

# OntoH2G: A semantic model to represent building infrastructure and occupant interactions

Richard Chbeir<sup>1</sup>, Aitor Corchero<sup>2</sup>, Yudith Cardinale<sup>3,1</sup>, Pierre Bourreau<sup>4</sup>,  
Khoulood Salameh<sup>1</sup>, Nathalie Charbel<sup>1,4</sup>, Lara Kallab<sup>1,4</sup>, Chinnapong  
Angsuchotmetee<sup>1</sup>, and Gulben Calis<sup>5</sup>

<sup>1</sup> Univ Pau & Pays Adour, LIUPPA, EA3000, 64600, Anglet, France  
rchbeir@acm.org, {ksalameh, ncharbel, lkallab, cangsuc1}@univ-pau.fr

<sup>2</sup> Eurecat Technology Centre, Barcelona, Spain  
aitor.corchero@eurecat.org

<sup>3</sup> Dpto. de Computación y Tecnología de la Información, U. Simón Bolívar, Venezuela  
ycardinale@usb.ve

<sup>4</sup> NOBATEK, Esplanade des Arts et Métiers / 33405 TALENCE Cedex  
pbourreau@nobatek.com

<sup>5</sup> Ege University, Faculty of Engineering, Depart. of Civil Engineering, Izmir, Turkey  
gulben.calis@ege.edu.tr

**Abstract.** In order to reduce the energy gap originated by the difference between existing tools estimations and real energy consumption, HIT2GAP European H2020 project aims at advancing on building control tools by providing a newer decision-making technology. Technically, HIT2GAP offers a platform inspired by previous reference architectures (e.g., Haystack) and complements them through a knowledge-based model, called OntoH2G, to store building information under a common vocabulary and consequently to enable fine-grained vision of the building with its equipment and occupants. OntoH2G allows to query building data so to extract required information and interesting events (through a set of services). OntoH2G advances over existing models on two main aspects: (i) being compliant with well-known ontologies in different domains in order to cover all energy building concepts, and (ii) its ability to represent user/occupant behaviour, preferences, and interactions.

**Keywords:** ifcOWL, SSN, Building Infrastructure, Occupant Behaviour

## 1 Introduction

Nowadays, a smart building can be perceived as a cyber-physical system in which physical (machines and humans) and digital entities exist and interact. In order to provide smart building management, intelligent systems need to be injected to cope with energy optimization and occupant comfort. In this scenario, the HIT2GAP H2020 European project (Highly Innovative building control Tools Tackling the energy performance GAP) addresses the energy performance gap from the commissioning phase to the operational phase, in the context of building usage.

Considering the high volume, variety, and velocity of the data generated within a building and needed to detect certain critical and important events, Building Management Systems (BMS) demand careful design of the data storage model in order to be sustainable and efficient. As a response, HIT2GAP offers an open toolbox inspired by key referent architectures (e.g., Haystack<sup>1</sup>), which relies on a multi-layered middleware, called Hit2Gap Core Platform, in charge of storing the data associated to building management and occupants' interactions, and enabling building systems and modules coordination and integration [4]. The architecture of the HIT2GAP framework is depicted in Fig. 1.

