who use heterogeneous local terminologies. The challenge is addressing this heterogeneity by aligning, and subsequently integrating engineering models (e.g., ontologies), at terminological and/or instance level. This integration is a prerequisite for supporting the analysis, automation, and improvement of multi-disciplinary engineering processes that rely on this data. Model integration is also necessary for creating discipline-crossing Engineering Tool Networks that enable interacting appropriately within an engineering network covering different engineering disciplines, engineers, and their tools.

In the area of production systems, ontology-based data integration methods are widely used to integrate engineering models [22] and, more recently, Knowledge Graphs were proposed for addressing this task [23]. Although there are several ontologies available for different engineering domains relevant to smart buildings, there is a need for ensuring interoperability and improved interaction processes among these different domains [24], for example by ontology alignment [25]. To support the engineering of smart grids, an ontology matching process is proposed to align the Common Information Model and IEC 61850 standards [26].

Model consistency management refers to the task of detecting defects and inconsistencies in the digital model, including models of individual engineering disciplines as well as across interrelated models from diverse engineering disciplines. In smart manufacturing, SPARQL queries are used to check for inconsistencies across engineering models integrated by using OWL [27]. The AutomationML Analyzer tool relies on Linked Data principles to provide an interface for browsing and query-based consistency checking of integrated engineering models [28]. More recently, the Shape Constraint Language (SHACL) is used for consistency checking of engineering models [29]. In the smart building domain, different types of models are used, e.g., architectural models, models for technical equipment and functional models. Several projects focus on consistency checks among different models to guarantee a well-working BAS. Examples are: an environment for semantic rule checking of building models [30] and an ontology-based approach for conformance checking in construction [31].

Flexible comparison for artifact reuse focuses on the identification of reusable system components within an engineering project. The core task is the evaluation of component models to decide about their potential usability within a CPS. To support the problem of parts exchange in an evolving manufacturing system, in particular, checking the compatibility of the old and the new part, SysML models are translated into OWL ontologies and then compatibility constraints are checked through SPARQL queries [32]. SWTs enable flexible comparison among products and production processes during product ramp-up in order to identify a suitable production process at a target site, which enables producing a product with the same quality as at a source site [33]. The design phase of a BAS requires the comparison of devices and the automatic evaluation of their interoperability. This is enabled by ontology-based device descriptions covering both hardware and software characteristics [34].

3.2. Operation and Maintenance

The transition from engineering to operation is marked by the commissioning of the CPS: at the physical level this includes putting the system in operation (e.g., transporting and installing a previously engineered production system); at the digital level, the digital model is, ideally, also passed on to the operation phase (note: in reality, the digital model is often not shared with the stakeholders in the operation phase).

During operation (runtime), information is collected through sensing about the functioning of the CPS through various sensor streams. For example, in a production system, information about the materials used, the current process, and the positions of the industrial robots can be recorded. In smart grids, active and reactive power in the distribution grid is recorded. This runtime information complements the "static" digital model created during engineering and results in the Digital Twin of the CPS, which enables a variety of use cases that could be supported by semantics.

Monitoring and anomaly detection focuses on the acquisition and interpretation of dynamic system data with the goal of the early detection of anomalies or faults in the system operation. In smart factories, a constraint- and reasoning-based approach identifies those production processes where too much material is used (in comparison to quotas defined at engineering time) for creating a product [35]. In smart grids, reasoning on data streams collected from distribution network field devices (e.g., battery energy storage systems) helps identifying voltage levels that are defined as dangerous by operators through a rule management interface [36] at an Aspem Smart City Research settlement. Smart buildings require detecting and reasoning about fault propagation in BAS. For example, diagno-
sis in terms of cause-effect relations in smart buildings was enabled by extensions to the SSN ontology [37]. In [38], the BAS knowledge is enriched with causal relations between building components and data points by relying on building information models and causalities automatically derived through a set of rules.

