Activity A3 Information Modelling for Data-Driven Smart Buildings

IEA ANNEX 81 PLENARY, 13 OCTOBER 2022, GOTHENBURG

Pieter Pauwels, Associate Professor, Eindhoven University of Technology
Gabe Fierro, Assistant Professor, Colorado School of Mines

Department of the Built Environment, Information Systems in the Built Environment
Presentation Outline

1. Activity A3 Introduction
2. A3 activity key concepts
3. Value Proposition: position statement
   - General questions and comments
   - What is missing?
   - Which tasks need to be done by whom to finalize the work?
4. Tool chains for data creation, curation and validation: how to obtain a high-quality metadata-driven smart building
   - Overall approach
   - Quality indicators
   - SHACL pipeline
   - Quality validation procedure
5. Next steps - closure
Activity A3 – Modus operandi

- **Members:** see member list

- **Output:**
  - Deliverables published by shortlist of contributors, under the lead of 1-2 main authors (= editors)
  - Articles published by activity members
  - Open-source tooling and software in ownership of individual members

- **Meetings:**
  - Monthly meetings: US/Australia timezone + Europe timezone
  - More need for pro-active input
Activity A3 – Our subject

Activity A3 aims to provide the knowledge, standards, protocols, and procedures for low-cost, high-quality data capture, sharing, and utilisation in buildings, particularly focusing on ‘Data Information Management’.

⇒ Data – Data – Data
⇒ And metadata 😊 namely: Smart building metadata
### Extract from BMS (labels):

Mapping of sensors to rooms in BMS building model:

<table>
<thead>
<tr>
<th>Column</th>
<th>ItemName</th>
<th>ItemDescriptionEnglish</th>
<th>Spaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>6358</td>
<td>11NR008FT-003FLW</td>
<td>ROOM TEMPERATURE 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>6359</td>
<td>11NR008FT-003FLW</td>
<td>ROOM TEMPERATURE 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>2089</td>
<td>ACTUEEL SETPOINT REGELING 9_Z01</td>
<td>CURRENT SETPOINT SCHEME 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>6821</td>
<td>STAND ZONERING 9_Z01</td>
<td>STAND AWNINGS 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>5900</td>
<td>REGELSIGNAL VERWARMING 9_Z01</td>
<td>CONTROL SIGNAL HEATING 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>5899</td>
<td>REGELSIGNAL VERWARMING 9_Z01</td>
<td>CONTROL SIGNAL HEATING 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>4349</td>
<td>KOELINGSETPOINT 9_Z01</td>
<td>COOLING SETPOINT 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>4350</td>
<td>KOELINGSETPOINT 9_Z01</td>
<td>COOLING SETPOINT 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>3568</td>
<td>CPA VIA WANDMODULE RUIMTE 9_Z01</td>
<td>CPA VIA WALL MODULE ROOM 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>3567</td>
<td>CPA VIA WANDMODULE RUIMTE 9_Z01</td>
<td>CPA VIA WALL MODULE ROOM 9_Z01</td>
<td>9_Z01</td>
</tr>
<tr>
<td>6823</td>
<td>STAND ZONERING 9_Z01</td>
<td>STAND AWNINGS 9_Z01</td>
<td>9_Z01</td>
</tr>
</tbody>
</table>

---

How to make this smarter?

---

Activity A3 – Audience and Objective

Target audience – who do we do this for?
1. the person in charge of operational management at a building level
2. application developers at a portfolio or sector-wide level

Objective: Making portable applications that use metadata to mask the inherent complexities of each building will, in turn, allow the proliferation of value-adding applications and services
Diversity in Metadata Schemas for Data-driven Smart Buildings

- Internet of Things
- digitalization
- cyber-physical systems
- data-driven processes
- classification tags
- metadata
- telemetry
- data monitoring

- fault detection and diagnosis (FDD)
- control strategies
- performance measurement
- energy auditing

Data-driven Technology

Business opportunities and market opportunities

Outcomes and deliverables

1. Report delivered mid 2022:

2. Position statement end 2023: The Value Proposition – short paper

3. Third deliverable: Tool chains for data creation, curation and validation

Presentation Outline

1. Activity A3 Introduction
2. A3 activity key concepts
3. Value Proposition: position statement
   - General questions and comments
   - What is missing?
   - Which tasks need to be done by whom to finalize the work?
4. Tool chains for data creation, curation and validation: how to obtain a high-quality metadata-driven smart building
   - Overall approach
   - Quality indicators
   - SHACL pipeline
   - Quality validation procedure
5. Next steps - closure
Terms and Definitions in our Activity

BIM
BMS
EMS
Information Model
Instance
Metadata
Metadata Schema
Model
Telemetry

Definitions

BIM: Building Information Modelling provides a 3D digital representation of the building structure and the plant and equipment contained within. 4D, 5D, 6D and xD BIM extend basic design and construction to scheduling, cost estimation, facility management and performance evaluations, respectively. As of 2021, 6D and xD BIM use is limited, with most active BIM users typically found in the architectural, engineering and construction (AEC) professions.

