Semantic Interoperability to Enable Smart, Grid-Interactive Efficient Buildings

Harry Bergmann, U.S. Department of Energy
Cory Mosiman, Avijit Saha, Selam Haile, William Livingood, National Renewable Energy Laboratory
Steven Bushby, National Institute of Standards and Technology
Gabe Fierro, University of California Berkeley
Joel Bender, Cornell University
Michael Poplawski, Pacific Northwest National Laboratory
Jessica Granderson, Marco Pritoni, Lawrence Berkeley National Laboratory

ABSTRACT

Achieving a widespread transition to grid-interactive, efficient buildings (GEBs) depends critically on there being sufficient interoperability among connected building systems. While many critical elements already exist at the technical interoperability level (TCP/IP, BACnet, etc.), a lack of interoperability in the semantic level hinders streamlined integration of interdependent applications. Semantics refers to expressing information about “things” in a way that can be consistently understood by applications. Key components of formalized semantics include identifying what a “thing” is (its “type”), defining general information about that “thing” (its characteristics or properties), and defining the appropriate relationships of that “thing” to other “things” (its function or role in a larger system). Although this might seem initially trivial, the success of smart building applications is highly dependent on maintaining consistent self-descriptive notions of the “things”.

Without semantic interoperability, it is technically difficult, labor-intensive, and cost-prohibitive to enable three key objectives of GEBs: optimizing performance, automatically identifying and diagnosing faults, and delivering grid services. Industry, academia, and standards bodies have invested effort in developing information models to facilitate semantic interoperability, however, they have not been widely adopted across the U.S. commercial building portfolio.

This paper will present a pathway to drive semantic interoperability through a three-pronged approach to be led by the DOE Building Technologies Office in partnership with NIST and multiple national laboratories comprising: 1) industry engagement and coordination across existing efforts; 2) a semantic interoperability standard that empowers building owners to identify and require interoperable attributes when procuring equipment and applications; 3) tools to assist in implementation and a test framework to verify compliance of products with semantic interoperability specifications. This approach is designed to accelerate the timeline for adoption of semantic interoperability specifications. The intent is to reduce soft costs associated with implementing advanced controls, fault detection and diagnostics, and other smart building technologies and use cases as a necessary step in achieving an energy efficient smart grid future.

Introduction

Technical innovations in the building industry tend to lag the technology advancements of the information age. This is due, in part, to the fact that the building automation market is small compared to the market for information technology in general, and thus it is not a driver for technology change. It is also due to the fact that investments in building infrastructure are expected to have a useful life of a few decades instead of a few years like a personal computer or cell phone.

The development of low-cost microprocessors in the 1980s led to a long transition in buildings from pneumatic to digital controls. Early in that transition building owners found themselves locked into largely proprietary communications protocols and networking systems. BACnet ("ANSI/ASHRAE Standard 135-2016" 2020) was designed to solve this problem at the technical interoperability level (Hardin et al. 2015), and does so quite well for many building technologies. Properties (name, description, current value, units, etc.) of standardized objects (inputs, outputs, calculated values, schedules, etc.) can now be successfully communicated between devices of different manufacturing origin using standardized services. Traditionally stand-alone building systems such as heating, ventilating, and air conditioning (HVAC), lighting, access control, and fire protection can also be integrated. Other standardized technologies have also started to work their way into building devices and systems (e.g., IEEE 802.3 and 802.15.4). As a result, technical interoperability is no longer a significant challenge; therefore, the remainder of this paper will focus on the importance of structured metadata and the ability to ensure that the precise meaning of exchanged information is unambiguously interpretable by any other system.

Expanding on the BACnet example, while the 'name', 'units', 'value', and 'description' fields in a BACnet object are examples of low-level metadata that provide only limited information:

- Knowing units of Fahrenheit implies that a temperature is being measured, however the substance being measured (air, water), the general context of the measurement (in a duct, zone), or more precise contexts (Zone-12 on Level 3) are unknown,
- In many implementations, BACnet fields are left blank or incomplete. The low-level metadata is implicitly encoded using ad-hoc naming conventions that are inconsistently applied.
- Applying strict naming conventions can partly solve this problem but, in many buildings, this is not done well. The 'name' of an object could be "Bob." This is considered valid in BACnet, although we know this is not useful for understanding what this "thing" is.

Moreover, buildings are composed of highly interdependent systems and components. Enabling GEBs means that the complex notions of the relationships between building systems, their component technologies, their relative sequencing, and physical location need to be expressed cohesively. As demonstrated above, this cannot be accomplished by BACnet (or other building automation protocols) alone, which is why structured metadata is required.

