SEMERGY: APPLICATION OF SEMANTIC WEB TECHNOLOGIES IN PERFORMANCE-GUIDED BUILDING DESIGN OPTIMIZATION

SUMMARY: This contribution presents a comprehensive summary of the first phase of the research project SEMERGY, aimed towards exploring the potential of Semantic Web Technologies to accommodate and support building performance assessment and optimization. Previous research on the adoption of performance assessment tools by the design community suggested that the extensive amount of time, effort and expertise required to accumulate and provide input information for various performance evaluations, limits the application of such tools for improvement of building design in early stages. Even the most rudimentary performance inquiries require large quantities of data regarding the properties of incorporated building materials and systems, climatic conditions, building codes, etc. Much of this data is available on the internet, yet its use is severely hampered by lack of appropriate structure and format. The SEMERGY research and development project was conceived with the aim of acquiring and utilizing the abundant, yet ill-structured building-related data available on the World Wide Web, through implementation of Semantic Web Technologies. This web-based data is utilized within a flexible, multi-objective design optimization environment, incorporating various performance assessment applications. The present contribution describes some of the major challenges of the project, discusses the adopted approaches, and presents SEMERGY's main features.

KEYWORDS: Semantic Web, Building Performance Modelling, Decision Support Systems, Thermal retrofit


COPYRIGHT: © 2015 The authors. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 unported (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
1. INTRODUCTION

Architects, planners and building scientists today are confronted with multiple challenges in the design and amelioration of the built environment. On one hand, the accelerated exploitation of natural resources and man-caused greenhouse gas emissions (major contributors to the Global Warming) call for an energy-efficient and environment friendly building sector. On the other hand, clients and public building administrators suffer the effects of the worldwide economic crisis, therefore prioritizing the economic feasibility of construction projects over sustainability considerations. Although the past decade has witnessed an upsurge in the development of performance assessment tools, their implementation has been limited to certification and post-validation activities, rather than performance-oriented design decision making (Hensen et al. 2004). The lack of general enthusiasm for performance-oriented design, on the side of the design community, may be attributed to various economical and technical challenges. The substantial amount of time and expert knowledge required to investigate the energy and environmental performance of different design approaches, leads to additional expenses and longer design phases (Reference). Furthermore, both clients and designers are affected by insufficient and inefficient access to data, necessary for detailed energy and cost assessments and decision making, including technical properties and costs associated with building materials, systems and products, building codes and standards, available subsidies and funding opportunities, etc. Other influential challenges include planning fallacy (Kahneman 2011) and uncertainties about decisions and data (Augenbroe 2010). According to DENA (2013) severe lack of market transparency and information opportunities are considered to be two major hampering factors for clients to perform thermal retrofits of their buildings and properties. As such, any developmental attempt at overcoming these obstacles, should insure efficient data streams, reduce user effort through automated processes, increase speed and enable multivariate considerations of the problem at hand.

SEMERGY addresses these problems by providing a web-based decision support and optimization environment, operable by novice users (clients) as well as professional members of the building design community. This environment proposes to surmount the data-related barriers to performance-oriented design through utilization of Semantic Web technologies (W3 2014) for semi-automatic data retrieval from web-based sources. The applicability of this approach for retrieval and use of building product data has been demonstrated in the first (beta) release of the SEMERGY environment (SEMERGY 2014). A multi-objective optimization process, supported by automatic identification of alternative design scenarios, reduces user effort and time-cost of the planning process.

SEMERGY is not a new performance assessment tool. Rather, it is intended to incorporate different calculation or simulation methods (e.g., Energy Plus 2014, Archiphysik 2014, PHPP 2014) for evaluation of energy, environmental, and financial performance of buildings. SEMERGY is intended to address evaluation and optimization of building retrofit projects, as well as new constructions. However, due to certain fundamental differences between optimization of already existing and newly conceived structures, SEMERGY’s beta release is focused on the building retrofit problem. The second phase of the SEMERGY project (ongoing) concentrates on the integration of new constructions and enhancement of the existing computational modules.