BMS: A Building Management System is a combination of software, hardware and communications infrastructure, designed to support the building operation including the HVAC systems and subsystems such as fans, pumps and chillers. This is also referred to as a Building Automation System (BAS).

EMS: An Energy Management System (also called an EMS — Energy Management Information System) is a BMS system focused on monitoring, control and orchestration of large energy consuming devices within the building. When the focus is on whole-building energy management, the term Building Energy Management System (BEMS) is also used.

Information Model: a digital model that represents a collection of information, for example about a building (e.g., Building Information Model). A model follows a well defined schema.

Instance: The individual parts that together form a model, and that follow a particular schema (e.g., metadata schema).

“A data-driven building is one whose processes are automated and driven by the use of historical and/or live building telemetry and which receives digital commands through such a data-driven process.”

Inclusion of bidirectional communication

These processes not only include **data analytics and monitoring**, but also **decision processes** that affect the functionality of the building in a good way.

This means that **a form of feedback** is expected in order to achieve a data-driven smart building, as the use of data needs to inform automated (control) or human (users) decision-making.

Key criteria for a data-driven smart building

1. Support for storage of data rather than semantics.

2. Focus on building-centric data with a minor reference to user-centric occupancy data.

3. Focus on Energy Management data.

4. Focus on the operational phase.

5. Focus on the management of data.

35 sensor nodes monitoring Temperature (°C), Relative Humidity (%), Air Pressure (hPa), Indoor Air Quality (Total Volatile Organic Compounds (TVOC), ppb) and CO2 (ppm)), illuminance (lux) and motion

- Data storage in SQL database
- Data monitoring and visualization in Grafana

Categories of data?

• **Data streams:**
  - Ordered lists of values, typically floating point values
  - Large amount of data -> data lakes
  - Almost no semantics, at best a few labels for categorization
  - E.g. temperature measurements, system logs (e.g. triggering of actuators), etc.

• **Semantically rich and interconnected data:**
  - Seldom including large data streams
  - Long debates about the semantics of things -> standardisation
  - Complex and brittle (breaks easily)
  - Small amount of very important data
  - Easy to combine with rule-based and/or logic-based technologies (inference and query)
  - E.g. BIM models, semantic web ontologies, taxonomies, OTLs, etc.

• **Control models:**
  - Algorithms for control, parametric functions
  - Communication system, signal processing, direct control, low latency
  - Typically located on the edge (devices with embedded functions)
  - E.g. Control Description Language (CDL), modelica models

• **User data:**
  - Outside of the system
  - Different privacy and security requirements

• **Files:**
  - No semantically rich encodings, no data streams
  - Highly valuable, and seldom machine-processable
  - E.g. PDFs, Images, Geometry ‘blobs’
The current problem of data-driven smart buildings

- Buildings produce endless streams of sensor, meter, and IoT data
- Business-as-usual approach treats each building as a stand-alone entity, replete with bespoke engineering solutions and a mix of standardised and bespoke metadata schemas to describe objects and states within the building
- Data-driven smart buildings are different from building to building, from integrator to integrator

=> need for a common metadata schema
=> which one??

Targeted framework for AI-based smart buildings

**Important:**
- Include access control (ACL)
- ‘linking’ of data on system integration level
- Agreement and standardization of labels and metadata tags
- Feed back into control systems!

Core metadata schemas reviewed

1. Project Haystack
2. Brick Schema
3. Real Estate Core (REC)
4. BOT ontology and Linked Building Data (LBD)
5. SAREF (SAREF4BLDG)
6. SOSA / SSN
7. Google Digital Buildings

Haystack Tags

“We standardize semantic data models and web services with the goal of making it easier to unlock value from the vast quantity of data being generated by the smart devices that permeate our homes, buildings, factories, and cities.” (https://project-haystack.org/)

Haystack = stack of technologies:

- **Data Types**: a fixed set of data types for modeling information
- **File Types**: a set of text formats to encode and exchange those data types
- **HTTP API**: a protocol to exchange data over HTTP using those file types
- **Ontology**: a standard way to model common concepts such as buildings, equipment, and sensors
- **Defs**: a standard way to define and extend the ontology

Individual aspects of the technology stack can be used on their own. For example you can use the Haystack data types as an “enhanced JSON”. Or you could use just the terms in the ontology without the Haystack data types.