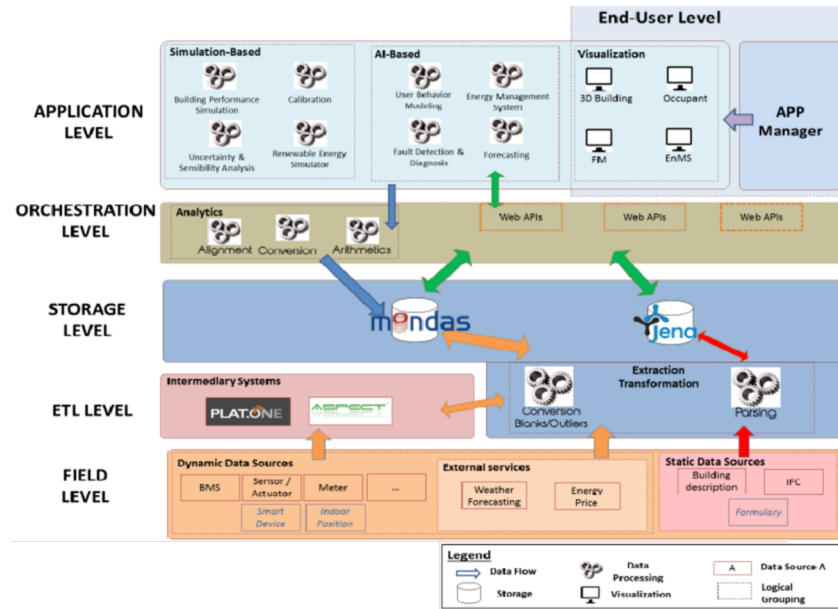


Fig. 1: Hit2Gap Platform

It is mainly structured into five levels: (i) **Field level**: contains all the data sources that are relevant for the analysis of the energy performance gap; (ii) **ETL (Extract-Transform-Load) level**: is in charge of pre-processing collected data and transfer them to the storage repositories; at this level, two different storage modes are available: a) *indirect*, in which data is first collected from the sources (e.g., BMS, sensors, occupants) by an intermediate system (e.g., PLAT.One or Cylon aspect), and transferred to the Core Platform through web APIs; and b) *Direct*, in which data is transferred to the Core Platform through the very same web APIs; the Core Platform only manages web protocols to allow direct access, connection, interaction to sensors and BMS through these standard protocols; the transformation modules identified in the project are conversion (so as to normalize all data collected from different devices), detection and interpolation of blanks and outliers (also known as data cleansing); (iii) **Storage level**: is in

<sup>1</sup> <http://project-haystack.org>

charge of storing all necessary data, considering the heterogeneity of collected data. This can be considered as an independent brick of the HIT2GAP solution for different (technical and commercial) reasons and may evolve in the future. Our current solution is dual in such a way that descriptive information (i.e., static data) are stored in an RDF store (Jena), while measures (i.e., dynamic data) are stored under Mondas storage (an HDF5-based system); (iv) **Orchestration level**: takes care of redirecting requests coming from the application level to the correct services; for instance, a client request can involve accessing data in the storage level, but also getting data from other services, whether internal to the HIT2GAP platform (i.e., generated on demand by HIT2GAP modules) or external to the platform. The orchestration level also contains built-in services to provide client some means to pre-process the data. They can be combined in a workflow specified by end-users (the module developer) through an API and potentially a Graphic User Interface (GUI); (v) **Application level**: contains all data processing services, which are in charge of producing some added-value. The modules of the application level communicate with the Core Platform through web APIs; a module at the application level can either be internal or external to the HIT2GAP platform; for instance, the OpenWeatherMap<sup>6</sup> service can be used to obtain weather forecasting information.

Due to the heterogeneous characteristics of the data required by the different modules (e.g., information on the building infrastructure, measures from sensors and actuators, occupants' interactions and preferences), the Hit2Gap Core Platform relies on a knowledge-based model, called OntoH2G, driven by an extension of standard and commonly-adopted ontologies to perform an appropriate management of the systems and modules (digital building world). To achieve this, OntoH2G establishes a common vocabulary for interrelating systems, data, occupants, and processes, capturing consequently the building behaviour. Thanks to its syntactic and semantic interoperability, OntoH2G is able to: (i) store, integrate, and aggregate heterogeneous data; (ii) provide reusability and extensibility; (iii) provide reasoning means to discover new information; (iv) cope with advanced data and information retrieval; (v) ease the implementation of intelligent algorithms and modules; and (vi) allow future evolutions and adaptations. In what follows, we give an overview of OntoH2G.

## 2 General view of OntoH2G

In order to provide appropriate data management functionalities in the Hit2Gap Core Platform, we have developed OntoH2G<sup>7</sup>, a dedicated ontology to support sharing data in a smarter and normalized way, homogenizing the existing building protocols under a common vocabulary.

### 2.1 OntoH2G Conception

For the OntoH2G conception, we were based on well-known methodologies for building ontologies [2,7]. In general, these methodologies divide the ontology

<sup>6</sup> <http://openweathermap.org/>

<sup>7</sup> <https://github.com/HIT2GAP-EU-PROJECT/HIT2GAPOnt/tree/master/owlModels>

