Information Integration and Semantic Interpretation for Building Energy System Operation and Maintenance

Hervé Pruvost, Olaf Enge-Rosenblatt and Jürgen Haufe
Division Engineering of Adaptive Systems EAS
Fraunhofer Institute for Integrated Circuits IIS
Dresden, Germany
herve.pruvost@eas.iis.fraunhofer.de

***Abstract*—Digitalization in facility management has become an active field of research for many reasons. Among others, it can provide high support by saving energy and cost in the building sector as well as for optimizing internal processes involved within the management of a facility. To enable this, many data about a building has to be collected and analyzed. Building management systems can already produce huge amounts of data that are analyzed for supervising, controlling and benchmarking buildings. However, even if numerous data are produced during building operation there is no much use of building information created during its design. In view of that, this research proposes a methodology and a software framework for closing the informational gaps between building design and operation. It follows data integration steps that use initial building information models and extend them with operation data as well as semantic models. These data models are used for interpreting energy system behaviors and performing predictive analyses. More specifically, a system ontology is introduced that supports reuse of knowledge for optimized building operation and maintenance.**

Keywords—building information modeling, knowledge management, maintenance, building operation, data analysis

I. Introduction

The digitalization process keeps rising in the building industry promising new perspectives for better planning and operating buildings throughout their whole life cycle. In this context, different ICT technologies are utilized in the research field as well as in industry that allow for the collection and analysis of data related to a building and its technical services. Data collection appears at different stages of the life cycle of a building and has accordingly different purposes. During building operation, building management systems (BMS) are usually in charge of gathering information about a building which are for a part related to its energy behavior. BMS data are gained through sensors and meters that provide information about e.g. the operational states of technical equipment, indoor temperature or energy consumption. Because of its highly time-dependent nature, this kind of information can be categorized as dynamic data about a building. In addition, BMS also uses static data which represent the building and its technical systems as they are i.e. as built physical entities. This information encompasses the energy system components, their technical characteristics and their layout in the building. This static information is intrinsically associated to dynamic data which can provide insights into the operational behavior of the physical system. In that context, digitalization can be somehow measured as the extent of static and dynamic information which is collected about a building. The actual trend shows an exponential increase in the use of dynamic data in building operation. In the actual state of technics, monitoring systems are able to provide very high amounts of data which are managed and analyzed with the help of established methods (cf. Big Data, Internet of Things).

Unlike the building operation phase, the design phase focuses on the collection respectively creation of static data. Digitalization in building design is nowadays associated to the field of building information modeling (BIM) [1] which has brought a set of data standards as well as a methodology for enhancing collaboration and interoperability by building projects. The emancipation and gradual establishment of BIM in the design phase of buildings results in the production of many information of different nature. Indeed, more than only the geometrical model of a building, BIM enables the enrichment of building data with e.g. technical specifications, analysis models e.g. for energy analysis and cost analysis, production schedules, etc. This information integration process is also known as nD modeling [2].

Despite the bright source of information offered by BIM, there still exists a large informational gap between the design and the operation phase. In practice, there is no reuse of BIM design information during the operation phase. Even if the design phase of a building engages much information and relies on high quality CAD models, the transfer of all created information to operation actors and systems is either poor or inexistent. It is generally recognized that software systems embedded into BMS or developed for computer-aided facility management (CAFM) could take benefit of such a stock of information, but in practice they still rely instead on well-delineated and proprietary data models [3]. Those data models necessitate manual efforts in the commissioning phase for configuring them and gathering all static information again. Related software systems are specifically customized for each

978-1-5090-6683-4/18/$31.00 ©2018 IEEE

new building. By their first configuration, empty data points are instantiated in some database memory units which must be then all manually initialized and associated with building layout drawings or mechanical, electrical and plumbing (MEP) schematics. As a matter of fact, the static information used by such systems is much more limited in scope than BIM information.

In view of this informational gap between building design and operation, the purpose of this article is to propose a methodology for reusing BIM data during building operation. In this context, an information integration and semantic enrichment process is formalized as a BIM process. Moreover, analysis use cases are introduced which focus on predictive maintenance. They rely on the analysis of operation data which, combined with BIM information, provides necessary outcomes for assessing energy system operational states and behavior. In view of that, it is explained how energy system operation and maintenance can be supported by BIM standards. Furthermore, the followed approach puts accent on the reuse of engineering knowledge that can be applied for planning FM actions and configuring BMS. That way, knowledge complements initial BIM static and dynamic data thus enriching it with engineering best practice.