Ontology Core:

Brick: A uniform metadata schema for buildings

Brick is an **open-source effort** to standardize semantic descriptions of the **physical, logical and virtual assets** in buildings and the **relationships between them**. Brick consists of an **extensible dictionary** of terms and concepts in and around buildings, a set of **relationships** for linking and composing concepts together, and a **flexible data model** permitting seamless integration of Brick with existing tools and databases.

[Diagram of Brick schema]

https://brickschema.org/
Linked Building Data, incl. Brick

Ontology Core:

Possibility to extend with other vocabularies:

Slide from
Cloud of linked building data

Brains for buildings: where to find all the relevant smart building data? - Pieter Pauwels
A qualitative comparison of metadata schemas

1. Structure of the model
2. Vocabulary organization and completeness and strictness/rigor
3. How is alignment handled
4. What is the role of the model in a hypothetical or reference data-driven building process?
5. Required tooling / software support / expertise
6. Creation / bootstrapping / maintenance

The Scale of AI Methods

STATISTICAL AI
- Plain labels
- Haystack
- REC

SYMBOLIC AI
- SAREF, SOSA, SSN
- LBD
- BRICK
- IFC

TAGGING
- Most Smart Buildings

METADATA

LINKING DATA

SEMANTICS

Presentation Outline

1. Activity A3 Introduction
2. A3 activity key concepts
3. Value Proposition: position statement
   - General questions and comments
   - What is missing?
   - Which tasks need to be done by whom to finalize the work?
4. Tool chains for data creation, curation and validation: how to obtain a high-quality metadata-driven smart building
   - Overall approach
   - Quality indicators
   - SHACL pipeline
   - Quality validation procedure
5. Next steps - closure
What is a Value Proposition statement

- A lot of value proposition has been described already in the survey of metadata schemas

- Difficult to comprehend for any non-technical person, however
- A Value Proposition needs to show Unique Selling Points (USPs) that are easy to grasp
- Need for a short industry-oriented white paper
The intended Value Proposition statement

- Short industry-oriented white paper

Business propositions
Value propositions

Section 0
Introduction

Section 1
Overview of metadata schemas
- scope and criteria
- terminology and definitions
list of core schemas related works

Section 2
Qualitative Comparison
1. model structure
2. completeness and formal rigor
3. alignments
4. position in reference software architecture
5. required tooling
6. creation and maintenance of models

Section 3
Additional schemas and models
1. asset management models
2. simulation models
3. protocols and communication-oriented data schemas
4. emerging and stand-alone initiatives

Section 4
Summary
Diversity in Metadata Schemas for Data-driven Smart Buildings

- Internet of Things
- digitalization
- cyber-physical systems
- data-driven processes
- classification tags
- metadata
- telemetry
- data monitoring

- fault detection and diagnosis (FDD)
- control strategies
- performance measurement
- energy auditing

Data-driven Technology

Business opportunities and market opportunities
Strengths and methods in Artificial Intelligence

**MACHINE LEARNING**

- Reinforcement learning: data comes from an available experimentation environment
- Neural networks: ANNs, GNNs, CNNs, ...
- Traditional ML:
  - Regression: predict next value
  - Classification: predict category / classification label
  - Clustering: group based on similarity
  - Association: identify sequences and combinations

**STATISTICAL AI METHODS**

- Semantic Web technologies:
  - Ontologies and formal vocabularies
  - Logics: Description Logic, Defeasible Logic, etc.
- Expert Systems
- Rule-based inference: if-then rules

**SEMANTICS**

Semantically rich and interconnected data

**SYMBOLIC AI METHODS**

Data streams

- Reinforcement learning: data comes from an available experimentation environment
- Neural networks: ANNs, GNNs, CNNs, ...
- Traditional ML:
  - Regression: predict next value
  - Classification: predict category / classification label
  - Clustering: group based on similarity
  - Association: identify sequences and combinations

- Semantic Web technologies:
  - Ontologies and formal vocabularies
  - Logics: Description Logic, Defeasible Logic, etc.
- Expert Systems
- Rule-based inference: if-then rules

Slide from
The Scale of AI Methods

Can we integrate these methods and use all of them?

