<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Building performance optimization using cross-domain scenario modeling, linked data, and complex event processing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>O'Donnell, James T.; Corry, Edward; Hasan, Souleiman; et al.</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2013-04</td>
</tr>
<tr>
<td><strong>Publication information</strong></td>
<td>Building and Environment, 62 (7): 102-111</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Elsevier</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/4842">http://hdl.handle.net/10197/4842</a></td>
</tr>
<tr>
<td><strong>Publisher's statement</strong></td>
<td>This is the author's version of a work that was accepted for publication in Building and Environment. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Building and Environment (62, , (2013)) DOI: <a href="http://dx.doi.org/10.1016/j.buildenv.2013.01.019">http://dx.doi.org/10.1016/j.buildenv.2013.01.019</a></td>
</tr>
<tr>
<td><strong>Publisher's version (DOI)</strong></td>
<td>10.1016/j.buildenv.2013.01.019</td>
</tr>
</tbody>
</table>
Building Performance Optimization Using Cross-Domain Scenario Modeling, Linked Data, and Complex Event Processing

Corresponding Author: James O’Donnell - jtodonnell@lbl.gov
Phone: +1-510-486-4146, Fax +1 510-486-4089
Building Technologies Group; Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90R3111
Berkeley, CA 94720

Author: Edward Corry - edwardcorry@nuigalway.ie
Phone +353 65 68 23721
Informatics Research Unit for Sustainable Engineering,
National University of Ireland, Galway

Author: Souleiman Hasan – Souleiman.hasan@deri.org
Phone +353 91 492973
Digital Enterprise Research institute, National University of Ireland, Galway.
IDA Business Park, Lower Dangan, Galway Ireland.

Author: Edward Curry - ed.curry@deri.org –
Phone +353 91 492973
Digital Enterprise Research institute, National University of Ireland, Galway.
IDA Business Park, Lower Dangan, Galway Ireland.

Author: Marcus Keane – marcus.keane@nuigalway.ie
Phone +353 91 492619
Informatics Research Unit for Sustainable Engineering
National University of Ireland
Galway, Ireland
Building Performance Optimization Using Cross-Domain Scenario Modeling, Linked Data, and Complex Event Processing

James O’Donnell\textsuperscript{a}, Edward Corry\textsuperscript{b}, Souleiman Hasan\textsuperscript{c}, Marcus Keane\textsuperscript{b} Edward Curry\textsuperscript{c}

Corresponding Author

\textsuperscript{a}Building Technologies Group; Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90R3111 Berkeley, CA 94720, jtodonnell@lbl.gov, – Phone: +1-510-486-4146

\textsuperscript{b}Informatics Research Unit for Sustainable Engineering,
National University of Ireland, Galway,
edwardcorry@nuigalway.ie, marcus.keane@nuigalway.ie
Phone 353 65 68 23721

\textsuperscript{c}Digital Enterprise Research institute, National University of Ireland, Galway.
IDA Business Park, Lower Dangan, Galway Ireland.
Souleiman.hasan@deri.org, ed.curry@deri.org, Phone +353 91 492973,

Abstract

The scenario modeling method empowers building managers by enabling comprehensive performance analysis in commercial buildings, but is currently limited to data from the building management domain. This paper proposes that Linked Data and Complex Event Processing can form the basis of an interoperability approach that would help to overcome technical and conceptual barriers to cross-domain scenario modeling. In doing so, this paper illustrates the cross-domain potential of scenario modeling to leverage data from different information silos within organizations and demonstrates how to optimize the role of a building manager in the context of his or her organization. Widespread implementations of cross-domain scenario models require a solution that efficiently manages cross-domain data acquisition and post processing underpinned by the principles of linked data combined with complex event processing. An example implementation highlights the benefits of this new approach. Cross-domain scenario models enhance the role of the building manager within an organization and increase the importance of information communicated by building managers to other organizational stakeholders. In addition, new information presented to stakeholders such as facilities managers and financial controllers can help to identify areas of inefficiency while still maintaining building function and optimized energy consumption. Two key challenges to implementing cross-domain scenario modeling are: the data integration of the different domains’ sources, and the need to process scenarios in real time. This paper presents an implementation approach based on linked data to overcome interoperability issues, and Complex Event Processing to handle real-time scenarios.

