

Proceeding Paper

The EN-TRACK Energy Efficiency Performance Tracking Platform for Benchmarking Savings and Investments in Buildings. Data Model Development [†]

Edgar Martínez-Sarmiento ^{1,*} , Stoyan Danov ², Eloi Gabaldon ¹ and Jordi Carbonell ²

¹ Building Energy and Environment Group, Centre Internacional de Mètodes Numèrics a l'Enginyeria, CIMNE-Lleida, Pere de Cabrera 16, Office 2G, 25001 Lleida, Spain; egabaldon@cimne.upc.edu

² Building Energy and Environment Group, Centre Internacional de Mètodes Numèrics a l'Enginyeria, GAIA Building (TR14), Rambla Sant Nebridi 22, 08222 Terrassa, Spain; sdanov@cimne.upc.edu (S.D.); jordi@cimne.upc.edu (J.C.)

* Correspondence: ealexis@cimne.upc.edu

[†] Presented at the Sustainable Places 2021, Rome, Italy, 29 September–1 October 2021.

Abstract: This paper is related to the H2020 project EN-TRACK, dedicated to developing a platform for gathering data on the performance of energy efficiency investments in buildings. The project aims to collect and harmonize data from different sources and provide services supporting investors and building owners in decision-making and de-risking building retrofit projects. The paper focuses on the methodology and the semantic technologies used in the development of the platform's data model, which enables the interoperability of data, and supports the service functionalities for tracking building energy performance and benchmarking savings from energy efficiency investments.

Keywords: energy efficiency; de-risking investments; building sector; data harmonization; data inter-operation; data model



Citation: Martínez-Sarmiento, E.; Danov, S.; Gabaldon, E.; Carbonell, J. The EN-TRACK Energy Efficiency Performance Tracking Platform for Benchmarking Savings and Investments in Buildings. Data Model Development. *Environ. Sci. Proc.* **2021**, *11*, 11. <https://doi.org/10.3390/environsciproc2021011011>

Academic Editor: Zia Lennard

Published: 24 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the principal challenges of increasing energy efficiency investments (EEI) is the lack of statistical data on the actual energy and costs savings achieved with them. On the other hand, the recent building and information technologies enable building owners, operators, facility managers, energy service companies, and financial institutions to collect increasing quantities of data. However, this data is still hard to access, aggregate, share, and utilize because it is housed in separate databases and stored in different formats, using ambiguous naming conventions [1]. Consequently, only a small part of the available information can be analyzed, compared, and benchmarked to produce reliable empirical evidence on the performance of the EEI necessary for decision makers and investors.

The EN-TRACK project addresses this challenge by developing a standard and an open data model for data harmonization from different sources, and using it in the core of its platform, thus enabling its interoperability. The alignment with most currently active databases and tools will lead to an unambiguous data exchange and service ecosystem with low transaction costs. This is a big step towards making energy efficiency investments a mainstream activity of the financial sector.

The present article briefly outlines the scope of the EN-TRACK project and focuses on the approach and semantic tools used in the development of the data model and presents the preliminary results and conclusions.

Section 2 presents the EN-TRACK project and objectives. Section 3 presents the methodology for the development of the data model. Section 4 presents the results and the initial data model. Section 5 presents the conclusions and future steps.

2. The EN-TRACK Project

The overall aim of the Energy Efficiency Performance-Tracking Platform for Benchmarking Savings and Investments in Buildings (EN-TRACK) is to enable an inter-operable ecosystem of data and tools in the area of building energy efficiency, supporting the technical and financial decision-making process of retrofitting the existing building stock. Built on an existing proven infrastructure (SHERPA [2] and EDI_Net [3]) and developed by the International Centre for Numerical Methods in Engineering (CIMNE), the EN-TRACK platform aims to enable massive data gathering, make them comparable and interoperable, and offer relevant results to key stakeholders by analyzing these data. This will be achieved by providing an open-source big data platform capable of acquiring and harmonizing data from multiple sources on the base of internal standardized description of building data and energy efficiency measures (EEM) and the standardized data-driven approach to evaluate EEI performance; attracting large public building owners in continuous collection and sharing of building data; enabling interoperability with most currently active databases and tools (DEEP [4], eQuad [5], EnerInvest [6], etc.); and developing and testing various business models and value allocation arrangements among the stakeholders in order to make the action self-sustainable and adaptable in the long term.

EN-TRACK shares objectives and is synergistic with other currently developed H2020 projects (EEnvest [7], Triple-A [8], Quest [9]); however, unlike others, its focus is placed on the data harmonization and interoperability between databases and platforms, which is achieved through the development of a common open data model for buildings.

