

Semantic Data Integration with DevOps to Support Engineering Process of Intelligent Building Automation Systems

Iori Mitzutani
iori.mitzutani@unisg.ch
University of St.Gallen
Switzerland

Ganesh Ramanathan
ganesh.ramanathan@siemens.com
Siemens AG
Zug, Switzerland

Simon Mayer
simon.mayer@unisg.ch
University of St.Gallen
Switzerland

ABSTRACT

The reliable infrastructure of building automation (BA) systems forms the foundation of smart environments and energy systems in our building towards increasing occupant comfort and safety while reducing the ecological footprint of buildings. This is achieved through the processing of data points collected from sensors and the control of installed actuators, and increasingly incorporates machine learning components. However, engineering of BA systems is intricately linked with the planning, installation, (pre-)commissioning, and operation of building services such as HVAC, and it requires an extensive amount of manual coordination which is often prone to errors, many of which are only detected late in the lifecycle and tends to lose transparency in data provenance. To address this, we propose the application of DevOps, a highly successful paradigm in the field of software engineering, to BA engineering process coordination. In addition, the possibility of using semantic data to develop artifacts such as requirements, construction, and devices of BA systems opens up the avenue of achieving continuous verification of the system as it is built and commissioned. Concretely, we propose a novel approach that integrates a semantic reasoner using the machine-understandable data of the building along with interactions facilitated by Web of Thing Thing Description to the DevOps workflow. The proposed approach is expected to ameliorate limitations of existing workflow management methods and thus provide transparency in the data provenance to gain trust for data-driven AI applications for BA.

CCS CONCEPTS

• **Software and its engineering** → *Software development process management*; • **Information systems** → *Information integration; Web applications*; • **Human-centered computing** → *Collaborative content creation*.

KEYWORDS

Building Automation, DevOps, Cyber-Physical Systems, Semantic Data, Provenance, Explainable CPS

ACM Reference Format:

Iori Mitzutani, Ganesh Ramanathan, and Simon Mayer. 2021. Semantic Data Integration with DevOps to Support Engineering Process of Intelligent Building Automation Systems. In *The 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '21)*, November 17–18, 2021, Coimbra, Portugal. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3486611.3492413>

1 INTRODUCTION

Modern buildings are integrating hardware and software components to create smart applications along with advances in machine learning (ML) models. This integration, however, requires transparency in the software components of the systems for both humans and machines to interpret the decision making process in the explainable artificial intelligence context [20]. However, opaqueness in cyber-physical systems (CPS) like BA arises not only from the software aspect, but also from the hardware constituting the system and its engineering [12]. For example, ML models in fault-detection diagnosis (FDD) systems rely on data points (DPs) collected from field devices, but credibility of the DPs can be affected by the engineering of the system (i.e., collected from which sensor, measured at what time, through which network, etc.). Realizing such data-provenance retrieval has been a major challenge for industry to date because of the heterogeneity of disciplinary domains involved in the engineering process [11]. To this end, standards such as ISO 16484 have helped planners, project managers, and engineers by encouraging them to organize the engineering process by breaking it down into distinct phases such as design, planning, installation, and commissioning [9] and specifying documentation artifacts to be generated in each phase. Fig. 1 shows the key activities and information flows of the engineering process. There are several phases in the workflow where synchronization between knowledge sources (both human and machines) is required. This brings us to question the adequacy of the system to be built: *Will the planned BA system meet the requirements and regulations? Will the system be installed as planned? Will certain changes cause other parts of the system, including software components, to fail?* Consider, for example, that the capacity of a central boiler in a building needs to be changed because of constraints discovered during a renovation.