The Scale of AI Methods

STATISTICAL AI

Plain labels

Haystack

REC

LBD

BRICK

IFC

SYMBOLIC AI

TAGGING

METADATA

LINKING DATA

SEMANTICS

Most Smart Buildings

Slide from
https://pure.tue.nl/ws/portalfiles/portal/202217259/BitsBricksBehaviour_P
Pauwels.pdf
Presentation Outline

1. Activity A3 Introduction
2. A3 activity key concepts
3. Value Proposition: position statement
   - General questions and comments
   - What is missing?
   - Which tasks need to be done by whom to finalize the work?
4. Tool chains for data creation, curation and validation: how to obtain a high-quality metadata-driven smart building
   - Overall approach
   - Quality indicators
   - SHACL pipeline
   - Quality validation procedure
5. Next steps - closure
Semantic graph of sensors

@prefix brick: <https://brickschema.org/schema/Brick#> .
@prefix inst: <http://linkedbuildingdata.net/ifc/resources20201208_005325/> .
@prefix ph: <https://project-haystack.org/def/ph/3.9.11#> .
@prefix phIoT: <https://project-haystack.org/def/phIoT/3.9.11#> .
@prefix phScience: <https://project-haystack.org/def/phScience/3.9.11#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

inst:11NR008LT-001PIRTM a brick:Occupancy_Sensor,
  phIoT:sensor ;
  rdfs:label "PRESENCE_8_128"^^xsd:string ;
  brick:hasLocation inst:space_892 ;
  ph:dis "PRESENCE_8_128"^^xsd:string ;
  ph:hasTag phIoT:his,
    phIoT:occupancy ;
  phIoT:spaceRef inst:space_892 .

inst:11NR008LT-003PIRTM a brick:Occupancy_Sensor,
  phIoT:sensor ;
  rdfs:label "PRESENCE_8_127"^^xsd:string ;
  brick:hasLocation inst:space_1023 ;
  ph:dis "PRESENCE_8_127"^^xsd:string ;
  ph:hasTag phIoT:his,
    phIoT:occupancy ;
  phIoT:spaceRef inst:space_1023 .

Targeted framework for AI-based smart buildings

Important:
- Include access control (ACL)
- ‘linking’ of data on system integration level
- Agreement and standardization of labels and metadata tags
- Feed back into control systems!

Can we avoid bespoke methods?

System Integrator
- API- and Web-based
- not semantics-based only
- not ML-based only
- human-centric

Semantic
data store

Data lakes

Control system incl. control logic

SYMBOLIC AI

STATISTICAL AI
How to create such information model??

1. Converters?
2. Creation from scratch?
3. Machine learning?

A better pipeline is needed --
And is it even valid???

- Quality indicators
- SHACL pipeline
- Quality validation procedure

van den Bersselaar, E. Automatic validation of technical requirements for a BIM model using semantic web technologies: Description of a linked building data workflow using SHACL as checking mechanism to improve information exchanges between actors in the AEC industry. MSc Thesis, Eindhoven University of Technology, 2022.
General architecture for model checking and validation

- Reporting
- Validation
- Quality checking
- Guideline
- Improvement of data

van den Bersselaar, E. Automatic validation of technical requirements for a BIM model using semantic web technologies: Description of a linked building data workflow using SHACL as checking mechanism to improve information exchanges between actors in the AEC industry. MSc Thesis, Eindhoven University of Technology, 2022.
Next steps - discussion
Presentation Outline

1. Activity A3 Introduction
2. A3 activity key concepts
3. Value Proposition: position statement
   - General questions and comments
   - What is missing?
   - Which tasks need to be done by whom to finalize the work?
4. Tool chains for data creation, curation and validation: how to obtain a high-quality metadata-driven smart building
   - Overall approach
   - Quality indicators
   - SHACL pipeline
   - Quality validation procedure
5. Next steps - closure
Outcomes and deliverables

1. Report delivered mid 2022:

2. Position statement end 2023: The Value Proposition – short paper

3. Third deliverable: Tool chains for data creation, curation and validation
Thank you

Pieter Pauwels
Associate Professor

Information Systems in the Built Environment (ISBE)
Dept. of the Built Environment
Eindhoven University of Technology

p.pauwels@tue.nl