The present contribution reports on the main features of the SEMERGY environment, its development process, as well as major challenges and adopted approaches. The following topics will be discussed in detail:

i. Software structure, goals and expectations.

ii. Creation of a suitable building representation scheme, with special regard to the numerous – sometimes contradicting – requirements of the incorporated computational methods.

iii. Acquisition of geometric data through a graphical user interface (tailored to the requirements and expertise of multiple user groups) or via import from BIM/CAD applications to populate the building data model.

iv. Incorporation of semantic web-technologies for semi-automatic extraction, reorganization, and population of semantic data.

v. Semi-automatic identification of legitimate alternative constructions as optimization candidates. These alternatives are evaluated and ranked towards their impact on the energy and ecological performance of the building as well as the required investment cost.
vi. Flexibility and adaptability for future integration of computational methods and codes of different countries.

2. RELATED RESEARCH

2.1 Competitor Analysis

There are quite some national and international approaches to offer design and retrofit decision support: The tools differ in their price policy (free / open source / priced), their scale (components, entire building, urban scale), their methodology (simple calculation, normative approaches, simulation), and their technical realization (browser-based tools, tools for download and installation, implemented technologies). Cetin and Mahdavi (2010) explored the availability and usability of web-based performance simulation tools, and identified a large number of different tools. The authors of this contribution – however – re-explored about three years later a major portion of these tools, to find that the majority of tools was no longer available. The following list is based on the competitors analysis continuously performed in the course of the SEMERGY development. It is neither meant to be exhaustive nor can it be representative of all offered tools in the World Wide Web:

- EnergyGlobe: The website energyglobe.com offers a set of tools for evaluation of new and existing buildings, as well as for determination of the energy demand of domestic appliances. Existing buildings can be evaluated via a “virtual energy certificate” based on a set of user-defined parameters. However, the user preferences are limited to predefined values and templates for most parameters. The exact specification of the building’s geometry is not possible.
- Online standard-based Energy-Calculators: German and Austrian vendors offer online building performance calculation services. However, for the most part these services do not include an automated calculation but rather data acquisition is facilitated through the World Wide Web while the actual computation is performed by experts in the field of energy assessment. Automated services offered need to be evaluated for accuracy and precision since the input data is limited to predefined templates (e.g. Express-Pass 2014). Furthermore, it needs to be mentioned that calculations purely based on rudimentary user-input do not fulful the quality standards of official energy certificates, even if claimed so.
- The newly announced Google spin-off project, Google Flux (Google 2014), intended to address similar questions as SEMERGY, confirms the validity of and global request for such an approach towards building design optimization.

2.2 Semantic Web Technologies

The notion of Semantic Web was first introduced by Berners–Lee et al. (2001) as potential future of the World Wide Web, in which data is offered in a machine-understandable format on the web, through incorporation of meta-data in the websites and establishing links and relations between various concepts and documents. The term ontology, borrowed from the field of linguistics, refers to “a specification of a representational vocabulary for a shared domain of discourse: definitions of classes, relations, functions, and other objects” (Gruber 1992). This concept evolved into a graph based web-supported data structure, in which objects (relevant to a certain field) are described through classes and properties, and relations between objects or classes can be defined. This “ontological” representation of data enables sophisticated queries as well as rule-based inference of new information. The utility of linked semantic data ontologies in the field of medical research and life sciences has already been well established (W3 2014).

Shayeganfar (2008) demonstrated the applicability of the same approach in retrieval and structuring of building product information with the example of various window products. However, a more comprehensive exploration of the potential of Semantic Web technologies for utilization of various web-based building-related data was aimed at in the SEMERGY project (Mahdavi et al.2012a, 2012b).

An effort toward the representation of building material and component data through Linked Data was reported by Radinger et al. (2013). Their project “BauDataWeb” examined the availability and location of building and construction materials within Austria aiming to offer e-business support to the building material industry. This project facilitates the search for product providers in the vicinity of the construction site, and enables product search based on product type and/or various properties of products (e.g., thickness, thermal conductivity).

ITcon Vol. 20 (2015), Pont et al., pg. 109
However, this project does not approach building products from a technical and engineering point of view. In a performance oriented decision support system, which calls for a sophisticated underlying logic for automatic identification of alternative design approaches, a higher range and resolution of data is required to properly describe various features of building products.