development into a set of phases: (i) specification; (ii) knowledge acquisition; (iii) conceptualization; (iv) integration; (v) implementation; and (vi) evaluation. The *specification* and *knowledge acquisition* phases (e.g., semi-formal definition of concepts and taxonomy of concepts, build relationship diagrams, gather knowledge), were done based on domain experts requirements, on the Core Platform specification, and on the revision of existing standard and proposals related to the building domain (IfcOwl<sup>8</sup>, SSN<sup>9</sup>, Haystack<sup>10</sup>, obXML<sup>11</sup> [3], and many others). In the *conceptualization* and *integration* phases, we created a glossary of terms and a dictionary of terms, we refined the description of all previously defined concepts and relations (by adding more properties and attributes), we enriched the data model by integrating other existing ontologies, and we validated concision, completeness, and consistency with the participation of domain experts and users. To *implement* OntoH2G, we used Protege<sup>12</sup> to generate the OWL representation, as well as the first version of the documentation<sup>13</sup>. Currently, we are conducting the *evaluation* phase, in which domain experts and end-users are involved. We are also working in the integration of OntoH2G into the HIT2GAP Core Platform for testing it with real data and further improvements.

## 2.2 OntoH2G Features

OntoH2G describes building infrastructure, including physical and digital entities, as well as user-building interactions not only in the form of activities but also comfort requirements, user preferences, and other aspects that motivate users to interact with the building. With this aim, OntoH2G takes into consideration the following aspects: (i) **Building infrastructure data** to represent different systems and infrastructures involved in the Building Information Modelling (BIM) that are required for the performance of the project tools (building simulation tools, building monitoring tools, visualization tools, etc.); (ii) **Sensor data and sensor network infrastructure** representations that are supported in BIM; and (iii) **User behaviour data** aimed at representing building occupants' DNA (Drivers, Needs, and Actions). Thus, OntoH2G extends various well-known and standard ontologies:

- Industry Foundation Classes (ifcOWL) [5]: to represent building information and life cycle <sup>14</sup>.
- Semantic Sensor Network Ontology (SSN) [1]: to represent sensors, actuation information, and data following Observation and Measurement patterns <sup>15</sup>.

<sup>8</sup> <http://openbimstandards.org/standards/ifcowl/>

<sup>9</sup> <https://www.w3.org/TR/vocab-ssn/>

<sup>10</sup> <https://project-haystack.org/>

<sup>11</sup> <https://behavior.lbl.gov/?q=obXML>

<sup>12</sup> <https://protege.stanford.edu>

<sup>13</sup> <https://hit2gap-eu-project.github.io/HIT2GAPOnt/>

<sup>14</sup> <http://www.w3.org/ns/ssn/>

<sup>15</sup> [http://www.buildingsmart-tech.org/ifcOWL/IFC2X3\\_Final#](http://www.buildingsmart-tech.org/ifcOWL/IFC2X3_Final#)

- Quantities, Units, Dimensions, Time (QUDT)<sup>2</sup>: to represent units and unit conversion for time series and sensor outputs (represented mainly with SSN).
- W3C Time Ontology (TIME)<sup>3</sup>: to contextualize time measurements and time instants. It complements QUDT regarding time measurements.
- W3C Geospatial Ontology (W3CGeo)<sup>4</sup>: to represent geospatial information for the building infrastructure and occupants’ positioning in the building.

Figure 2 shows the integration and extension of all these ontologies. Next section describes the main extensions.

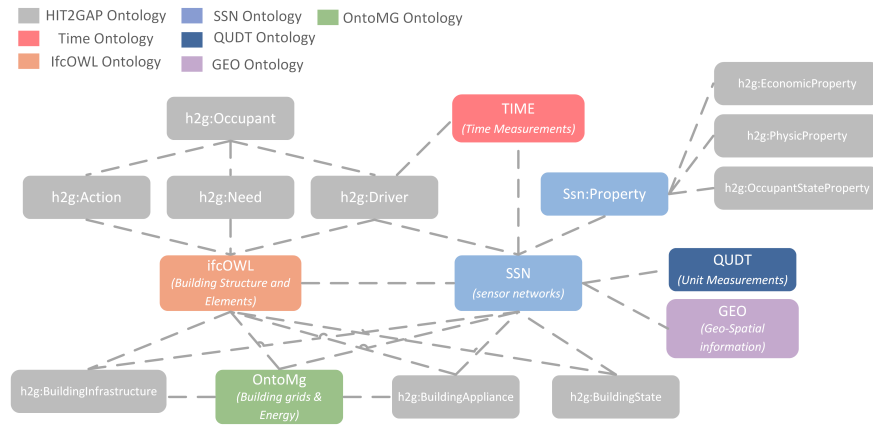


Fig. 2: OntoH2G overview

### 3 OntoH2G Adds

As shown in Figure 2, OntoH2G extends the existing solutions by proposing new concepts related to: (i) Occupant/User and simulation (*h2g:Occupant*, *h2g:Action*, *h2g:Need*, and *h2g:Driver*); (ii) Power system contextualization with OntoMG [6]; and (iii) Building and sensor network (*h2g:BuildingInfrastructure*, *h2g:BuildingAppliance*, *h2g:BuildingState*, *h2g:EconomicProperty*, *h2g:PhysicProperty*, *h2g:OccupantStateProperty*), with the aim of being able to cope with the project requirements, while remaining compliant with existing standards. The main adds are as follows.