II. Facility Management Information Workflow

The building life cycle involves different stakeholders who play specific roles and have their own impact on the operation of a building. Some main actors include building owners, users, facility managers, designers and engineers. They all use and share certain amounts of information that practically all belong to an overall workflow. For a rational use of information during building operation, all data about the building produced during design and commissioning phases shall be used as basis. More than a pure geometrical model, parts of the available information stock should include among others technical specifications, requirements and simulation results. Operating a building can then follow a BIM-centered workflow as the one from Fig. 1 expressed with the business process model and notation (BPMN) standard. In the diagram, the rows define interacting domains that represent the stakeholders' roles mentioned above. The additional BIM domain in the middle is responsible for building information management. This workflow is defined as a continuous and iterative process describing different tasks that are performed manually for a part by FM actors and automatically for another part by software applications or services. The resulting framework aims at automatizing building operation and maintenance as much as possible while maximizing the reuse of BIM information.

First of all, based on the technical characteristics of the building energy system and the performance requirements of users and owners, the FM domain accordingly configures the energy system and applies specific settings like e.g. set points. Requirements and settings can be gathered into a key metric model containing reference "to be" quantitative values (indoor temperature, supply temperature, nominal power of components, efficiency rate, etc.) and key performance indicators (KPI). The BMS fulfills its role of continuously collecting operation data, referred to as building automation and control system (BACS) data in Fig. 1, and storing them in a dedicated database. The number of data points available reflects a certain level of monitoring (LoM) that conditions the feasible analyses. An important step in reusing design information is to associate building data from CAD design with operation data so that the BIM model does not remain a static building system representation but becomes a dynamic information model. Then, the BIM model is converted into an ontology using an existing semantic web tool [4]. The use of an ontology enables to semantically enrich the BIM model later on in order to assist the targeted analyses and provide support for FM decisions. This new BIM system ontology is named