The predominant operational metadata schemas for buildings are Project Haystack ("Project Haystack" 2020) and Brick Schema ("Brick Schema" 2020), informally referred to as Haystack and Brick. These schemas have been designed to add structure to low-level metadata typically found in building management systems (BMS). A key issue in the adoption of structured metadata is the process of transforming or translating existing metadata --- such as ad-hoc BMS labels --- to a structured form. One approach, generally known as 'point mapping', is primarily implemented today in supervisory energy management and information systems (EMIS) for automated fault detection and diagnostics (AFDD) applications.

Point mapping is largely unautomated due to the unreliability of the low-level metadata and is often implemented manually by an engineer. Previous research has demonstrated costs associated with this point mapping exercise to range from $230/point (up-front) to $1880/point (5-year ownership) (Granderson and Lin 2016). Other existing efforts attempt to streamline the point mapping process using machine learning techniques (Bhattacharya et al. 2015; Koh et al. 2018; Hong et al. 2015; Gao, Ploennigs, and Berges 2015; Balaji et al. 2015; Mishra et al. 2020) or inferencing frameworks (Fierro et al. 2019). The significant costs and difficulty of the point mapping approach act as key barriers to the adoption of smart building technologies.

Additionally, limited educational resources or tools exist to assist those performing point mapping, and there hasn't been any substantial or best practice documentation generated for Haystack and Brick. This is problematic in a few ways, namely:

- Due to the lack of verifiability of the correctness of the generated metadata and the significant customization of the metadata schema in Haystack, vendors or end users who want to work with a site using Haystack have little confidence in the correctness of its implementation. Inconsistency and customization using the Haystack data model have been previously demonstrated (Fierro et al. 2019).
- Semantic web technologies, utilized by Brick, are generally the domain of computer science experts. As the need for and accessibility of metadata extends to a larger workforce, better tooling is needed to enable widespread adoption of these technologies.

Furthermore, point mapping generally does not leverage metadata captured at other points in the building lifecycle. For example, in new construction projects, significant metadata about where things are located in space (e.g., an air handling unit on the roof) as well as more specific information generated during procurement and submittal reviews (part numbers, equipment capacities, etc.) are available. In existing buildings where this is not the case, energy audits are a way for experienced engineers to capture and document this information. These other processes in the building lifecycle also have defined mechanisms for capturing building metadata (as expanded upon in the following section), however, the development of these different metadata schemas have been largely uncoordinated efforts.

Ideally, building metadata would be captured and strategically maintained throughout the building lifecycle via complementary metadata schemas in order to support value-added use cases. Metadata would enable design teams to more seamlessly transition from the design of a building, through construction and commissioning, and into operation and verification of building performance. Metadata would also enable more efficient and less expensive auditing, analysis and implementation of energy conservation measure (ECM), and measurement and verification of energy savings without endless tailoring of metadata. Furthermore, it would link AFDD to system design information created by the mechanical engineer and implemented by the controls contractor, saving money and minimizing error. Currently, however, most of this workflow does not exist without significant amount of human labor.

This paper reviews recent advancements in metadata schemas for buildings and discusses an approach to leverage and integrate them into an implementable and verifiable standard. The core focus is to create a workflow and use case driven approach for defining semantic interoperability and to develop assisting tools, reference implementations, and a conformance evaluation framework for promoting widespread adoption. The hypothesis is that developing metadata standards in a workflow and use case centric fashion will significantly reduce the soft costs associated with implementing next generation building applications, for example:

- Automated fault detection and diagnostics,
- Building system commissioning,
- Control system implementation,
- Digital twins,
- Energy audits,
- Optimization of energy use, and
- Smart grid interactions.

The rest of the paper is organized as follows: a review of existing metadata schemas for buildings is presented, followed by a discussion of the philosophy of use case driven standards; next, a description of the proposed solution along with its intended impacts is provided; and, finally, the paper concludes with a summary and next steps.

Review of Existing Metadata Schemas for Buildings

Significant effort has already been invested to standardize building metadata during the building lifecycle. These efforts utilize different underlying technologies as a means for expressing their metadata schemas, the majority of them relying on the eXtensible Markup Language (XML) or semantic web technologies, specifically the Resource Description Framework (RDF), RDF schema (RDFS), and the Web Ontology Language (OWL). Although a detailed exploration as to the pros, cons, and critical differences between these technologies is outside the scope of this paper, they are briefly summarized as follows:

- XML is a hierarchically structured data model, where the structure can be verified by comparison to an XML Schema Definition (XSD). Information about things (concepts) discussed by the data model is generally fully contained by a single XML document.
- RDF is a more flexible graph structured data model, designed as a distributed knowledge-based system. With its use of cooperatively shared resource identifiers, information about entities can be defined in different places, further classified, and linked to one another, enabling knowledge to be built and expanded upon. The other aspects of the semantic web technologies extend this capability to organize concepts in a more object-oriented fashion (classes and properties – RDFS) with more formalized description logics (cardinality and allowable relationships – OWL).