**HIT2GAP adds:** Several concepts and properties have been added to cope with various aspects. The economic property ‘*h2g:EconomicProperty*’ makes reference to all monetary and investment aspects that could affect the building related domains. For example, this part of the ontology allows one to represent water-related prices, energy bills, building investments, etc. These specific phenomena have been subdivided into: Cost and Investment. The ‘occupant

<sup>2</sup> <http://www.qudt.org/release2/qudt-catalog.html>

<sup>3</sup> <https://www.w3.org/TR/owl-time/>

<sup>4</sup> <https://www.w3.org/2005/Incubator/geo/XGR-geo-ont-20071023/>

property' *h2g:OccupantStateProperty* describes the state of an occupant of the building. For example, this part of the ontology allows to represent the pulse and stress of an occupant. The physical property '*h2g:PhysicsProperty*' represents all related physical aspects that can be observed: dimensional, volumetric, and radiation aspects that influence the building and occupant behaviour. These phenomena support the alignment and fusion of specific observations based on general measurements (e.g., temperature, volume, etc.).

**Extensions related to sensor network infrastructure:** SSN represents a (building) sensor network as a system composed of a set of sub-systems or sensors. Sensor measurements (*ssn:Observation*) refer to measurements generated either by building devices or by people interacting with the building (e.g., changes in temperature set-points). An Observation is characterized by: (i) the sensor performing the measure under the concept *ssn:Sensor*; (ii) the object/property measured *ssn:Property* that typically refers to the physical value captured by the sensor (e.g., temperature, CO2 rate, etc.); (iii) the context of the object being monitored or observed (*ssn:FeatureOfInterest*), that can refer to a specific zone in the building; and (iv) the specific measurement as a result of the observation (*ssn:Result*). An observation can be associated with an external ID (which refers to the time series data) in order to allow distributed storage solution. This model allows contextualizing all related sensor information, the type of information that is being observed, as well as the information regarding the sensor/device/system installed (thanks to the inheritance between elements, in particular *ifc:Building*, *h2g:BuildingAppliance*, *ifc:SpatialElement*, *ontomg:MicroGrid*, *ifc:BuildingElement*, and *ssn:FeatureOfInterest*). The *ssn:FeatureOfInterest* is fundamental for aligning the different ontologies. The sequence of observations can be analysed, since it forms a time series that can be represented using a combination between the classes offered by the SSN ontology (to specify the values of the time series) and the QUDT ontology (to provide mechanisms for representing specific measures and data conversions). Time intervals, series and temporal parameters of the observations are represented using W3C-Time ontology. This facilitates the representation, abstraction, and conversion of temporal measurements. In detail, the *time:Instant* and *time:DateTimeInterval* are used to enable an observation represent a single value at an instant or multiple values (i.e., time series) during a certain period of time.

**Extensions related to occupants:** OntoH2G Occupants module relies on the combination of Drivers, Needs, and Actions with the representation of the building infrastructure (i.e., aligning the model with the ifcOWL) and different observations that could occur in the building (SSN ontology). Building Actors (*ifc:Actor*) make reference to people that interact with the building through a set of devices (managers) or events to determine the occupancy of different zones (*h2g:Occupant*). Hence, the class *ifc:Actor* is represented as an entity that interacts with building devices (*ssn:Device*) for sensing building infrastructure or even interacts with different elements of the building to satisfy their defined needs (through *ifc:Actuator*). All actors interactions depend on their role (*h2g:Role*) that varies from being a building manager, a building owner, a building tech-

nician, to a simple occupant. We distinguish a specific actor (*h2g:User*) that refers to any person who uses the HIT2GAP application. Such users are able to receive notifications (e.g., recommendations, alerts, etc.) regarding the building systems and able to gather feedback about energy-related status of the building according to their preferences (e.g., SMS, Email, etc.). Drivers (*h2g:Driver*) correspond to measurements of building systems that interact with occupants. Therefore, drivers are the glue between actions (*h2g:Action*) and occupant needs (*h2g:Need*). Thus, they are member of sensed observations (*ssn:Observation*), as if each of the observations required for simulating a user behaviour. Four different kind of drivers are considered: (i) Behaviour Model: to measure occupants habits and activities; it is defined to register and contextualize the occupant events (building interactions with the different systems); (ii) Activity Model: for registering the occupants activities (related to the occupancy); (iii) Event Model: to register the events that can be detected from sensor readings; and (iv) Equipment Model: for measuring the building features and environmental factors including indoor and outdoor parameters.