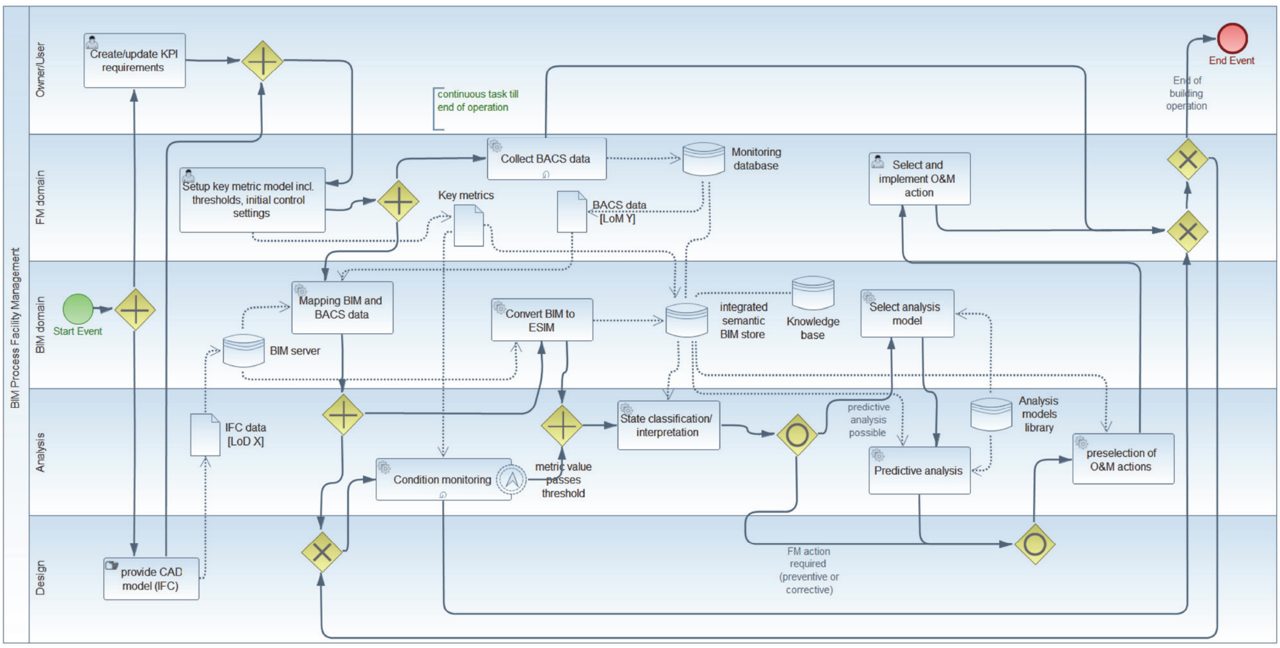

Fig. 1. BIM process for facility management

energy system information model (ESIM) as it is centered on the energy system. The ESIM ontology is complemented by additional information from heterogeneous sources and acts as an integrated semantic BIM model. Part of additional information is characterized by knowledge models that are useful for analyzing, interpreting and curing the energy system operational state. In practice, this knowledge can be generic i.e. independent from buildings and FM stakeholders (stored in BIM domain), but also firm-specific (contained in FM domain or in analysis domain). Using that knowledge and BACS data, an inference engine can then classify and interpret energy system operational states. By need and if corresponding analysis models are available, a predictive analysis can be performed which shall give insights into the aging and reliability of energy system components. Finally, on the basis of the identified operational states and the analysis results, predefined operation and maintenance measures can be retrieved from the knowledge base for supporting FM decisions.

III. Integration of Static and Dynamic Data

A. Integration of design and operation data

An efficient use of building information models within the operation phase relies on their association with collected operation data. This association opens a certain number of possibilities. In our case, it shall give support for performing data-driven analyses in the context of energy efficiency analysis, aging and failure analysis, as well as decision making. For coupling building information models and real data, we apply a procedure that enables direct reuse of building CAD design models and semi-automatic linking of both static and dynamic data. In that context, it is assumed that CAD models produced within the design phase reach a certain degree of quality and a sufficient level of development [5]. Indeed, in building design different disciplines are involved in the planning process like architectural, structural, MEP and BACS design. This latter discipline, which plays a major role in rather late design or in commissioning stages, shall provide sufficient information about the building automation system (e.g. sensors, meters and controllers) in order to map it with monitoring data.

For the integration of static and building data we rely on the use of the open BIM standard IFC (Industry Foundation Classes) [6]. This choice is made to ensure a maximum interoperability with usual CAD programs used by practitioners. Indeed, the IFC format is commonly supported by most CAD vendors. The IFC data specification covers a wide variety of domains and project relevant information in a single data format. It is divided in several hierarchical class layers and domains. For our purpose, we focus on the so called building controls domain which provides several concepts related to the field of building automation. More specifically, it defines data objects for representing e.g. sensors, meters and controllers. For modeling such technical components in a CAD model, one can utilize BIM component catalogues. Many BACS components from several manufacturers are represented and available in such catalogues. Nevertheless, many specific components still do not exist. Moreover, because each BIM catalogue product disposes of its own custom set of properties

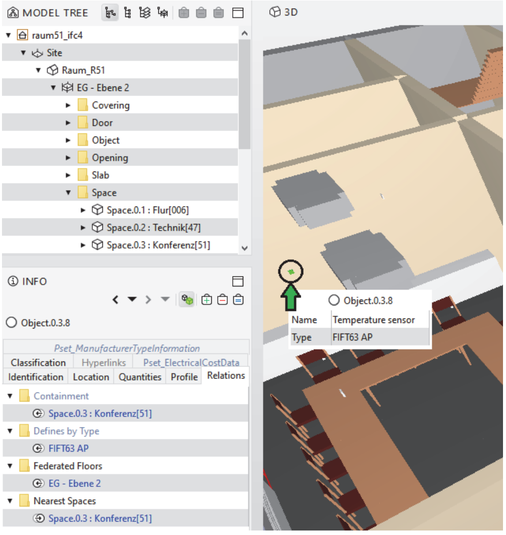

Fig. 2. View of an IFC 3D CAD model containing a temperature sensor and information about its relations to building topology.