In short, XML imposes structure on information using a nested tree structure and is designed to be self-contained; semantic web technologies utilize the distributed nature of the web to share resource graphs that define, reference, and expand upon information in a highly formalized semantic fashion. These notions become critically important when evaluating the metadata landscape for buildings provided in Table 1. Many of these schemas overlap in their domain and touch points in the building lifecycle. Complementary efforts are encouraged by semantic web technologies, as the schema technology was specifically designed to promote extensibility and semantic interoperability via the semantic web suite of tools. XML based schemas, however, do not offer the same benefits, since XML was primarily designed to be self-contained[1]. As XML is the technology used in the majority of building design related schemas,

---

[1] Although the XLink and XPointer specifications provide mechanisms for linking between XML documents, it was not part of the original design intent (Bray et al. 2008)

this becomes important to consider when evaluating the ease of concept alignment across schemas, specifically when transitioning from XML to semantic web. The simultaneous redundancy and misalignment of schemas requires human intervention to reinterpret and rebuild portions of metadata to support different smart building applications, say, when the building shifts from the design phase (Industry Foundation Classes (IFC)) to the operational phase (Haystack or Brick). As the majority of smart building applications and GEB use cases are operationally oriented, the two predominant operational metadata schemas for buildings (Haystack and Brick) are explored in more detail in the following two sections.

Table 1 Summary of Building Metadata Schemas and their Purposes

Name	Primary Domain	Schema Technology	Primary Purpose
Industry Foundation Classes (IFC)	Building Information Modeling (BIM)	XML Schema (XSD), has OWL representation	To enable designers (primarily engineers and architects) to exchange models between different BIM applications during the design phase.
Green Building XML (gbXML)	Building Energy Modeling (BEM)	XML Schema (XSD)	To enable BIM to BEM and BEM to BEM interoperability for the purposes of conveying system and building information needed to simulate building energy consumption.
Building Energy Data Exchange Specification (BEDES)	Multiple	Informal - XML or CSV download	To provide a standard set of terms to facilitate the exchange of information on building characteristics and energy use (“Building Energy Data Exchange Specification (BEDES) | BEDES” 2020)
BuildingSync	Commercial Energy Auditing / Operations	XML Schema (XSD)	To standardize the reporting format of information obtained through commercial building energy audits, namely, ASHRAE Standard 211, building on the terms defined by BEDES.
Green Button	Utility Data Access	XML Schema (XSD)	To provide utility customers (residential and commercial) access to their utility data in a standardized format.
DNAs Framework	Occupant Behavior	XML Schema (XSD)	To provide ontology for energy related occupant behavior.

Project Haystack 3.0	Operational Applications	Custom text format	To provide a standard set of terms for describing sites, equipment (primarily HVAC), points, and the relationships between them. Primarily focused on representing information from building management systems (BMS).
Brick	Operational Applications	RDF / RDFS / OWL	To provide a graph-oriented, standardized ontology for representing the physical, logical, and virtual assets in buildings and the relationships between them ("Brick Schema" 2020), primarily building on the terms defined by Project Haystack.
Project Haystack 4.0	Operational Applications	RDF / RDFS / OWL	To add formal mechanisms for defining terms, a taxonomy, and an ontology on top of Haystack 3.0.
Building Topology Ontology (BOT)	Operational Applications	RDF / RDFS / OWL	Similar to Brick in purpose and scope, however, developed by the Linked Building Data W3C Community Group.
Semantic Sensor Networks (SSN and SOSA)	Sensors and Actuators	RDF / RDFS / OWL	To describe sensors and actuators, along with the features, observations, samples, and procedures generated by or acted upon by them. Not specific to the buildings domain but specific to the sensing domain.
RealEstateCore	Operational Applications and Real Estate	RDF / RDFS / OWL	To blend components primarily from IFC, Haystack, and SSN, as well as to introduce building ownership concepts into an ontology. Efforts are complementary to Brick.
Smart Applications REFerence for oneM2M (SAREF)	Smart Applications	RDF / RDFS / OWL	To describe smart applications. Not specific to the buildings domain.