Keywords: Linked data, complex event processing, performance metrics, building performance analysis

1 Introduction

Unavailable, unreliable, and inaccurate building performance information is a major cause of inefficient building operation [1,2]. Information used by building managers must be reliable, but no standards are currently available for
the analysis and interpretation of building performance data. In addition, current methods and tools fail to account
for the profile of building managers, both in terms of the operational context of their role and their typical technical
and educational background [3]. As a result, the information communicated to building managers can result in
decisions that are often ad-hoc, arbitrary, and incomplete [4].

In practice, analysis methods applied to building operations vary in complexity. Performance benchmarks usually
originate from prescriptive code compliance, externally established energy performance guidelines, whole-building
energy simulation results, and rules of thumb or conventional wisdom [5,6]. Such benchmarks are typically difficult
to disaggregate, as they fail to contain enough metadata for a comparison with systems and components in a given
building. Mismatches between the functionality provided by information systems and that required by designers
during the design stage have been recognized in the architecture, engineering, and construction (AEC) industry [9].
This lack of functionality is also evident in the operational phase of the building lifecycle, as building managers seek
to operate complex facilities in the absence of specifically tailored information flows driven by the profile of the
building manager [3]. For example, building managers might compare a building’s annual gross energy
consumption with data about its previous year’s performance or with normalized data from similar buildings, but
they may not have the data or technical resources to break down energy end use by type. Normative comparison
methods include those from the Netherlands Normalization Institute (NEN 2916), ENERGY STAR, and CIBSE
Guide F: Energy Efficiency in Buildings, as [10–12], as well as the more advanced U.S. General Services
Administration (GSA) Building Performance Assessment Toolkit, providing objective performance indicators that
are communicable between different project stakeholders [13].

However, these methods can fail when they involve building-to-building comparison. For example, if all the
compared buildings operate inefficiently, or the comprehensive, unique nature of the building, such as highly
specialized low energy heating and cooling systems, is not taken into account by the measurement method. In
buildings that use a calibrated whole-building energy simulation model, measured performance is extremely difficult
to compare with predicted performance, due to difficulties in creating like-for-like comparisons, especially at
different levels of granularity [14–16].

Structured methods applied by experts have resulted in energy conservation measures that have, on average, saved
over 20% of total energy costs and over 30% of heating and cooling costs, in more than 100 studied U.S. buildings
[17]. Experts brought in to “fix” building systems provide one possible solution to identifying causes of inefficient
operation in the building stock [18,19]; although, when the magnitude of the global building stock is considered, the
time, skills, and resources of experts needed to attend to inefficient operation suggest that an expert-led solution is
impractical.

O’Donnell [20] established that a global solution must primarily involve the stakeholder responsible for commercial
building energy management. A properly enabled building manager should be able to achieve savings similar to
those of an expert consultant. O’Donnell defined a tailored technique called scenario modeling that accounts for the
knowledge, skills, and experience of building managers, together with the available building data, and this technique
can be used by a building manager to drive the optimization process.

1.1 Scenario Modeling

Scenario modeling enables the explicit and unambiguous coupling of building functions with other pivotal aspects of
building operation (Figure 1), in a method that specifically considers the education and technical expertise of
building managers, for example, to evaluate the impacts of operational strategy on a building’s carbon dioxide (CO₂)

1 Data models such as Industry Foundation Classes (IFC) or SimModel offer potential solutions in this area [7,8].
emissions. The technique removes the need for subjective interpretation of limited data or expert intervention with respect to holistic performance analysis. Building managers can minimize the time spent aggregating performance data and instead focus on optimization strategies. The scenario modeling method uses an easily navigable, holistic, and reproducible checking mechanism that compares actual performance with predicted performance and completes the “plan-do-check-act” cycle for building managers. The underpinning logic captures, transforms, and communicates the complex interdependencies of environmental and energy management in buildings but presents this information in a format appropriate for building managers [21].