**Extensions related to Power System:** Bearing in mind that a smart building can be seen as a network connecting smart equipment that are mainly power-based ones, OntoH2G proposes an energetic contextualization of a building within a microgrid through the integration OntoMG [6], an ontology-based information model for modelling power system equipment. The grid (*ontomg:Microgrid*) is based on several branches that link the building infrastructure with the generated building renewable flows, non-renewable flows, and the different building systems linked with that flows. It was aligned to ifcOWL through several concepts:  $\langle \textit{ontomg:BranchController}, \textit{isA}, \textit{ifc:Controller} \rangle$ ;  $\langle \textit{ontomg:InfraBranch}, \textit{isA}, \textit{ifc:DistributionCircuit} \rangle$ ;  $\langle \textit{ontomg:DERBranch}, \textit{isA}, \textit{ifc:DistributionCircuit} \rangle$ ;  $\langle \textit{ontomg:ESBranch}, \textit{isA}, \textit{ifc:DistributionElement} \rangle$ ; and  $\langle \textit{ontomg:ELBranch}, \textit{isA}, \textit{ifc:DistributionElement} \rangle$ .

## 4 Discussion

Despite numerous tools, systems, and knowledge-driven models that have been constructed and published, there exists a gap between simulated energy and real consumptions. HIT2GAP platform aims at solving this reduction by offering a scalable platform, that relies on OntoH2G, a knowledge-based model. OntoH2G uses the strengths of standards and/or commonly adopted models to represent buildings, sensors, and information exchanged. Moreover, the model goes beyond modeling users and occupants by including their interactions with the building. In detail, OntoH2G also advances in: (i) the alignment of well-known ontologies on different domains to convey energy building concepts; (ii) the simplicity (by filtering out complex relations found in ifcOWL); and (iii) the user/occupant behaviour, preferences, and interaction representation. By interrelating these aspects, the HIT2GAP Core Platform complements existing building management tools being extensible and scalable to a broad of buildings. Thus, the platform is capable of conveying different systems from different vendors allowing the decision-making model to be more integrated and adaptable to the current

building situation. The ontology was mainly designed by LIUPPA and Eurecat based on the partners requirements, and on the Core Platform APIs specifications (brought by Nobatek). An iterative validation is currently in process, in which LIUPPA, Eurecat, Ege, and Nobatek are all involved <sup>16</sup>.

## Acknowledgements

The HIT2GAP project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 680708.

## References

1. M. Compton, P. Barnaghi, L. Bermudez, and et al. The SSN ontology of the W3C semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17:25 – 32, 2012.
2. M. Fernández-López, A. Gómez-Pérez, and N. Juristo. Methontology: From ontological art towards ontological engineering. In *Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series*. American Association for Artificial Intelligence, Marzo 1997. Ontology Engineering Group ? OEG.
3. T. Hong, S. D'Oca, S. C. Taylor-Lange, W. J. Turner, Y. Chen, and S. P. Corgnati. An ontology to represent energy-related occupant behavior in buildings. part ii: Implementation of the {DNAS} framework using an {XML} schema. *Building and Environment*, 94, Part 1:196 – 205, 2015.
4. L. Kallab, R. Chbeir, P. Bourreau, P. Brassier, and M. Mrissa. Hit2gap: Towards a better building energy management. *Energy Procedia*, 122:895 – 900, 2017.
5. P. Pauwels and W. Terkaj. Express to owl for construction industry: towards a recommendable and usable ifcowl ontology. *Automation in Construction*, 63:100–133, 2016.
6. K. Salameh, R. Chbeir, H. Camblong, G. Tekli, and I. Vechiu. A Generic Ontology-Based Information Model for Better Management of Microgrids. *Advances in Information and Communication Technology*, 458:451–466, 2015.
7. M. Uschold and M. King. Towards a methodology for building ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI*, 1995.

---

<sup>16</sup> Partners' data is confidential and cannot be published