The impact of such a change might propagate not only to the ML models in action, but also to different levels of components and stakeholders – the control strategy, the physical construction, and the energy management. Under the traditional manual management of engineering processes, technicians need to find relevant semantic data including information about tools to validate the change and their reasoning behind, i.e., provenance. In the absence of an unifying framework, synchronisation activities (e.g., project meetings) use these artifacts by largely relying on traditional

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

BuildSys '21, November 17–18, 2021, Coimbra, Portugal

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-9114-6/21/11...\$15.00

<https://doi.org/10.1145/3486611.3492413>

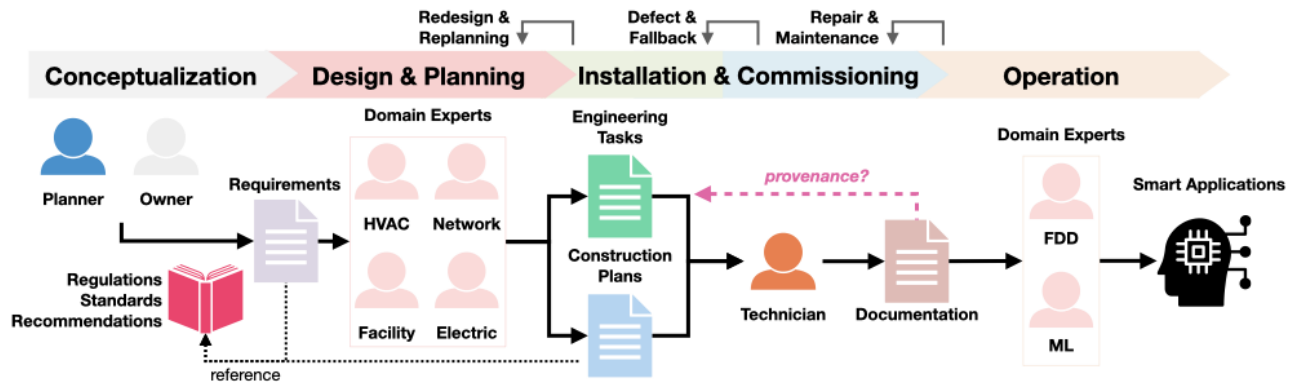


Figure 1: Engineering workflow with extensive manual verification towards smart applications.

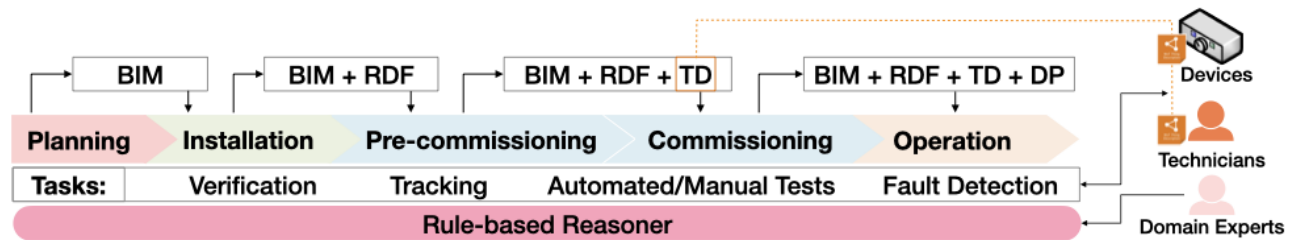


Figure 2: CI/CD pipeline for BA engineering process coordination.

human-to-human interaction where a technical project manager functions as the hub. For this reason, the design, integration, and deployment of industrial CPS remain a highly complex endeavor. To deal with such complexity and to enhance transparency of the provenance of data used for smart applications in BA systems, we propose to support BA engineering processes with semantic data by employing the concept of DevOps – a successful paradigm of software development and IT operations management, which has also been adopted by other fields such as deployment management of network engineering (DevNetOps [7]) and ML models (MLOps [1]).