Using this method, building managers have more reliable information that can be communicated to other stakeholders at the tactical and strategic levels of organizations [22], enabling more informed energy-related decisions by upper management, who require a return on investment for any new method or technology.

Figure 1: Holistic building performance analysis, as evaluated by building managers, relies on an understanding of five key performance aspects and their interdependencies. These performance aspects form the basis of scenario modeling.

A building-wide implementation of scenario modeling [23] requires a software framework that is capable of representing the explicit class structure of the scenario modeling technique, illustrated in Figure 2a and accompanied by an example Scenario Model implementation in Figure 2b, and transforming the underpinning data through defined algorithms, to derive meaningful operational information. Available data contained in other organizational silos would further enhance the scope of scenario models. Linking building information silos has long proved to be a complicated exercise, and a mechanism that can effectively link these data sources is required. We explore the linked data concept as a solution to this problem.
There is a strong motivation for cross-domain data sharing of building data, and the benefits of performance data, presented at a granular level, have long been recognized (Section 2). We now offer a technical implementation that leverages the principles of linked data and complex event processing, to provide a decision support system for building managers. A demonstration of the proposed implementation focuses on an owner-occupied office building in Ireland that has limited measurement points (Section 2.1, 2.2).

2 Cross-Domain Scenario Modeling

Scenario modeling presents customized information for end users, but like all data processing techniques, this method relies on the availability of relevant data. In practice the vast majority of performance analysis is based on measured data, but benchmarks may also be derived from predictive models. Data other than that found in Building Management Systems (BMS) and utility information is seldom, if ever, used [24]. Sources such as whole building energy simulation models generate enormous volumes of data at differing levels of granularity, but post processing is lacking, especially the comparison of simulation outputs with measured data in buildings [3,25,26]. Utility pricing information often does not leave the financial domain of an organization, although access to such information may lead to an adjustment in operating strategy by building managers.

One of the key restrictions to the use of such data is that many of these external data sources use heterogeneous data definitions that may require technical expertise and significant time resources to parse into a performance framework. For example, comma separated value (.csv) files are often used to store common performance information such as time-series data, and the format of these data can take many different forms. Other data, such as weather or financial data, may also be retained in formats that are not immediately accessible, requiring manual intervention on the part of the building manager. Data may also be retained in functions of the organization, and a cross-domain data sharing framework may not exist.
Within organizations, many untapped sources of data exist that could be of enormous value to building managers, including occupancy counts, human resource data, room scheduling data, cleaning crew information, and design data that have not been updated over the building life cycle. For example, actual occupancy values may vary from those intended at design. An example from the National University of Ireland, Galway (NUIG) campus highlights the issue at hand (Figure 3). The air handling unit for a large lecture theater schedules the system to be available from 08:30 to 11:00 and 15:00 to 16:00 from Monday to Friday and assumes that the zone is not occupied when the system is off. Based on this information alone, a building manager performing a rudimentary trend analysis of zone conditions would not be concerned by a slight rise in temperature during unoccupied periods, but an expert may question these data. The CO₂ readings clearly indicate activity in the room. In the absence of the HVAC system, the CO₂ levels rise to almost 2,000 parts per million (ppm), almost twice the traditionally held figure of 1,000–1,200 ppm taken to indicate a deterioration in air quality [27]. In this case, a separate room scheduling system, maintained by the admissions office, contains a different occupancy schedule for the zone in question, and this clearly indicates
that the room is heavily used when the HVAC system is turned off (Figure 4). As in this case, the majority of building managers in commercial buildings do not have access to data other than that sent from the BMS.