**Maintenance and replacement engineering** focuses on finding and replacing faulty components, potentially after an anomaly detection phase. This use case requires the combined use of dynamic runtime data (for detecting anomalies) and static engineering data about the structure of the system (the Digital Model) and the characteristics of the faulty component that needs to be replaced. Device exchange in the context of power plants was considered in [39] with a solution based on the AutomationML language.

**Adaptation through optimization and reconfiguration** is the “ability of the CPS to achieve an intended purpose in the face of changing external conditions such as the need to upgrade or otherwise reconfigure a CPS to meet new conditions, needs, or objectives” [3]. We distinguish different levels of adaptation. **Optimization** changes the system behaviour through control if operation conditions change: if some condition is not fulfilled, then there is an actuation action (e.g., reduce temperature). **Reconfiguration** is a more advanced notion of adaptation, where the system set up itself changes to respond to new goals or external factors. For example, the runtime flexibility of production systems in order to produce new products, requires the integration of knowledge about the production system and the product to enable the use of advanced techniques such as configurators [40]. In building automation, the smart control ontology (Colibri) follows a service-centric approach for modeling data and functionality [41] and enables the generation of optimization problems. For optimizing end-user energy consumption, an ontology for smart grid interactions with Building Energy Management Systems was designed [42].

A use case that spans both the operation and engineering phases of CPS, concerns enabling backflow from operation experience to improve the engineering of similar CPS. The goal is to bring back operation data and analysis results from systems into the engineering environment in order to improve the engineering of new systems based on experiences with the performance of already built systems (e.g., frequent defects and affected devices) [43].

### 4. Lessons learned from the application of SWTs

#### 4.1. Benefits of Semantic Web technologies

The following SWT capabilities were perceived as beneficial in the CPS life-cycle:

- **Formal and flexible semantic modeling** refers to the capability of explicitly capturing a universe of discourse. Unlike other (semantic) modeling approaches (e.g., UML, SysML), Semantic Web knowledge representation languages offer unambiguous, formal semantics that enable reasoning. Additionally, the ability to evolve an ontology on the schema and instance levels at runtime, provides a high degree of flexibility [44]. This capability is at the basis of all semantic-enabled use cases, with ontologies having been used to represent a variety of engineering knowledge [8, 10].

- **Intelligent, web-scale knowledge integration** is enabled by ontology matching, Linked Data, and ontology-based data integration techniques. This capability addresses those use cases where the heterogeneity of engineering data requires model and data integration. In particular, ontology-based data integration is widely-used to integrate engineering data [22].

- **Quality assurance of knowledge with querying and reasoning** support data validation and consistency checking both during CPS engineering (e.g., model consistency management) and operation (e.g., monitoring and fault detection). The formal semantics of ontologies and links between ontology concepts and instances (e.g., owl:sameAs) enable quality assurance tasks such as consistency checking (e.g., by SPARQL constraints) or data validation (e.g., with SHACL).

- **Browsing and exploration of distributed data sets** is enabled by Linked Data technologies [28]. This capability can be used to efficiently browse and explore both engineering models internal to an organization and external data sources, such as Web resources of third-party providers, supporting, e.g., artefact reuse.

#### 4.2. Challenges in using Semantic Web technologies

Some requirements and assumptions of CPS are fundamentally different from typical Semantic Web application areas, which focus on the integration of Web-scale data. This leads to a number of challenges when using SWTs in CPS engineering and operation.

- **Lack of knowledge acquisition interfaces** that are easy to use by engineers is a major challenge for the adoption of SWTs and has only been mitigated with partial solutions. For example, widespread en-
engineers or modeling languages (e.g., SysML, SysML4Mechatronics) are used as a “front-end” to acquire engineering models which are then translated into ontologies. Excel is often used in practice by engineers as knowledge acquisition tool, however, it misses the means for formal semantics definition. Therefore, the short-term benefit of using Excel can become a major liability as it is hard to check the semantic correctness of the data [45]. Finally, domain-specific knowledge acquisition tools are built, such as SOMM [35].