2 CHALLENGES

Semantic data has been adopted to create knowledge bases with artifacts such as requirements, regulations, construction plans, network configurations, and device capabilities [21]. As seen in industrial applications using data-driven ML-based approaches [22], semantic data lends itself well to the use of a knowledge reasoner [6] to validate changes and assist in establishing coordination between the project participants. However, there are two major challenges to integrate BA engineering processes and with semantic data: (i.) BA engineering takes place in a *fragmented project organization*, with multiple domains and standards, making it difficult to create a comprehensive knowledge base, and (ii.) BA component integration requires the intervention of experts from different domains because it involves *heterogeneous information* such as building construction plans, device descriptions, and control logic descriptions, whereas the integration of software components is possible through formal validation and defined interfaces.

3 APPROACH

Our approach centers around integrating semantic data of BA Engineering artifacts in the concept of a Continuous Integration / Continuous Delivery (CI/CD) pipeline as shown in Fig. 2. To address the challenges, we propose a two-pronged approach by (i.) integrating automation devices, equipment, and domain knowledge from human experts as first-class members of the development ecosystem by representing them using *machine-understandable* descriptions based on the the W3C Web of Things (WoT) Thing Description (TD), which can then be managed and exchanged through the CI/CD pipeline, and (ii.) harnessing the semantic data to create a rule-based reasoner that serves as a foundation to achieving semi-automated verification layer for the CI/CD pipeline by also assuming that domain experts, technicians, and even devices are source of knowledge. These two aspects enable automatic validation of the BA engineering processes in the different phases.

To illustrate this concept, consider an example of changes in boiler capacity and a smart application monitoring the BA system: Domain rules in HVAC states the relation of a boiler's capacity to the requirement of the consumers and control of the distribution network, whereas rules from the electrical system relates it to the power supply and peak load restrictions. Each engineering task associated with a change in boiler capacity will produce a corresponding artifact (e.g., BIM) to be committed to a repository, and the CI/CD pipeline triggers semantic reasoning on the artifact, validating the changes made to the system. The description of the capabilities of the boiler, the electrical panel, the hot water distribution pumps, and the link to the responsible technical experts are

available as TDs. The semantic reasoner uses the knowledge graph to understand the chain of consequences and points to relevant resources via TDs which can provide confirmation or runtime data to help validate the change. This systematic workflow ensures the transparency of the system that emits and collects DPs, which in turn improves the trust of the smart application in its operational state.

3.1 CI/CD Pipeline for BA Engineering Phases

We propose that the development phases in BA should be integrated in a five-phase CI/CD pipeline as depicted in Fig. 2. By managing both the BA engineering artifacts and the knowledge base with pull-based development model [8], the semantic knowledge base is built and enhanced incrementally during the planning, installation, pre-commissioning, commissioning, and operational phases. Discrepancy discovered during each phase result in the step failing and, if required, the previous phase to be reviewed.

Planning: As the project progresses, domain experts develop the building plan using *Building Information Model* (BIM) according to the requirements defined by the project owners and planners. Domain experts produce the BIM data in an file exchange format as specified in the ISO 16739 Industry Foundation Classes (IFC) standard, which includes drawing of the building construction (e.g., AutoCAD DXF) and metadata of the building components. The IFC file is then committed to a code repository (e.g., Git) so that the CI/CD pipeline can support this phase by translating requirements to RDF and verifying compliance with relevant regulations, standards, and recommendations using the ifcOWL ontology [14], for example.

Installation: Given the output of the planning phase, the installation can progress while committing changes to the BIM repository, which invokes a semantic reasoner to detect discrepancies between the installation and the description in the BIM data. A tracking task that can be ported or integrated from classical project management tools use the reasoner to confirm the prerequisites for all dependent downstream tasks. These tasks are instantiated as TDs linked to semantic descriptions of the BIM data and the technician. Through TDs, the technician can relate manual tasks to the prerequisite inferred by the reasoner, which also ensures transparency for humans to better understand the context of the engineering tasks. Every installation task committed (e.g., as a form of documentation through TDs using a WoT client application) and successfully passed by the CI modules triggers the BIM data to be updated with the Brick ontology [16], containing semantic descriptions of building components (e.g., the valve actuator or the room sensors). For example, once a radiator and its valve have been installed, then the pipeline clears the way for the actuator to be mounted and tested using their semantic descriptions.