3. SEMERGY STRUCTURE

As mentioned before, the SEMERGY project was not aimed towards development of new computational engines for various performance aspects of buildings. Rather, it was envisaged as a core component of a flexible and extensible environment using existing methods and calculation engines. This core incorporates a central data model, a multi objective optimization component, and three interfaces establishing links to various other components that can be plugged into SEMERGY, thereby expanding or customizing its functionalities. These three interfaces are the following:

1. The reasoning interface, providing smooth transactions with various performance assessment engines (including heating demand calculator, cost estimator, OI3 index calculator, etc.)
2. The semantic interface, enabling access to semantic data ontologies, thereby facilitating utilization of the scattered web-based resources pertaining to building products and materials, codes and regulations, funding opportunities and subsidies, etc.
3. The user interface, enabling user interaction and the communication of design and user preferences to SEMERGY.

Figure 1 illustrates the modular structure of the SEMERGY environment.

![SEMERGY's principal structure](image)

**FIG. 1:** SEMERGY’s principal structure.

4. SEMERGY BUILDING MODEL

In order to facilitate the transfer of data between various components of the SEMERGY environment, and to provide a solid framework for the formation of semantic data ontologies, a building representation had to be adopted. This representation had not only to comply with the requirements of various intended performance assessment procedures, but also to enable extensions for future functionalities. Initially, our intention was to select among the existing standard building representations.
Prior to the investigation of existing standard building models, some intended performance computational engines (thermal simulation engines, steady state heating demand calculators) were reverse-engineered to determine the scope and format of the required input information. In a second step, the Industry Foundation Classes (IFC 2014) developed by the Building Smart group as a universally acknowledged platform-neutral building data model, and Green Building XML (gbXML 2014), developed to facilitate transfer of building data from BIM software for analysis purposes, were analysed in detail.

The IFC model was dismissed due to its lack of a space-based structure, a fundamental requirement for a majority of performance analysis procedures, and the high computational cost of the necessary conversions for the procurement of such a structure from an IFC model.

Having been developed with performance assessment in mind, gbXML was structurally compliant with SEMERGY's functionalities. However, the impossibility of extending the model (for instance with more detailed material representations) and the multiplicity of representation formats within the model rendered it unsuitable for the project.

Additionally, a third model was studied: The Shared Object Model SOM (Mahdavi 1999, Mahdavi et al. 2002). It was developed in the 1990's as part of the SEMPER project, an early effort towards automation of performance assessment. The SOM was developed through an iterative bottom-up approach. It was not intended to be a general building model for a broad array of building applications. Rather, it was designed to allow users to generate a building model that allows the conduction of performance simulations with a number of preselected simulation engines. Since the outline of SOM was developed based on a range of performance assessment procedures (thermal, acoustical, etc.), and since representation of space, its boundaries and their properties are fundamental in such computations, SOM offers a space-based view of a building. Although not a standard building representation schema, it provides a relatively uncomplicated structure, which was adopted as a basic framework and expanded according to the previously identified informational requirements of the targeted computational procedures. The resulting building data representation, known as SEMERGY Building Model (SBM, Ghiassi 2013), is compliant with a wide range of performance inquiries (thermal, environmental, financial), is systematically map-able to standard performance compliant data formats and can be extended for compliance with other performance calculations (e.g., visual and acoustical). Within the SEMERGY environment, SBM is partly populated by user-provided data acquired through the user interface and partly by the data selected from the semantic data ontologies based on user preferences or system defaults. A structural overview of the SBM is provided in Figure 2.

Roughly, the SBM structures building data into four categories of objects and classes:

- **Calculation parameters:** These objects need to be understood as links to external repositories of calculation and weather data templates. Influenced by the user-provided physical data such as location and construction system, relevant parameters and values are extracted from these external repositories, stored, and supplied to the corresponding assessment tools. While initially provided manually to serve in the general proof of the SEMERGY concept, this is currently changed in terms of the general architecture of SEMERGY: Data is retrieved via the semantic interface, stored and administrated in the SEMERGY kernel and provided to the reasoning interface.

- **Operational data:** These objects include all settings regarding internal conditions of zones. It includes HVAC data and settings such as schedules and set points for temperature, relative humidity, ventilation and infiltration, internal gains through lighting, occupancy and equipment use, and information about the zone’s equipment and furniture. For novice users and simplified calculation engines pre-defined templates are provided to populate the corresponding classes in the SBM based on case-specifications. Furthermore, a property to space was added in the structure named “function” that serves as shortcut to predefined templates of operational data following the space’s function. For instance, if “function” is set to residential, default internal gains of 3.75 W.m⁻² and a ventilation rate of 0.4 h⁻¹ are assumed (as stated in Austrian standards).