Facility Smart Grid Information Model	Operational Applications and Grid Interaction	UML[2]	To define an abstract representation of the energy consuming, producing, and storage systems found in residential, commercial, and industrial facilities (“ANSI/ASHRAE/NEMA Standard 201-2016 (RA 2020),” 2020)

Project Haystack

Project Haystack is a popular building metadata schema based on the concept of “tags”. When first introduced into the market, its simple methodology and domain-oriented linguistic style was designed to provide controls contractors and field technicians (those who would be implementing the schema) a tool that was quickly understood and implemented. One of the primary critiques has been that its lack of formal structure prohibits a verification methodology. It is easy to claim that one uses Haystack, but this can have significantly different meanings since Haystack represents multiple things: a data model (implementation of Haystack tagging), an application programming interface (API) specification, and data serialization formats. The following points provide a short summary of the Haystack data model:

1. The primary domain of Haystack is the description of HVAC systems and metering infrastructure. Terms for lighting and other systems have also been incorporated.
2. Haystack uses sets of marker tags to describe ‘typing’ concepts. While definitions for individual tags are defined, the meaning of sets of tags is neither well-articulated nor machine interpretable, meaning that implementation is based more on folksonomy[3] than machine interpretable concepts for declaring the specific ‘type’ of something in a consistent manner.
3. Neither best-practice documentation nor thorough examples are available for how to implement Haystack across diverse and complex systems, resulting in lexical overlap, the ability to compose ambiguous tagsets, custom tag definitions, and significant user interpretation as to how to apply the schema.
4. The main verification that can currently be accomplished (Haystack 3.0) is to define mutually exclusive tags (a thing can’t be both a ‘point’ and an ‘equip’) and verify that entities don’t break those rules. However, these verification style rules are not embedded in the schema nor machine interpretable.
5. Consequently, thorough verification of a Haystack implementation is nearly impossible.
6. The development of Haystack 4.0 is intended to migrate the data model to a machine interpretable ontology.

Brick Schema

The other metadata schema gaining significant traction in the smart buildings space is the Brick Schema, which was designed in consideration of the atomic terms defined by Haystack 3.0. The initial design of Brick utilized an application-oriented approach to designing the schema, defining entities and their relationships necessary for eight applications, including

---

[2] Unified Modeling Language

[3] A user-generated system of classifying and organizing online content into different categories by the use of metadata such as electronic tags.

occupancy modeling, energy apportionment, web displays, model-predictive control, participatory feedback, fault detection and diagnostics, non-intrusive load monitoring, and demand response (Balaji et al. 2016). The following points provide a short summary of the Brick data model:

1. It has been designed from the onset using semantic web technologies.
2. An abstraction between individual tags and concepts is defined, such that a unique set of tags directly maps to a unique Brick concept.
3. More generic relationships have been defined (compared to Haystack 3.0) with formalized rules as to the domain and range[4] of those relationships.
4. To be maximally extensible, Brick defines a deep and specialized class hierarchy without having any fixed or required data properties. Brick v1.1 addresses this issue by explicitly parameterizing concept definitions for Points. For example, an “Air Temperature Sensor” is a Sensor that measures the “Temperature” property of the “Air” substance, and so on. Subsequent releases will expand this parameterization for other classes such as equipment and location.
5. Brick does not currently provide many textual definitions of concepts. This might not be necessary for all classes, but things such as `brick:On_Command`, `brick:On_Off_Command`, `brick:Off_Command`, and `brick:Start_Stop_Command` need to be clarified to ensure implementers can clearly delineate between said classes and their functional intent.
6. As mentioned previously, the suite of semantic web technologies may be perceived as complex and would benefit from assistive tools to improve accessibility for implementers.

To summarize; significant domain overlap in the Haystack and Brick schemas exists; they utilize different modeling approaches; they were designed to solve different problems[5]; they are at different stages of development[6]; they are both a step in the right direction compared to traditional point naming conventions; both are in need of a more formalized validation approach; and, both are in need of improved documentation, examples, tooling, and workforce education to enable widespread adoption.

Both of these metadata schemas are primarily implemented in the operational phase of a building and don’t account for building metadata provided by the other schemas mentioned in Table 1. Due to new building codes, in new construction projects, significant investment is made in digital documentation via BIM (IFC) and BEM (gbXML) during the design phase. Moreover, as a means to significantly reduce the operational energy of existing buildings, legislation, such as New York City’s Local Law 87 and San Francisco’s Energy Performance Ordinance, are requiring commercial buildings to undergo mandatory energy audits. Data from these buildings is primarily being stored in BuildingSync XML files, providing another valuable data stream to be used to bootstrap the creation of the operational metadata models.

---

[4] Semantic web relies on a graph representation of information with directed edges (relationships/predicates) between nodes, conveyed in a triple style format as (subject predicate object). The domain of an edge refers to the allowable type of the subject, while the range refers to the allowable type of the object.

[5] The design of the Haystack API is primarily a means to make finding entities of interest simpler for users. Compare this with the semantic web version of querying data (SPARQL), and the advantage in simplicity is a big win

[6] Haystack was established in 2014; Brick released their first paper in 2016

Proposed Solution

Based on the review of the existing work and the current state of building technologies, in order to enable a smart building future, a strategic and comprehensive approach to addressing semantic interoperability is required, including:

1. Harmonizing existing metadata schema efforts.
2. A governance structure which ensures a coordinated push forward in direct collaboration with industry partners, providing the means to create an informed, industry relevant, and widely adoptable semantic standard.
3. A solution and supporting resources which are accessible to technicians, service providers, and other critical workforce members without significant additional training.