While offering a wide scope of benchmarking and investment de-risking services and continuous gathering of empirical information on EEIs' performance, EN-TRACK will enable joint services with third-party platforms through unambiguous data exchange, setting the base of unprecedented ecosystem of tools for EEI de-risking and decision-making support.

3. Methodology

3.1. Establishing of the Initial Scope

The data modeling methodology consist of four activities. First, in order to characterize EN-TRACK data model's data requirements, a collaborative process was driven, identifying main stakeholders and a set of potential services to define use cases. Based on this, the data sources and their relationships with third party databases were detected. Data fields from DEEP, eQuad and Enerinvest platforms, and the available data sources at the project pilots were revised and included to form an initial dataset that can satisfy the envisaged functionalities. Later, a conceptual data model was formulated by grouping concepts into classes and sub-classes and by defining attributes, which implied data type selection and development of taxonomies for the required enumeration fields.

3.2. Semantic Technologies and Terminology Alignment

With data interoperability in mind, class relationships were established by merging the Smart Application REference (SAREF [10]) and the BIM-based holistic tools for energy-driven renovation of existing residences (BIMERR [11]) ontologies. Buildings and energy-related terminologies were adjusted to the Building Energy Data Exchange Specification (BEDES [1]) dictionary and the Industry Foundation Classes (IFC [12]) data model. Additionally, financial terminology was aligned based on the Investor Confidence Project (ICP [13]) protocols to avoid the occurrence of redundancies and ensure compatibility with the third-party applications and databases where possible. These activities were done in parallel with a visual implementation of the model in UML.

3.3. Data Model Assessment

The data model assessment is an iterative process. Each iteration ends after version evaluation is completed. This fourth and last activity of the cycle was conducted by cross-checking the data model against the data sources and its compatibility with the current

data model version. Theoretically, the iterative process stops when no updates on the model are needed. The data model development process is outlined on Figure 1.

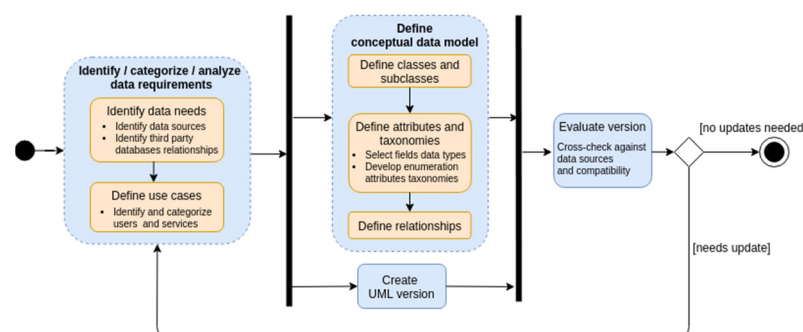


Figure 1. The EN-TRACK data model development process.

4. Results and Data Model

4.1. Data Sources

The EN-TRACK platform is fed with data from different sources through automated data transfer APIs, uploading data from files or manual data entry. The two project pilots' data considered for the data model formulation are the Spanish pilot, which consists of information regarding energy consumption from energy bills, applied energy efficiency measures, energy performance certificates, cadaster, electricity consumption, asset management and register of the properties; and the Bulgarian pilot, which comprises building renovation projects financed under various national programs, energy certificates and energy consumption from bills. Both pilots' data are complemented by manual entry data and open weather data, forming a heterogeneous dataset in different formats, such as spreadsheets, XML, web forms and databases endpoints. EN-TRACK will require a minimum set of essential data necessary to enable the services. By introducing additional optional data, additional services will be enabled for the users.

4.2. Data Model

The data model development started from the identification of the data needs for the implementation of the EN-TRACK functionalities. These comprise the gathering of statistical evidence on the performance of EEI in buildings and provision of services incentivizing the main stakeholders (building owners and operators, financial institutions, and policy makers) to provide the necessary data. The principal service groups include:

- Benchmarking of building energy performance;
- Evaluation of energy savings and financial return from EEMs;
- Benchmarking of EEMs' performance;
- Recommendations of EEMs.

The concepts recognized during the data definition phase were revised to identify overarching conceptual groups (classes) and their data fields (attributes). In an iterative and incremental process, the classes and attributes were extended by including items for compatibility with third-party databases, such as DEEP, and relevant existing ontologies. Special attention was placed on the taxonomies representing the constrained list of values for enumerated data type. Among many others, a new comprehensive taxonomy of EEMs was developed, for which no standard classification currently exists.

Once a wider picture of the classes was established, the data structure was designed in a unified modelling language (UML) diagram representing the classes' relationships and cardinalities, as seen in Figure 2. Despite the resemblance to a high complexity model, every single class and attribute has been studied and is supported by ontological artifacts. This model doesn't represent a final version and is meant to evolve and be extended over the time. The aforementioned methodology has enough flexibility and its application will ensure retro-compatibility with previous versions.