- **Physical Data - Geometry:** These data classes include the whole geometrical representation of the building including building’s location and site, its sections, zones and spaces, and the representations of its enclosures, partitions and apertures. SEMERGY offers different opportunities to populate these

*ITcon Vol. 20 (2015), Pont et al., pg. 111*
classes, which are a web-based drawing tool and data import from CAD/BIM-tools. Moreover, data import from Geographical Information Systems is currently under development.

- Physical Data - Semantics: This category summarizes all classes with semantic information about building constructions, layers and material necessary for calculation and simulation purposes. To populate these kinds of information, the SEMERGY environment utilizes a comprehensive ontology of building products, supported by semantic web technologies and structured in a SBM-compatible way. For novice users – comparable with the operational data – default options are integrated. For instance, if a user does not know about the constructive details of his existing building, SEMERGY assumes default data about constructions from the specific construction years. Similar, simplifications concerning classes such as shading, screens and blinds are implemented in the SEMERGY environment.

**FIG. 2: Structural overview of the SEMERGY Building Model (SBM)**

*ITcon* Vol. 20 (2015); Pont et al., pg. 112
5. POPULATING SBM

5.1 The Graphical User Interface

As mentioned before, SEMERGY is intended both for novice users and professionals. Therefore, the simplicity, speed, and clarity of data entry methods are of utmost importance. For the acquisition of the geometrical representation of the building and user preferences regarding construction, internal conditions, systems and optimization criteria, and threshold values, a web-based Graphical User Interface (GUI) was developed. The location of the building, main usage, period of construction, number of floors above and below ground, roof type and condition, and other general information are entered by choosing among the available options (often accompanied by graphical clues).

The SEMERGY GUI allows users to draw floor plans in a very simple CAD-like environment using different line types for elements with different material compositions (e.g., external walls, fire walls, internal walls). Doors and windows are drawn in a similar fashion and the sill height and element height are entered in terms of numerical values. Once the plan is drawn, SEMERGY's internal logic identifies spaces, space enclosures, and apertures and constructs a three dimensional representation of the building in SBM format. The user is asked to select a “function” for each identified space from a pre-defined list of options (in case of residential buildings, bedroom, kitchen, unheated storage room, etc.). These functions act as shortcuts in the system back-ground allowing the retrieval of the associated internal conditions from the internal template repositories. Figure 3 shows a screen-shot of SEMERGY’s GUI, illustrating identified spaces and apertures as well as the selection options for different room functions and room heights.

![Screen-shot of SEMERGY's GUI](image)

**FIG. 3**: SEMERGY’s Graphical User Interface.

Once the geometry and the space functions are defined, user preferences regarding the semantics of various building elements are entered. In order to simplify the acquisition and entry of material related data, the user is
asked to select the desired overall construction type of the building (e.g., load bearing masonry walls, wood frame construction). Accordingly, the user is presented with a list of construction configurations for each building element (e.g., brick wall with external insulation, or double layered brick wall with air cavity). The user can modify the thickness of layers in the selected configuration to obtain the desired construction (Figure 4).

These templates are supported with detailed properties and descriptions of the composing materials, provided by the building product ontology through SEMERGY Semantic Interface.

FIG. 4: SEMERGY’s Graphical User Interface: Definition of construction properties.

5.2 Data import from CAD & BIM applications

Although the input option via GUI can be of interest to professional users in the early design stages, a systematic link to commonly used CAD tools is necessary to ensure usability for the professional design community. As mentioned before, SBM and gbXML, the industry standard data exchange format for BIM applications, share the same space-based view of buildings. As such, despite their differences in capturing various aspects of a building, the two schemas can be conveniently mapped to each other. A mapping schema has been developed to transform
gbXML geometry data into SBM format. However, due to the low resolution of construction descriptions in gbXML, the semantics are not fully acquired from this data model. The import function is to be implemented in the SEMERGY environment in the next phase of the project.