(e.g. name, manufacturer, manufacturer reference…), we define higher-level BIM entities i.e. templates that represent a certain category of components and that dispose of a common set of properties. For the example of sensors, we define a specific sensor template which contains general properties. Among these properties, some are defined for mapping purpose with the monitoring database while some others are defined for proper representation of the sensor in the IFC data format. More precisely, we use the IFC classes *IfcSensor*, *IfcSensorType* and *IfcSensorTypeEnum* [6]. In practice, each virtual sensor, as the one shown in Fig. 2, refers to some template and disposes of a valid representation in IFC.

Because the IFC are not sufficient and adapted to represent dynamic information, we couple the CAD model with a monitoring relational database that is used to collect all data related to building operation from a BMS. This data are usually collected in the form of time series with different time intervals and additional metadata about their type, unit, related data device and position in the building. Because both virtual sensors from the CAD model and data points in the database share similar information, an internal database service performs then an automatic mapping to create explicit associations between static and dynamic data. More concretely, the mapping is based on the relations the sensor data points have with technical components as well as with building topology. As a result, the data points get linked to virtual sensors from the IFC model. For that purpose the IFC globally unique identifier (GUID) are used as association keys.

B. Integration of Additional heterogeneous Information

In addition to mapping virtual building models with their operation data, further data integration steps are necessary for allowing and especially automating the targeted analyses. Indeed, there are several reasons why integration of additional information is of interest. First of all, even if the IFC schema specification covers a wide variety of information domains, these still are limited in scope, hence extensions are necessary for performing specific analyses. As an example, in the field of

BIM-based energy analysis several works like in [7] have demonstrated that additional non-IFC information about e.g. occupancy, climate forecasts or energy behavior of HVAC components is necessary to make reliable predictions. The IFC schema is not designed to represent such kinds of information.

In a similar manner and for the purpose of BIM-based operation and maintenance, we aim in our research at extending the IFC information space for following reasons:

- For fastening and automating predictive analyses within FM, we couple the CAD model with a library of analysis models. These models consist of aging models for different HVAC components and reliability models.
- We integrate also a knowledge model that provides support in analysis of the operational state of a building and in planning of FM actions. This model covers three main information areas:
    - a classification of operational and behavioral states of building energy systems and their components which are associated to key metrics,
    - causality relations between system and component states,
    - operation and maintenance templates that describe predefined classes of FM measures like maintenance plans or control strategies.
- For managing previously mentioned heterogeneous information and for driving the process described in Chapter 2, there is a need for a metamodel that (1) represents all involved information at a more abstract level, (2) that provides means of interlinking it and (3) that defines rules and constraints for processing it (e.g. which sensor data are required for which analysis model at which building information level of detail?).

For the above mentioned purposes, we introduce an ontological data structure which is implemented in the energy system information model (ESIM) mentioned earlier and described in next chapter.

IV. Ontology-Based Energy System Representation

The energy system information model (ESIM) is a domain-specific model that represents energy systems at building level as well as at urban level, and that includes automation and control equipment. It has been initially developed in the scope of the EU project eeEmbedded [8]. It provides a data structure that enables functional, structural and physical descriptions of entire energy systems. Fig. 3 shows the core superclasses and property groups of the ESIM. Besides the description of complex systems, it has been also developed for enabling integration of further system-related information. Concretely, the ESIM is built as an ontology that applies semantic web standards [9] like RDF/OWL. Beside their functionalities, these standards can guaranty interoperability with several systems in future applications of the information model. Fig. 4 shows the five main functions of the ESIM. First, this information model is meant to enable a comprehensive system description which encompasses the physical components, their hierarchy and their

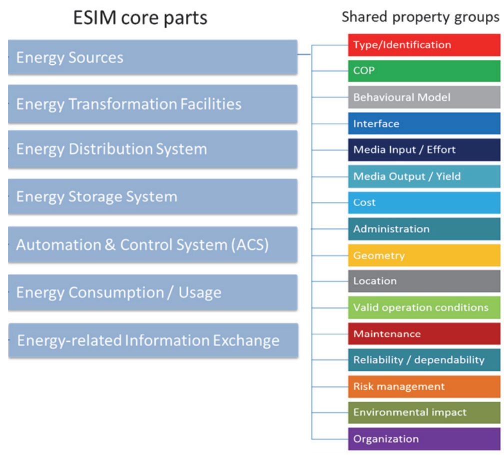

Fig. 3. Core taxonomic parts and property groups of the energy system information model [8].