The U.S. Department of Energy, Building Technologies Office, in partnership with the National Institute for Standards and Technology (NIST) and several national laboratories, will provide the central thrust for a semantic interoperability effort with the ultimate goal of significantly accelerating the timeline for reaching a GEB future. Our proposed solution is a three-pronged approach:

1. Develop a workflow and use case driven semantic interoperability standard that integrates and builds upon the successes of existing schema technologies;
2. Create a conformance verification framework for evaluating correctness of semantic implementations with respect to use cases, providing a mechanism for easier incorporation into owner's project requirements and engineering specifications; and
3. Provide necessary tools, documentation, and reference implementations to assist the workforce with a) metadata migration at key pain points in the building lifecycle, and b) implementation of the standard to defined use cases.

Outcomes of an Integrated Standard

The development of an integrated standard provides the following benefits:

1. Semantic metadata becomes consistent, shareable, and easily interpreted.
2. Supportive tooling can be created to make the job of the installer, commissioning agent, technician, controls engineer, or other individuals working on the building significantly easier.
3. The manufacturers producing building technologies can embed semantic information into their applications with predictable results and compatibility.
4. Use cases can be specified to clearly articulate minimum functional requirements or semantic sufficiency in order to deliver a certain level of functionality by a building system.

The following subsections will discuss this strategy in further detail.

Enabling Strategic Workflows

Through outreach and engagement with industry stakeholders, pain points in the generation of *operational* metadata schemas will be identified. Alignment of metadata concepts across relevant schemas will be undertaken and toolkits will be created to assist users in metadata migration from pre-operational schemas (IFC, gbXML, BuildingSync) to operational metadata schemas. An example narrative of a workflow is given below:

Existing Buildings: Energy Auditing to Measurement and Verification

Energy auditing will be an increasingly important mechanism to increase the energy efficiency of existing buildings, with ASHRAE Standard 211 providing a standard that cities can mandate through a legislative process. Auditors are required to capture building information, propose measures, and identify potential savings, all of which can be captured via BuildingSync XML. After energy efficiency measures are implemented, auditors, owners, and mandating bodies want to understand the effectiveness of the implementation – whether the implemented measures had the desired effects. This can be documented via the CalTRACK methods (Ngo 2019), but it requires data from the audit, as well as performance data from an EMIS. This is one prime workflow question that we intend to address. Other example workflows include:

- Control system programming to automated commissioning
- Automated commissioning to monitoring based commissioning, AFDD, and control sequence optimization
- BMS point discovery and mapping to monitoring based commissioning, AFDD, and control sequence optimization

Using Use Cases to Define Necessary Concepts for Semantic Sufficiency

Use cases are an effective way for capturing business processes and functional requirements. Means for formally capturing use cases have been defined by many industries. Textual, structural, and visual modeling techniques for specifying use cases started to emerge in the mid-1980's and became widely adopted in object orientated software development in the 1990's (Jacobson et al. 1992). The usage of use cases to capture functional requirements has since become prevalent in the development of health information technology (Bourquard et al, 2017), cybersecurity threat profiling – where they are used to decompose the system that might be attacked and identify attack vectors (Muckin et al. 2019), and standards and specification development (Hattachi et al. 2015). The telecommunication industry has long used use cases to drive the development of next generation cellular technology standards and specifications; a stand-alone organization – the Next Generation Mobile Networks Alliance – was created with the sole and specific purpose to "ensure that the standards for next generation network infrastructure, service platforms and devices will meet the requirements of operators and, ultimately, will satisfy end user demand and expectations" (i.e., develop use cases).

Use cases are not foreign to building systems. They have been used to some degree to prioritize work scope in the development of Brick and the IFC schema ("Model View Definition (MVD)" 2020), and to verify BuildingSync data requirements for different ASHRAE Standard 211 auditing levels. Using use cases to define the concepts necessary for semantic sufficiency will be invaluable for verifying interoperability from an end-user perspective, similar to how the BACnet Testing Laboratories certification process validates that building automation products are technically interoperable. Use cases drive plugfests, whereby both upstream and downstream application service providers gain confidence that they can use the information being provided to them by other applications.

Use cases are particularly useful for prioritizing scope of work for complex interdisciplinary undertakings with large numbers of possibilities, stakeholders, and challenges.

In our approach, use cases for building systems would be developed and prioritized based on semantic sufficiency. Examples of candidate use cases are presented below. Notably, each section below merely contains an introduction to the given use case. A fully documented use case would contain much more detail about roles, systems, and interactions.