BIM applications offer many advantages. Nonetheless, they are typically rather complex. Hence, the architectural design community is still strongly dependent on "less intelligent" drafting software. To accommodate users with a preference for conventional CAD, a plug-in for AutoCad (AutoCad 2014) was developed. The plug-in asks the user to pre-process the drawings and plans in AutoCad, using a layer system provided by the plug-in and a set of guidelines. The resulting .dwg file can then be imported into SEMERGY, where space identification is done and space attributes are specified. This working scheme supports the rapid extraction of information from existing CAD-plans and potentially also disburdens the implementation of paper-only plans of historic buildings for building performance evaluation.

6. UTILIZATION OF SEMANTIC WEB TECHNOLOGIES

Web-based data sources can be instrumental to addressing the challenge of data availability for various performance inquiries. However, the unstructured and haphazard trend of data presentation adopted currently by many data providers hampers the automatic and systematic retrieval of necessary data at a reasonable speed. The current structure of the web is centred on human cognitive abilities and does not focus on machine-readability. Most web-based documents are provided for human users and as such, lack the structure and embedded intelligence required for machine usability. In order to be able to use this vast informational potential in automated processes, measures have to be taken to restructure the available data into machine-readable format.

Semantic Web Technologies offer an opportunity towards restructuring the current Web of Documents to Web of Data by providing means to inter-link data from various web sources, to re-categorize the data according to a more consistent logic (via ontologies), and to augment its intelligence through various semantic rules and reasoning processes.

As a proof of concept, SEMERGY focuses on the available pool of building product information. In order to demonstrate the abilities of the adopted strategy, two major web-based data repositories, BauBook (BauBook 2014) and MASEA (MASEA 2014) were selected. Theoretically, the Semantic Web approach can be applied to any web-based resource (e.g., manufacturer websites). However, inconsistencies in formatting and presentation on the side of data providers can render such efforts fruitless. Therefore, data repositories with consistent data structures and a large quantity of data were preferred for the demo.

The BauBook categorization schema was adopted and refined to serve as the reference class hierarchy and basis of the SEMERGY Ontology. This categorization schema identifies groups of materials with similar functional attributes (e.g., insulation materials, bricks).

The representation of materials in SEMERGY Building Model (SBM) was adopted for the building product ontology to insure data compatibility. Data pertaining to each product was obtained from the web-source and organized under the corresponding attribute. In case of presence of a single product in both repositories (i.e. BauBook and MASEA), the data derived from both sources merged into a richer product profile. In a second phase, cost data from various manufacturer or reseller websites were added to the acquired physical attributes (Figure 5). The material profiles preserved links to the original data sources, providing an opportunity for regular updates and acquisition of new information, as long as the original data structures remained constant.

To accommodate the requirements of the automatic alternative identification procedure (explained in the following sections), the Ontology categories were augmented with new attributes and characteristics (such as format or application position), which enabled the intended rule-based queries. This additional information was inherited by each product under that specific class. This inheritance capability allows for the effortless assignment of the required attributes to new materials from new web repositories through simple assignment of the corresponding super class.

Based on the preferences of users as well as technical considerations and restrictions regarding the material composition of different building elements, the product Ontology populates SBM with the detailed material information of the base case. Using its inherent rule-based logic, SEMERGY then identifies alternatives for this base case. The data pertaining to alternative scenarios is subsequently retrieved from the Ontology and supplied to SBM for optimization computations.
7. UTILIZATION OF SEMANTIC WEB TECHNOLOGIES

For the optimization of building’s performance, SEMERGY needs to be able to suggest alternatives for incorporated materials and constructions. As mentioned before, SEMERGY’s current release focuses on refurbishment cases. Therefore, approaches for optimization of the building’s performance are limited by constraints of the existing building. For instance, certain parts of the existing building need to be excluded from optimization, as per the client’s request, or due to heritage protection. Furthermore, the thermal improvement of certain constructions (e.g., floor slabs, basement walls) might not be feasible due to unreasonably difficult and thus cost-intensive realization in existing structures. Another hampering issue for improvement strategies might be the building owner’s lack of knowledge on detailed composition of the building’s components.

SEMERGY offers based on the year of construction, default construction options for the building’s thermal envelope. These can be accepted or modified by the user. Users have the possibility to change details of these templates. For instance, if an old building component has previously undergone retrofit, it is possible to consider this adjustment in the base-case evaluation. Additionally, the user can determine which constructions (and layers thereof) should be included in or excluded from optimization efforts.