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-09:00</td>
<td>237</td>
<td>237</td>
<td>237</td>
<td>200</td>
<td>237</td>
</tr>
<tr>
<td>09:00-10:00</td>
<td>237</td>
<td>237</td>
<td>237</td>
<td>200</td>
<td>237</td>
</tr>
<tr>
<td>10:00-11:00</td>
<td>237</td>
<td>180</td>
<td>180</td>
<td>145</td>
<td>237</td>
</tr>
<tr>
<td>11:00-12:00</td>
<td>237</td>
<td>200</td>
<td>237</td>
<td>200</td>
<td>149</td>
</tr>
<tr>
<td>12:00-13:00</td>
<td>237</td>
<td>145</td>
<td>145</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td>13:00-14:00</td>
<td>221</td>
<td>237</td>
<td>145</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td>14:00-15:00</td>
<td>221</td>
<td>120</td>
<td>160</td>
<td>140</td>
<td>140</td>
</tr>
<tr>
<td>15:00-16:00</td>
<td>149</td>
<td>250</td>
<td>160</td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td>16:00-17:00</td>
<td>200</td>
<td>160</td>
<td>160</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17:00-18:00</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Projected Occupancy Pattern for Lecture Hall 01 from the room scheduling system. This pattern does not match the supervisory operation schedule in the Building Management System.
In the context of organizational integration, a link between a room scheduling database and the operation of the HVAC system would provide significant organizational benefit. This example highlights the potential for interlinking different silos of information within an organization and creating an interpretive, richer scenario model for building managers and other organizational stakeholders. A robust, flexible solution must therefore consider a number of technical requirements that include: acquisition of large volumes of data stored in heterogeneous formats, processing of these data, and making all of the data available for pertinent stakeholders.

### 2.1 Use Case for Cross-Domain-Based Scenario Modeling

Scenario modeling using linked data and complex event processing ensures that data can be queried in a variety of previously unavailable ways. By doing so, building managers can leverage significant benefits in terms of operational understanding and building controls. The efficacy of the proposed approach depends on the range of data sources available in a particular building. We now demonstrate the concept as applied to the Digital Enterprise Research Institute (DERI) building, a utilitarian office block with stand-alone heating and ventilation units servicing the office spaces. The facility has a limited number of metered data streams. It serves exclusively as an office building, with a mix of open plan spaces, offices, and meeting rooms spread over three floors around a central core. The occupant profile suggests a 9 to 5 pattern of occupancy for the most part, with occupants based at personal computers throughout the day.

This case study describes the implementation of scenario modeling on the DERI building, located in Galway, Ireland, and reflects a potential use case for many organizations: the financial controllers at DERI are concerned by the rise in energy costs and are considering a pro rata billing system, based on the volume of energy consumed by each research group. Before such a system is implemented, the financial controllers wish to understand the current breakdown of energy consumption and assign the task to the building manager. There are 14 distinct sections in the building, but members of any group can sit in more than one section. Not all sections are equal in terms of energy consumption, so proportioning bulk energy is inappropriate.

Traditionally, in order to complete such a task, the building manager would track down the utility bills for a 12-month period and divide total figures by the number of groups within the organization. However this approach would not proportion the energy consumption fairly. To complete the given task, the building manager would have to locate the required information through file searches, phone calls, and meetings. He may have access to energy consumption values based on 14 predefined zones in the building (Table 1) but would not have space allocation details as defined in Table 2. Combining the information provided by these two tables could be a tedious process that includes data-parsing and data-processing activities. As the data required to carry out the analysis are spread across a number of domains, it becomes very difficult and time consuming to capture and interpret the data.