relationships in terms of physical and communication interactions. With a similar goal, parallel researches try to establish information models like e.g. IEC Common Information Model (CIM), Smart Grid Architecture Model (SGAM) or Facility Smart Grid Information Model (FSGIM) [10]. These models have been developed for information exchange among energy management systems and stakeholders of the energy market. However, they focus more exclusively on electrical networks. As a second function, the ESIM model is built for integrating other data structures by the means of links. Especially, it allows a full integration of the IFC data structure by using the ifcOWL ontology format which belongs to the open BIM standards [4]. That way, the ESIM ontology can provide extensions to the initial IFC CAD model developed within the design phase. Among these extensions, the ESIM ontology can be interlinked with a requirement model including key metrics (KPIs, normative restrictions, control settings, reference values and thresholds) as well as a cost model. For the sake of data integration again, the third and fourth functions of the ESIM cover respectively the integration of analysis models and operation data. Analysis models are organized in a library that includes different computational algorithms about e.g. energy behavior, aging and failures of MEP components. As those algorithms are data-driven and rely on building operation data, the ESIM provides an interface to the monitoring database. This relies on the mapping described in Chapter 3 between the IFC CAD model and data points. The ESIM ontology includes additional metadata to sensor templates like e.g. time intervals which are missing in native IFC. Templates compose the fifth function and represent whether abstract or pre-defined energy system components and

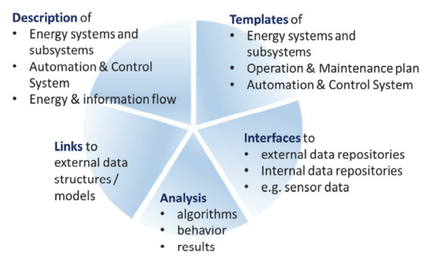

Fig. 4. Five main functions of the energy system information model [8].

BACS elements as well as O&M strategies and procedures. They are used for modular and fastened design of energy and BACS systems as well as for supporting FM decisions.

V. Energy System State Analysis And Prediction

Through several steps of the FM BIM process presented in Chapter 2, the ESIM model is configured as an integrated BIM model. For that purpose, building information from initial CAD model is associated to related operation data and then converted into a system ontology described by the ESIM. For performing the further tasks presented next, the ESIM integrates also knowledge and analysis models. These tasks consist of (1) the identification of energy system operational states, (2) a predictive analysis about future outcomes of such states and (3) a contingency plan implmented as FM actions.

A. System state classification and interpretation

For assessing the operational state and behavior of the building energy system, we rely on expert knowledge. This knowledge reflects facility managers' and engineers' know-how and experience about technical building systems. It is formalized and contained in a knowledge model that is associated to the ESIM ontology. The retrieval and application of that knowledge relies on semantic reasoning. For that purpose, we apply semantic web reasoning technics [9]. Practically, an inference engine takes facts about the energy system as input and produces statements as output on the basis of the expert knowledge. The facts are provided by the ESIM ontology and are of two kinds. Static facts consist of building information containing the type of technical systems, their description and interactions. Dynamic facts consists of operation data that are handled in condition monitoring. The purpose of the condition monitoring task is to analyze the evolution of different data points and identify certain changes in energy system behavior. The analysis of these changes relies on the key metric model previsouly introduced. This model describes several variables and measurements that have a strategic character and that are associated with a thresholds system. As illustrated in Fig. 5, when a metric passes a threshold value, a changed condition can be identified and by need a preventive or corrective O&M action can be triggered. These metrics can be derived directly from raw monitoring data (e.g. temperature, rotation speed, vibrations…) or indirectly computed (e.g. efficiency rate). Alternatively, metrics and metric ranges may also be computed by data analysis algorithms using e.g. machine learning methods like clustering by complex and ambiguous data sets.