1) **Energy Performance Reporting**

Energy use data is the most accurate way to evaluate how buildings perform, determine when corrective action needs to be taken due to subpar building performance, or implement data-driven energy management schemes – whereby new algorithms or other approaches can be incrementally implemented and validated or rejected based on energy reports. Energy data might also be used to validate energy performance measures incentivized by utility energy efficiency programs or contracted by energy service companies, issue renewable-energy certificates, and financially settle grid services.

The use of metadata is essential for exploring where and when energy is being consumed and helps software developers to easily deploy their tools without having to modify them for each customer. Further, when multiple stakeholder groups might want or need to access data that describes the energy performance of a building, metadata might facilitate reporting performance in different ways, thereby simultaneously meeting use case needs and maintaining privacy. For example, while a building owner might need to know the lighting or HVAC energy performance for a particular floor or conference room, they might prefer for outside parties to only see the data at a less granular level – so as to limit the ability to infer occupancy.

2) Automated Fault Detection and Diagnostics (AFDD)

It is estimated that 5 % to 30 % of the energy consumed by commercial buildings is wasted due to faults and errors in system control and operation (Granderson et al. 2017). In addition, a considerable amount of mechanical equipment must leave service before the end of its useful life because these faults and errors are not repaired in a timely manner. Although AFDD is a powerful approach to ensure efficient operation, it is still in its early stages of adoption. The current practice of applying AFDD via BMS involves manually going through a list of control points, which are usually different from building to building and from contractor to contractor, matching each point to the correct equipment and building space, identifying the set of points corresponding to each fault type, and then applying the corresponding fault rules or diagnostics.

The adoption of a semantic interoperability standard can automate the above processes and reduce the engineering time and money needed for deployment. Enforcement of standard naming approaches will enable control companies to deploy their pre-defined fault rules and methods to building installations without the need for any modification.

3) Automated Commissioning

Commissioning is a rigorous process following equipment installation that ensures human comfort, indoor air quality, energy efficiency, and reduced operations and maintenance costs. The current commissioning practice requires a human with specialized training – a commissioning agent – to manually verify that a building system is operating as intended, typically by reviewing the design intent and/or sequence, and stepping through scenarios that require the system to adjust and adapt. Commissioning agents also typically check whether the system data outputs are within their expected range. This process takes days and sometimes months, depending on the amount of equipment to commission and complexity of the system.

Common, well-defined semantic models and expected operational ranges of points should enable applications to be developed that can automatically run equipment through varying scenarios, thereby expediting the commissioning process to get buildings up and running in a shorter time frame and possibly with better performances. An example of such a tool is HVAC-Cx, developed by NIST (Milesi-Ferretti, Galler, and Bushby 2016).

4) Automatic Verification of ASHRAE Guideline 36 Control Sequence Implementation

ASHRAE Guideline 36 (Guideline-36 2018) was developed with the intent of reducing control programming and commissioning time through the adoption of best practices that are already proven to perform efficiently. If the model exposed by a BMS or controller could be verified as semantically sufficient for this use case, a third party AFDD application could have confidence plugging into this system, extracting the necessary data and points, and validating the controls implementation. Parallel effort of expressing control sequences in a vendor-neutral format through Modelica based control description language (CDL) can be leveraged in this regard.

5) Sensor data sharing

The developers of many building systems have explored and commercialized the use of sensors for the purpose of optimizing system performance and energy consumption. For example, HVAC and lighting system installations have successfully used occupancy data to improve their performance. The use of sensors is far from the norm, however, due to many factors – including inconsistent sensor-driven system performance, and sensor cost. In sophisticated buildings, multiple building systems (i.e., HVAC and lighting) might be deployed, each with its own sensors.

Semantic metadata standards would allow for the sharing of sensor data between building systems, which should reduce overall capital costs (i.e., first-costs and commissioning costs) and operational costs (e.g., maintenance, re-commissioning due to occupant complaints or needs). Further, they would facilitate the integration of sensors into building systems that are best suited for a particular measurement (e.g., lighting systems for occupancy sensors, HVAC systems for air quality sensors).

Tools and Reference Implementations

A standard is only effective when it is accessible and understandable to those implementing it. Providing technicians, engineers, system integrators, and other stakeholders with reference implementations of semantic models based on use cases and tools to assist with the specification and implementation is vital for widespread adoption. Learning from large datasets of building system data containing machine readable metadata (Example: Fierro et al. 2018) is helpful in this regard. Once workflows and use cases have been strategically identified and the semantic model developed to enable them, the development of easy-to-use tools will make the standard significantly more accessible. Engagement with industry will also remain vital to developing tooling that is designed for both new construction and existing building implementations.