Pre-commissioning: The functionality of peripheral devices such as sensors and actuators, connectivity at various network levels are tested and confirmed. The domain experts define checklists and steps to verify dependencies and constraints of the devices. Pre-commissioning of a room automation implementation includes the equipment in the room (i.e., radiators, lights, blinds, etc.), and the associated sensors, actuators, automation controllers, as well as the network to which they are connected. The pipeline supports

verification of required measurements and identification of sensors by creating a visual inspection task for the technician or using a temperature change test. The TDs of the devices is also generated in this phase to describe how to obtain the sensor measurements – an automated test agent in the role of a WoT client application would be able to interact with the devices using the TDs to manage the tasks.

Commissioning: The commissioning phase implements process goals into steps in the control software. Intention of the control software can be verified by obtaining the functional description of the control logic and its interfaces [17] through the TDs. Alternatively, akin to software unit testing, virtual simulation environments can be used to validate the control software [18]. For example, the control software for a room automation can be verified to see if it binds to the right sensors and actuators, and has functions like temperature or ventilation control loop to satisfy the thermal and air-quality requirements present in the semantic data. The commissioning completes once all the artifact committed to the repository passed the tests defined by the domain experts.

Operation: The BA system starts its operating after the previous engineering phases. Operation of BA systems includes monitoring using automated FDD which is aware of the construction, the control software, and the process goal that has gone through the pipelines, and hence provides inputs for tracing the provenance for DPs used in data-driven ML-based applications. Such smart applications continuously verify the consistency between BA and requirements, and raise issues in the pipeline when anomalies are detected. To resolve an issue, the engineers can also trace the tests verified in any of the previous phases, which support the domain experts and the technicians to identify the BA components related to the anomaly detection using the data provenance described in the semantic data of DPs. Distinguishing from explainability of ML models, which is typically achieved by explaining the predictions of classifiers (e.g., LIME [15]), semantic data integration provides fine-grained insight into CPS from an understanding of the BA infrastructure.

3.2 Rule-based Semantic Reasoner

During each BA engineering phase, the BIM data is annotated with links to requirements using semantic data [5]. Semantic data can be stored as RDF triples in a serialized format such as XML, Turtle, or using the Web Ontology Language (OWL). A BA domain ontology defines functional requirements of BA components on top of other ontologies such as Brick and ifcOWL [2]. Concretely, domain experts assign functions (e.g., lighting and HVAC conditions) to spaces (e.g., a room), by editing the BA ontology through the BIM development software, which also generates the updated ontology in a serialized format.

To infer missing (engineering) requirements for the assigned functionality, the functional requirements can be encoded as constraints – Description Logics (DL) using the Semantic Web Rules Language (SWRL) [13]. An OWL DL reasoning tool extended for SWRL (e.g., Pellet [19]) then provides the rule-based reasoning service as a CI module to validate if all the required information is stored in the repository to complete the engineering phase. The CI/CD pipeline integrates such rule-based semantic reasoning as a

CI module implemented as scripts [12] and enables an incremental and iterative development workflow throughout the different BA engineering phases.

Semantic reasoning can be used to automate service composition in a goal-driven way [10]. The semantic reasoner infers logical consequences from the semantic requirement specifications, allowing to infer potential compositions of building components as well as their engineering tasks. For example, given the requirement to maintain thermal and air-quality comfort in a room, the planner and HVAC experts can submit the plan to the pipeline, which then will invoke the reasoner to use the description of the construction and verify if equipment for heating, cooling, and ventilation are available and if they satisfy the need for thermal and ventilation conditioning.