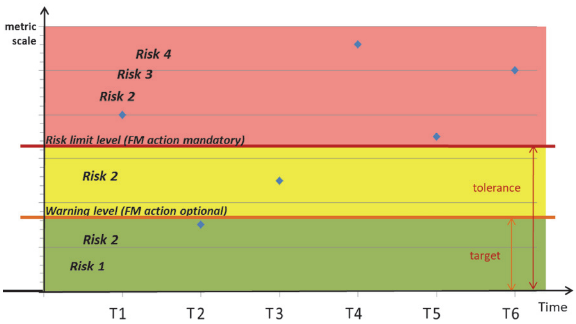

Fig. 5. Association of risk states with several metric values using two thresholds: risk tolerance and risk target.

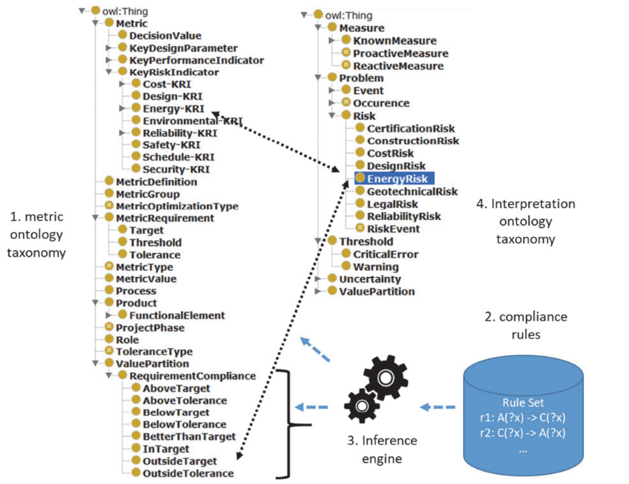

Fig. 6. Ontology classes structure of the key metric and the interpretation knowledge models.

These methods represent further development steps of the presented research. As for the example of risk events which is illustrated in Fig. 5 and 6, different operation risks are catalogued and grouped in an interpretation ontology using the RDF data standard and more precisely its overlying Web Ontology Language (OWL). More specifically, they are associated to different metric ranges using the ontology modeling principle of value partition. As illustrated, each metric can refer to several risks according to the range its value belongs to. That way, each energy system condition can be automatically classified into specific risks and criticalities. In that case, the key metrics are used as main knowledge access points. For performing this state classification, a set of rules are defined using the semantic web rule langage (SWRL) and executed by the inference engine on the basis of the facts provided by the integrated semantic BIM model. Using this approach different risks and conditions can be identified like excessive stress, aging or energy inefficient usage.

B. Predictive analysis

According to the system states interpreted within the previous task a predictive analysis can be necessary. In case of predictive maintenance, the overall aim is to allow for an estimation of the remaining working time in a certain efficiency range or before failure of a component or a system. For that purpose, the BIM process relies on a library of analysis models. Each analysis model has a certain function and implements a specific mathematical approach e.g. a physical aging model or a stochastic model. They are associated to certain types of components or sub-systems that are formalized in the ESIM taxonomy. The selection of an analysis model depends then on the component or system of interest, its identified condition as well as some informational constraints like the level of monitoring (LoM) and the level of detail (LoD) of the building information model. As stated by [11], there exist two causes of failure in reliability theory: aging and excessive stress. As a result of the combination of both phenomena, failures can be differentiated into three categories: gradual failures, aging failures and sudden failures. Gradual failure propagation may be directly measured by one or several indicators (condition monitoring - CM). In that case no subsequent analysis model is necessary. Nevertheless, it requires a certain LoM for the CM task to be performed i.e. the related data points must be available in the monitoring database. If the LoM is not sufficient, the failure might be simulated through a stochastic model. By aging failures, the failure probability is age dependent, that is, there is a predictable wear-out limit which can be computed by an aging model. In contrast, sudden failures are of complete randomness. Such failures cannot be predicted by condition monitoring and aging analysis. In that case, components' failures can be simulated using stochastic models in which time to failure is described by e.g. an exponential or a Weibull distribution. As for components, an analysis model can be used at system scale. In that case, the computed failures of single components can be used as inputs for analyzing their effects on the whole system. As shown in Fig. 7, this can rely on the system graph that is intrinsically provided by the ESIM ontology and based on IFC. To deduce failure occurrences of the whole system, it is necessary to describe the interactions between its components. For that purpose, an ontological representation is composed at instance level of the HVAC model contained in the ESIM and, at an abstract level, of a reliability knowledge model that describes known failure dependencies according to component interaction types.