In addition to the mentioned user-specified input, SEMERGY needs rules for identification of legitimate construction alternatives. These rules facilitate on one hand the compliance with the building codes and necessities of building construction (U-value, avoidance of condensation, acoustical parameters), and reduce on the other hand the number of potential refurbishment options to a realistic and manageable size. The rule-based optimization is illustrated for a historical attic floor construction in Figure 6.

For each historical template a set of potential retrofit designs are designated, and in each option, allowed building materials and properties for each layer are stated. To ensure the conformity of the resulting combinations with construction considerations, U-value and water vapor diffusion calculations are carried out. Combinations that do not fulfill the requirements are removed from the set of alternatives.

As illustrated in the example of Figure 6, SEMERGY suggests three constructive refurbishment concepts for the existing attic floor (that in its current state would not be able to fulfill the applicable thermal requirements): i) Adding insulation to both sides, including water vapour barrier and finish below; ii) Adding finish above and insulation and finishing below. iii) Adding insulation below.

Note that for all three refurbishment concepts, SEMERGY examines the suggested material combinations in view of their thermal and hygric behaviour. This means that constructions without water vapour barrier and condensation risk will only be suggested with suitable materials (e.g., insulation boards capable of absorbing condensation without damage risk).
This contribution illustrated efforts toward utilization of semantic web technologies in the field of semi-automated building performance assessment and optimization. The applicability of semantic data ontologies for the retrieval, restructuring, and enrichment of building product data has been demonstrated in SEMERGY’s beta rollout (SEMERGY 2014). The developed data ontology has been used along with an elaborate rule-based logic to identify and generate viable building component alternatives based on user preferences and technical considerations. A multi-objective optimization procedure helps approximating pareto-optimal design solutions in view of overall heating demand, environmental impact, and investment cost. To our knowledge automated identification of various design and retrofit options and automated multi-objective optimization are currently not supported by any competitive development for building performance oriented design. Moreover, no competitor or ongoing research project known by the authors, offers user the possibility to enter high-resolution building data through a simplified web-based graphical user interface. SEMERGY automatically augments the user-provided two-dimensional drawings into a space-based three-dimensional building representation; thereby reducing the complexity of data entry while maintaining the possibility of running complex performance enquiries.

Previous states of the progress of this research project in general, as well as detailed descriptions of various features are presented in Ghiassi (2013), Ghiassi et al. (2012, 2013, 2014), Hammerberg et al. (2013), Heurix et al. (2013), Mahdavi et al. (2012a, 2012b), Pont et al. (2013), Pont (2014), Pont et al. (2014) Sadeghi Kafri (2014), Shayeganfar et al. (2013) and Wolosiuk et al. (2013).

9. FUTURE DEVELOPMENT & ENRICHMENT OF THE SEMERGY ENVIRONMENT

The presented achievements are the results of the first phase of the SEMERGY project. The second phase of the project started in June 2014 and will focus on the following: i) Integration of optimization for new buildings. For the new buildings, changes and expansion of the templates and ontology are expected. Specifically, the templates of constructions will be expanded to cover a large number of typical up-to-date constructions. ii) Integration of building systems and services into the SEMERGY environment, including active solar components. iii) Integration of new calculation modules addressing summer overheating and partial building codes compliance check. iv) Improving and refining the cost calculation scheme currently integrated in SEMERGY.
Both the SBM and building product Ontology were developed with extensibility in mind. Therefore, expected modifications applied to these features will only be of a minor nature. SEMERGY is currently undergoing a systematic usability evaluation to help identify necessary future adjustments.

10. ACKNOWLEDGEMENTS

The SEMERGY project is funded under the FFG Research Studio Austrian Program (grant No. 832012) by the Austrian Federal Ministry of Economy, Family and Youth (BMWFJ). In addition to the authors, the SEMERGY team includes A. Anjomshoaa, K. Hammerberg, I. Merz, T. Neubauer, F. Shayeganfar, C. Sustr, M. Taheri, D. Wolosiuk and A. Wurm.

11. REFERENCES


ITcon Vol. 20 (2015), Pont et al., pg. 118