Validation achieved by the rule-based reasoning includes not only the integrity of the required engineering tasks, but also the dependency resolution among BA components. For example, when cooling and humidification are assigned to a room, an air conditioner and a humidifier are to be installed in the room, which influence the control strategy of each other. A CI module for virtual engineering with physical simulation model [18] verifies such requirements that dynamically change during the planning phase.

4 EVALUATION METHOD

To evaluate effectiveness of BA engineering through the CI/CD pipeline with the rule-based reasoner, we plan to compare the development workflow of the proposed approach with an alternative method using the Business Process Model and Notation (BPMN) standard, for instance. While BPMN can be used for coordination of human tasks (e.g., [4]), the standard lacks conceptual streamlining in semantically relevant concepts and a seamless systematic mechanism for refinement from conceptual to executable BPMN models [3]. Our approach leverages the rule-based reasoners to reduce interdefinitions of semantic descriptions and the plethora of existing DevOps tool-chain to realize systematic workflow for the coordination of engineering processes within a predictable total cycle time, which is achieved by automatic verification implemented as CI modules.

5 CONCLUSION

Documentation about BA engineering process contains numerous ambiguities and incompleteness in the description of semantically related concepts, leaving room for incompatible interpretations in the design, analysis, and use of business processes towards reliable intelligent BA systems with transparency. In the heterogeneous and complex ecosystem where BA engineering takes place, we proposed to integrate the potential of using semantic knowledge-base for semi-automated reasoning by the DevOps concept. The reasoner serves as the foundation for automated verification and identification of dependencies and allows the CI/CD pipeline to access them during the BA engineering tasks. As an avenue for future work, we continue to implement and deploy this idea in an active industry research, and aim to collaborate with other researchers in the area of semantic modeling and data-driven FDD applications which use ML models for fault classification and root-cause identification.