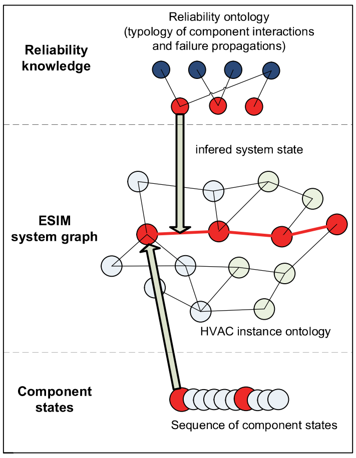

Fig. 7. Exploitation of the ESIM system graph for analyzing failures and malfunctions at system level based on system reliability knowledge.

C. *Knowledge-Based O&M Decision Support*

The last task of the BIM process consists of supporting the facility manager in taking decisions for an optimized maintenance and a more efficient use of energy. As for the system state identification, this task relies on rule-based semantic reasoning that takes the outcomes of the previous tasks as input for pre-selecting best fitting O&M measures. These measures are formalized and predefined as knowledge templates in the knowledge model. The O&M templates are described by several properties like the following:

- Type: preventive or corrective,
- Action: repair, replace, exhaust…
- Used resources: technical parts, men power, costs…
- Maintenance schedule: period, sequence of actions...
- Control settings: temperature set points, AHU settings (pre-heater, pre-cooler, valve positions)…

As each building use case is specific and unique, the knowledge templates mainly represent types of O&M measures which are pre-instantiated after semantic reasoning. They are only pre-selected according to some criteria as possible FM action candidates that can then be further elaborated and implemented by FM managers.

VI. Conclusion

The presented methodology intends to bridge building design and building operation with the support of digitalization. For that purpose it relies on BIM standards and proposes a building operation workflow in which several FM stakeholder roles interact through a software framework. As central component, an integrated semantic BIM model is built as an ontology and is used for data integration, knowledge management as well as energy system behavior analysis. Some parts of the framework are at prototypical state while others are at conceptual state and do not exist yet.

Acknowledgment

This research work has been enabled by the German Federal Ministry for Economic Affairs and Energy (project FMopt, Id-No. 03ET1372A) and the European Commission (project eeEmbedded, Id-No. 609349).

References

[1] C. Eastman, P. Teicholz, R. Sacks and K. Liston, "BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors", 2nd Edition, John Wiley & Sons, ISBN 978-0-470-54137-1, 648 pages, July 2011.

[2] A. Lee, S. Wu, A. Marshall-Ponting, G. Aouad, R. Cooper, J.H.M. Tah, C. Abbott and P.S. Barrett, "nD modelling road map: A vision for nD-enabled construction", University of Salford, 2005.

[3] P. Teicholz, ed., "BIM for facility managers", John Wiley & Sons, 2013.

[4] P. Pauwels and W. Terkaj, "EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology", in Automation in Construction 63, pp. 100-133, Elsevier, 2016.

[5] BIMForum, "Level of Development Specification for Building Information Models, Version 2017", available online: www.bimforum.org/lod (accessed 22nd June 2018).

[6] buildingSMART International Ltd. "Industry Foundation Classes Release 4 (IFC4)", available online: http://www.buildingsmart-tech.org/ifc/IFC4/final/html/ (Accessed 20th January 2016).

[7] P. Stenzel, J. Haufe and N. Jimenez, „Using a Multi-Model for a BIM-based Design and Operation of Building Energy Management Systems", 10th European Conference on Product and Process Modelling (ECPPM 2014), Vienna, Austria, Sep. 2014].

[8] J. Kaiser and P. Stenzel, "ESIM Multi Model", public project deliverable D4.2, © eeEmbedded Consortium, Brussels, 2015.

[9] W3C, "Semantic Web: Linked Data", available online: http://www.w3.org/standards/semanticweb/data (accessed 18th May 2017).

[10] D.S. Shafiullah, T.H. Vo, P.H. Nguyen. and A.J.M. Pemen, "Different smart grid frameworks in context of smart neighborhood: A review", 52nd International Universities Power Engineering Conference (UPEC 2017), IEEE, 2017.

[11] M. Rausand and A. Høyland, "System reliability theory: models, statistical methods, and applications", second edition, vol. 396, John Wiley & Sons, 2004.