Conformance Testing

Verification procedures ensure compliance of products with a standard. A standard testing and verification framework coupled with a process to recognize the compliance (for example a certification label) provides customers an easy means to procure compliant products. Giving the end users and building owners this capability is an effective strategy to create market pull for adoption. The example of BACnet's adoption is of particular relevance here. Despite its initial publication in 1995, it took BACnet more than two decades to reach its current widespread adoption (64% global market share according to BACnet International). A key element in accelerating the adoption was the formation of BACnet testing laboratories and the BTL label to verify compliance of products with the specifications in the BACnet standard.

Following this model, we propose that a verification framework for the semantic standard that can be constructed to evaluate the ability of a given system to deliver specific use cases like AFDD, calendaring-based space optimization, or other functionality based on the building's semantic graph, will play an instrumental role in minimizing the adoption time for this standard.

In summary, we believe that a semantic interoperability standard based on existing semantic specifications and use cases, supported by assistance tools and reference implementations and a verification framework to evaluate compliance - can provide the thrust to enable semantic interoperability in buildings.

Impacts of a Semantic Standard

The development of a semantic interoperability standard enables two critical pathways forward: easy procurement of interoperable systems and testing and validation of equipment for interoperability. With a workflow and use case driven approach to standard development, and the model views that establish the "semantic sufficiency" required to deliver that corresponding functionality, support tooling can be made to test the semantic model contained within a building's operating system (EMIS, BAS, or other supervisory system layer) and determine whether the system is actually capable of delivering the identified service(e.g. AFDD).

Similarly, building owners and managers will be able to include that model view as a requirement in procurement documents. For instance, if a manager wants the system installed in a new building to be able to support calendaring coordination, or system reconfiguration to account for future build-outs, they can reference the semantic standard to ensure that the minimum functional requirements, in terms of information conveyed through the building's system, are met and can operate appropriately.

Finally, the owner or manager can specify a use case and simply identify the responsible parties. Today, those parties must reach consensus on their own regarding how they will coordinate semantic information in their networks and across their software platforms. In the future, they will have a neutral, industry-driven standard to reference, which they can leverage to save time and cost in installation and provide a higher level of quality assurance when delivering a final product to the customer.

The standard being proposed would result in a transformative moment for the connected device industry. While there are parallel efforts such as Project Connect Home over IP, they are specifically focused on the residential space and are not going to bring much to the table for commercial products, which are significantly more complex. There is also a need to break into the multifamily and small/medium commercial markets, where the systems are not as complex as those found in large commercial buildings, but for which significant opportunities still exist for energy efficiency improvements and performance enhancements. The standard proposed here can serve as the framework enabling residential service providers to move into the larger and marginally more complex multifamily and light commercial markets with lower technical barriers to engagement.

Next Steps

A Semantic Interoperability Working Group has been formed within the ASHRAE BACnet committee to develop the needed standard. This standard, now under development, has been given the designation ASHRAE Standard 223P, where the "P" indicates that the standard is proposed. Standard development can take some time, but the resulting industry consensus is needed to lay the groundwork for industry to build smart building applications in a meaningful and cost-effective manner. Following the success story of BACnet penetration into the market, work is needed in several key areas:

1. Creating a consistent ontology which can be applied via the semantic standard but is extensible across other building energy data use cases and is compatible with more advanced design systems like BIM;
2. Field validation and demonstration projects with quantifiable financial impacts; and
3. Workforce education and development to support the impact that semantic information can have on installation of new technologies, commissioning and configuration of new systems, and incorporation of existing and legacy technologies into the smart building concept.

It is critical for technologies to not simply be operational themselves, but to be easily adopted, installed, and leveraged by the buildings' workforce.

Conclusion

For data driven smart buildings to be both cost-effectively delivered and functionally useful, the standardization of semantic information must take place. This paper outlined a compelling path forward, to be led by the U.S. Department of Energy, NIST, and several national laboratories and leading colleges and universities. Though demonstration projects illustrating the value of semantic information through efforts like Brick are limited, the value of semantic standardization is clear and undeniable. That value can be increased significantly when standardization further reduces costs and the time commitments necessary to deliver a smart, connected building.

A semantic standard, leading to improvements such as automated commissioning, energy performance reporting, and more, reduces the time required to set up a new system once installed and ensure it is running properly. This means less technician time, less disruption to the use pattern of the facility, and improved ability to identify issues where maintenance is needed. Semantic standardization provides a lifetime benefit to new and connected technologies and will be a critical piece of the puzzle as buildings move further into the internet age. As semantic standardization gets adopted at EMIS, BAS, and embedded devices, we get closer to a smart, grid-interactive, efficient built environment.

Acknowledgements

This work was supported by US Department of Energy Building Technology Office under Semantic Interoperability R&D project (CPS agreement number 34579).