REFERENCES

- [1] Sridhar Alla and Suman Kalyan Adari. 2021. What Is MLOps? In *Beginning MLOps with MLFlow*. Springer, 79–124. https://doi.org/10.1007/978-1-4842-6549-9_3
- [2] Jakob Beetz, Jos Van Leeuwen, and Bauke De Vries. 2009. IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Ai Edam* 23, 1 (2009), 89–101. <https://doi.org/10.1017/S0890060409000122>
- [3] Egon Börger. 2012. Approaches to modeling business processes: a critical analysis of BPMN, workflow patterns and YAWL. *Software & Systems Modeling* 11, 3 (2012), 305–318. <https://doi.org/10.1007/s10270-011-0214-z>
- [4] Marco Brambilla, Piero Fraternali, and Carmen Vaca. 2012. BPMN and design patterns for engineering social BPM solutions. In *Business Process Management Workshops*. Springer Berlin Heidelberg, 219–230. https://doi.org/10.1007/978-3-642-28108-2_22
- [5] Björn Butzin, Frank Golatowski, Christoph Niedermeier, Norbert Vicari, and Egon Wuchner. 2014. A model based development approach for building automation systems. *19th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2014* (2014). <https://doi.org/10.1109/ETFA.2014.7005365>
- [6] Henrik Dibowski, Joern Ploennigs, and Klaus Kabitzsch. 2009. Automated design of building automation systems. *IEEE Transactions on Industrial Electronics* 57, 11 (2009), 3606–3613. <https://doi.org/10.1109/TIE.2009.2032209>
- [7] Andrej Dyck, Ralf Penners, and Horst Lichter. 2015. Towards definitions for release engineering and DevOps. In *2015 IEEE/ACM 3rd International Workshop on Release Engineering*. 3–3. <https://doi.org/10.1109/RELENG.2015.10>
- [8] Georgios Gousios, Martin Pinzger, and Arie van Deursen. 2014. An Exploratory Study of the Pull-Based Software Development Model. In *Proceedings of the 36th International Conference on Software Engineering (ICSE 2014)*. Association for Computing Machinery, 345–355. <https://doi.org/10.1145/2568225.2568260>
- [9] H. R. Kranz. 2008. Initial and Continuous Commissioning of Building Automation and Control Systems (BACS) -Preview EN ISO 16484-. <https://hdl.handle.net/1969.1/148654> (Accessed on 17/10/2021).
- [10] Simon Mayer, Dominic Plangger, Florian Michahelles, and Simon Rothfuss. 2016. UberManufacturing: A Goal-Driven Collaborative Industrial Manufacturing Marketplace. In *Proceedings of the 6th International Conference on the Internet of Things (Stuttgart, Germany) (IoT'16)*. Association for Computing Machinery, 111–119. <https://doi.org/10.1145/2991561.2991569>
- [11] Iori Mizutani, Jonas Brütsch, and Simon Mayer. 2021. Towards Provenance Integration for Field Devices in Industrial IoT Systems. In *Provenance and Annotation of Data and Processes*. Springer International Publishing, 250–255. https://doi.org/10.1007/978-3-030-80960-7_21
- [12] Iori Mizutani, Ganesh Ramanathan, and Simon Mayer. 2021. Integrating Multi-Disciplinary Offline and Online Engineering in Industrial Cyber-Physical Systems through DevOps. In *11th International Conference on the Internet of Things*.
- [13] Bijan Parsia, Evren Sirin, Bernardo Cuenca Grau, Edna Ruckhaus, and Daniel Hewlett. 2005. Cautiously approaching SWRL. , 30 pages. https://cs.uwaterloo.ca/~gwendell/cs848/SWRL_Parsia_et_al.pdf (Accessed on 10/17/2021).
- [14] Pieter Pauwels and Walter Terkaj. 2016. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in construction* 63 (2016), 100–133. <https://doi.org/10.1016/j.autcon.2015.12.003>
- [15] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. “Why should I trust you?” Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- [16] Brick Schema. [n. d.], BrickSchema/Brick: Uniform metadata schema for buildings. <https://github.com/BrickSchema/brick> (Accessed on 10/17/2021).
- [17] Georg Ferdinand Schneider, Pieter Pauwels, and Simone Steiger. 2017. Ontology-Based Modeling of Control Logic in Building Automation Systems. *IEEE Transactions on Industrial Informatics* 13, 6 (2017). <https://doi.org/10.1109/TII.2017.2743221>
- [18] Georg Ferdinand Schneider, Hendro Wicaksono, and Jivka Ovtcharova. 2019. Virtual engineering of cyber-physical automation systems: The case of control logic. *Advanced Engineering Informatics* 39 (2019). <https://doi.org/10.1016/j.aei.2018.11.009>
- [19] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. 2007. Pellet: A practical owl-dl reasoner. *Journal of Web Semantics* 5, 2 (2007), 51–53. <https://doi.org/10.1016/j.websem.2007.03.004>
- [20] Seshadhri Srinivasan, Pandarasamy Arjunan, Baihong Jin, Alberto L. Sangiovanni-Vincentelli, Zuraimi Sultan, and Kameshwar Poolla. 2021. Explainable AI for Chiller Fault-Detection Systems: Gaining Human Trust. *Computer* 54, 10 (2021), 60–68. <https://doi.org/10.1109/MC.2021.3071551>
- [21] Walter Terkaj, Georg Ferdinand Schneider, and Pieter Pauwels. 2017. Reusing domain ontologies in linked building data: the case of building automation and control. In *8th International Workshop on Formal Ontologies meet Industry*. <http://hdl.handle.net/1854/LU-8578705>
- [22] Baifan Zhou, Yulia Svetashova, Andre Gusmao, Ahmet Soylu, Gong Cheng, Ralf Mikut, Arild Waaler, and Evgeny Kharlamov. 2021. SemML: Facilitating development of ML models for condition monitoring with semantics. *Journal of Web Semantics* (2021), 100664. <https://doi.org/10.1016/j.websem.2021.100664>