**Weak support for modeling engineering-specific knowledge structures.** Modeling engineering knowledge is characterized by the need to model system components with their roles, as well as part-of or other connections between them [46]. While modeling these structures is not straightforward in ontology engineering languages (e.g., there is no OWL part-of relation with formal semantics), ontology design patterns [47] could offer a solution to translate engineering modeling needs into ready-to-reuse modeling solutions.

**Lack of support for mathematical calculations,** which are frequently required in engineering-specific settings relying on processing of numeric data. While SWTs focus mostly on logistics-based knowledge representation and do not have a strength in advanced processing of numeric data, hybrid solutions are often proposed that combine SWTs with techniques more suitable to mathematical data processing, such as data mining, statistical analysis or Relational Constraint Solvers (RCS) for solving cardinality problems.

**The Open World Assumption (OWA) is not a natural fit to the engineering domain,** as traditional engineering approaches, e.g., databases and quality assurance methods, rely, in general, on a Closed World Assumption (CWA). This issue is partially addressed by mechanisms that combine open- and closed-world reasoning [48] such as expressing negations in SPARQL 1.1 queries.

**Difficulties for the integration of SWTs with existing enterprise systems** are two-fold. First, they stem from differences between object-oriented methods in the business environment, which typically rely on task-specific models, and ontologies that are conceptual domain models. Second, the lack of SWT skills among engineers hampers the adoption of these technologies. A solution for enabling software engineers to develop enterprise systems on the basis of an ontology relies on an adjustable transformation from OWL to Ecore, which allows authoring of and programmatic access to a reference ontology, through a familiar development environment (e.g., Eclipse) [44].

### 5. Looking forward: future research trends

To conclude, we discuss selected emerging trends in CPS that will require substantial, long-term research.

**Cross-domain applications** require solving semantic challenges in CPS spanning several domains, for example, a Smart Factory occupying a Smart Building supplied by the Smart Grid; or an e-Car charging scenario, where product information (of the e-Car), building energy management system, and smart grid information are equally important. For such scenarios, it will be necessary to dynamically discover available sensors, to access, interpret and integrate their data. FIWARE[3] and Web of Things (WoT)[6] are notable emerging efforts towards large-scale, semantic sensor discovery and integration. Common to these efforts is a predominant focus on using light-weight ontologies (e.g., taxonomies) and knowledge graphs as opposed to logically-complex (heavy-weight) models.

**Ontological foundations of CPS.** The realisation of cross-domain applications will require a meaningful harmonization of CPS viewpoints and terminology across domains. Future research should focus on establishing such a shared conceptual framework across CPS domains and creating a corresponding ontological viewpoint of CSP that can act as a basis for cross-domain research.

**Cyber-Physical Social Systems (CPSS)** consist not only of software and raw sensing/actuating hardware, but are fundamentally grounded in the behaviour of human actors involved in the system [49, 50]. A fundamental change in CPSS is that data is both received from (e.g., from mobiles, social networks, medical sensors) and distributed to human actors. This opens up challenges related to: (1) integrating data about the social component of the system from a variety of sources (e.g., social networks, mobile operators, wearables) both during the design and operation of the system; and, therefore, (2) data privacy, such as privacy-aware data integration and presentation, e.g., based on explicit, semantically represented usage policies [51].

**Explainable CPS.** CPS increasingly enable complex infrastructures, and, often directly interfere with every-day activities of a large user base (e.g., as in the case of smart energy grids). It becomes therefore important that these systems can provide (on demand) comprehensible explanations of the reasons for their status/behavior to a range of stakeholders includ-

---

ing end-customers and infrastructure operators (e.g., through a relevant chain of causation events). Among others, explainable CPS will require novel functionalities for detecting, representing, and reasoning about events within and outside the boundaries of the system—a knowledge intensive task which can benefit from SWTs (e.g., knowledge representation, provenance tracking, data integration).

Acknowledgements. We acknowledge the financial support of FFG PoSyCo (867276), FFG CitySPIN (861213), Asperrm City Research, the Christian Doppler Association, the Austrian Federal Ministry for Digital&Economic Affairs and the National Foundation for Research, Technology and Development.

References