References

"ANSI/ASHRAE/NEMA Standard 201-2016 (RA 2020)." n.d., 836.

Balaji, Bharathan, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, et al. 2016. "Brick: Towards a Unified Metadata Schema For Buildings." In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*, 41–50. BuildSys '16. Palo Alto, CA, USA: Association for Computing Machinery. https://doi.org/10.1145/2993422.2993577.

Balaji, Bharathan, Chetan Verma, Balakrishnan Narayanaswamy, and Yuvraj Agarwal. 2015. "Zodiac: Organizing Large Deployment of Sensors to Create Reusable Applications for Buildings." In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys '15*, 13–22. Seoul, South Korea: ACM Press. https://doi.org/10.1145/2821650.2821674.

Bhattacharya, Arka A., Dezhi Hong, David Culler, Jorge Ortiz, Kamin Whitehouse, and Eugene Wu. 2015. "Automated Metadata Construction to Support Portable Building Applications." In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys '15*, 3–12. Seoul, South Korea: ACM Press. https://doi.org/10.1145/2821650.2821667.

Bray, Tim, Jean Paoli, C.M. Sperberg-McQueen, Eve Maler, and Yergeau Francois. 2008. "Extensible Markup Language (XML) 1.0 (Fifth Edition)." November 26, 2008. https://www.w3.org/TR/xml/#sec-origin-goals.

"Brick Schema." 2020. Brick Schema. March 20, 2020. https://brickschema.org/.

"Building Energy Data Exchange Specification (BEDES) | BEDES." n.d. Accessed March 19, 2020. https://bedes.lbl.gov/.

Fierro, Gabe, Jason Koh, Yuvraj Agarwal, Rajesh K. Gupta, and David E. Culler. 2019. "Beyond a House of Sticks: Formalizing Metadata Tags with Brick." In *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 125–34. New York NY USA: ACM. https://doi.org/10.1145/3360322.3360862.

Fierro, Gabe, Marco Pritoni, Moustafa AbdelBaky, Paul Raftery, Therese Peffer, Greg Thomson, and David E Culler. 2018. "Mortar: An Open Testbed for Portable Building Analytics," 10. https://doi.org/10.1145/3366375.

Gao, Jingkun, Joern Ploennigs, and Mario Berges. 2015. "A Data-Driven Meta-Data Inference Framework for Building Automation Systems." In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys '15*, 23–32. Seoul, South Korea: ACM Press. https://doi.org/10.1145/2821650.2821670.

Granderson, Jessica, and Guanjing Lin. 2016. "Building Energy Information Systems: Synthesis of Costs, Savings, and Best-Practice Uses." *Energy Efficiency* 9 (6): 1369–84. https://doi.org/10.1007/s12053-016-9428-9.

Hardin, Dave, Eric G. Stephan, Weimin Wang, Charles D. Corbin, and Steven E. Widergren. 2015. “Buildings Interoperability Landscape.” PNNL--25124, 1234792. https://doi.org/10.2172/1234792.

Hong, Dezhi, Hongning Wang, Jorge Ortiz, and Kamin Whitehouse. 2015. “The Building Adapter: Towards Quickly Applying Building Analytics at Scale.” In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys ’15*, 123–32. Seoul, South Korea: ACM Press. https://doi.org/10.1145/2821650.2821657.

Koh, Jason, Bharathan Balaji, Dhiman Sengupta, Julian McAuley, Rajesh Gupta, and Yuvraj Agarwal. 2018. “Scrabble: Transferrable Semi-Automated Semantic Metadata Normalization Using Intermediate Representation.” In *Proceedings of the 5th Conference on Systems for Built Environments - BuildSys ’18*, 11–20. Shenzen, China: ACM Press. https://doi.org/10.1145/3276774.3276795.

Milesi-Ferretti, Natascha S, Michael A Galler, and Steven T Bushby. 1924. “Evaluating the Initial Field Performance of HVAC-Cx for Air Handling Units.” *NIST Technical Note*, 19. https://doi.org/10.6028/NIST.TN.1924.

Mishra, Sakshi, Andrew Glaws, Dylan Cutler, Stephen Frank, Muhammad Azam, Farzam Mohammadi, and Jean-Simon Venne. 2020. “Data-Driven Metadata Tagging for Building Automation Systems: A Unified Architecture.” *ArXiv:2003.07690 [Cs]*, February. http://arxiv.org/abs/2003.07690.

“Model View Definition (MVD).” n.d. BuildingSMART Technical. Accessed March 19, 2020. https://technical.buildingsmart.org/standards/mvd/.

Ngo, Phil. 2019. “CalTRACK Documentation.”

“Project Haystack.” 2020. Project Haystack. March 20, 2020. https://project-haystack.org/.