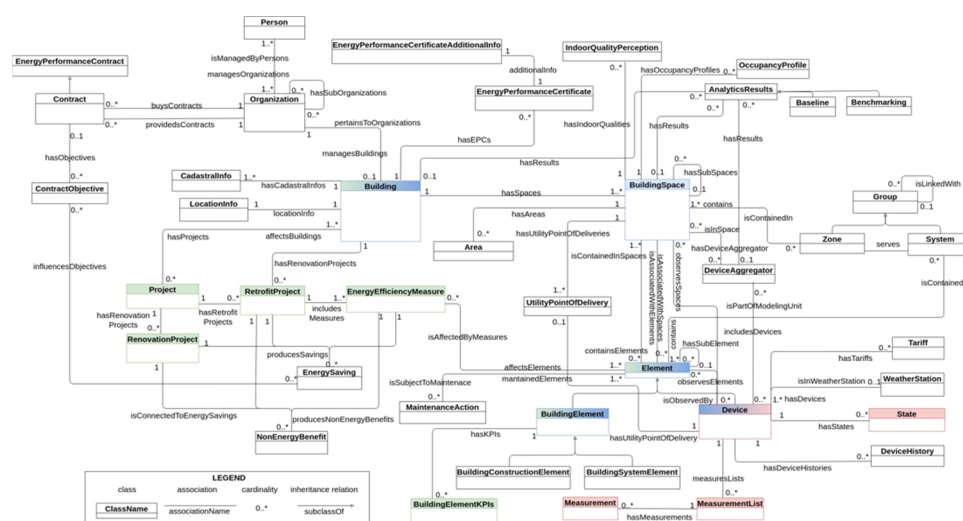


Figure 2. Simplified UML of the current EN-TRACK data model version.

5. Discussion and Conclusions

A methodology for the development of EN-TRACK's data model was successfully defined, which will also cover future iterations on the data model. Given that the model is planned to be extensible for capturing additional use cases and data inputs, the data modeling process would require updates periodically.

By making use of semantic artifacts (ontologies and dictionaries), the development of an inter-operable common data model is possible. Features of the semantic technologies will enable us to match data fields in the data model implementation phase.

Author Contributions: Conceptualization, E.G. and J.C.; methodology and writing—original draft preparation, E.M.-S.; writing—review, editing and supervision, S.D. All authors have read and agreed to the published version of the manuscript.

Funding: The EN-TRACK project received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement number 885395.

Acknowledgments: Edgar Martínez-Sarmiento acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the “Severo Ochoa Centre of Excellence (2019–2023) under the grant CEX2018-000797-S funded by MCIN/AEI/10.13039/501100011033” programme.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. DOE *Building Energy Data Exchange Specification Scoping Report*; Energy Efficiency & Renewable Energy: Washington, DC, USA, 2013.
2. SHERPA—SHared Knowledge for Energy Renovation in Buildings by Public Administration. Available online: <https://sherpa.interreg-med.eu> (accessed on 12 August 2021).
3. EDI-Net—The Energy Data Innovation Network. Available online: <https://www.edi-net.eu/> (accessed on 12 August 2021).
4. DEEP—De-Risking Energy Efficiency Platform. Available online: <https://deep.eefig.eu> (accessed on 18 August 2021).
5. EQuad Platform: A Bridge to Sustainable Energy Efficiency Financing. Available online: <https://www.eu.jouleassets.com/about-equad> (accessed on 2 September 2021).
6. EnerInvest: Spanish Sustainable Energy Financing Platform. Available online: <https://www.enerinvest.es/en/> (accessed on 2 September 2021).
7. EEnvest: Risk Reduction For Building Energy Efficiency Investment. Available online: <https://www.eenvest.eu/> (accessed on 2 September 2021).
8. Triple—A Project Enhancing at an Early Stage the Investment Value Chain of Energy Efficiency Projects. Available online: <https://www.aaa-h2020.eu/> (accessed on 9 September 2021).

9. Quest—Quality and Effectiveness in Science and Technology Communication. Available online: <https://questproject.eu/> (accessed on 9 September 2021).
10. SAREF—Smart Applications REference Ontology, and Extensions. Available online: <https://saref.etsi.org/> (accessed on 21 September 2021).
11. BIMERR—BIM-Based Holistic Tools for Energy-Driven Renovation of Existing Residences. Available online: <https://bimerr.iot.linkeddata.es/> (accessed on 21 September 2021).
12. IFC—Industry Foundation Classes. Available online: <https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/> (accessed on 23 September 2021).
13. ICP—Investor Confidence Project. Available online: <http://www.eepperformance.org/> (accessed on 23 September 2021).