**Table 1: Breakdown of the available electrical meters at the DERI building, as available from the traditional Building Energy Performance (BEP) silo. Note one meter per zone and a total building meter.**

<table>
<thead>
<tr>
<th>PointName</th>
<th>Measured Object Type</th>
<th>Measured Object Name</th>
<th>PointType</th>
</tr>
</thead>
<tbody>
<tr>
<td>DERI_GF_SW</td>
<td>Space</td>
<td>Ground Floor South Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_GF_WW</td>
<td>Space</td>
<td>Ground Floor West Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_GF_NW</td>
<td>Space</td>
<td>Ground Floor North Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_FF_SW</td>
<td>Space</td>
<td>First Floor South Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_FF_WW</td>
<td>Space</td>
<td>First Floor West Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_FF_NW</td>
<td>Space</td>
<td>First Floor North Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>DERI_SF_SW</td>
<td>Space</td>
<td>Second Floor South Wing</td>
<td>Electricity Use</td>
</tr>
<tr>
<td>ID#</td>
<td>Floor</td>
<td>Zone</td>
<td>Name</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>308a</td>
<td>2</td>
<td>South</td>
<td>Researcher 1</td>
</tr>
<tr>
<td>310</td>
<td>2</td>
<td>West</td>
<td>Researcher 2</td>
</tr>
<tr>
<td>311</td>
<td>2</td>
<td>West</td>
<td>Researcher 3</td>
</tr>
<tr>
<td>313</td>
<td>2</td>
<td>North</td>
<td>Researcher 4</td>
</tr>
<tr>
<td>314</td>
<td>2</td>
<td>North</td>
<td>Researcher 5</td>
</tr>
</tbody>
</table>

Table 2: A subset of zone allocation information based on the human resources data silo.

In Section 2.2, we illustrate how linked data techniques might be used to interrogate diverse data sources, allowing data analysis to be performed easily.

2.2 Implementation of Scenario Modeling Using Linked Data and Complex Event Processing at DERI

DERI has implemented a linked data infrastructure which connects different silos of information through common contexts. This infrastructure has been extended to consider energy usage in the building. Figure 5 illustrates how the various data silos may be combined. For example, the zone context is common to three silos, and the person context is common to two silos (Figure 5). Different association relations are notable for the person context in each silo. With regard to the human resources silo, a person is an occupant of a group and a member of a group, while in the inventory silo, a person has a laptop. The “same as” relationship links the person context in both silos. This “same as” mechanism ensures that the end user has the ability to query data from all linked silos.
In this case, the end user is the building manager, and we propose that the building manager would use a single scenario model which represents total building energy consumption but focuses on an analysis of Research Group 1.
In doing so, it uses two performance aspects: energy consumption and legislation. This approach ensures that the energy consumption analysis also considers the legislative costs associated with CO$_2$ emissions. This particular scenario leverages whole building energy consumption as captured by electricity meters [28] and accounts for the zone-level electricity measurements depicted in Table 1. The scenario model explicitly captures the following information:

- Total building energy consumption
- Energy consumption for Research Group 1
- Cost of utilities for Research Group 1
- Associated CO$_2$ emissions for Research Group 1
The combination of the scenario modeling technique with different silos of information, connected through a linked data approach, dramatically assists the building manager with performance analysis activities. The system

<table>
<thead>
<tr>
<th>Performance Aspects</th>
<th>Building Objects</th>
<th>Performance Objectives</th>
<th>Performance Metrics</th>
<th>Formulae</th>
<th>Datum Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Consumption</td>
<td>DERI building</td>
<td>Maintain use within 5% of benchmark</td>
<td>Energy consumption</td>
<td>= Sum (Datum 1 + Datum 2)</td>
<td>Datum 1: Building Electricity (kWh)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Datum 2: Building Gas (kWh)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Utility consumption data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance Aspects</th>
<th>Building Objects</th>
<th>Performance Objectives</th>
<th>Performance Metrics</th>
<th>Formulae</th>
<th>Datum Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Consumption</td>
<td>DERI building</td>
<td>Determine use by research group</td>
<td>Energy use per research group</td>
<td>$\sum_{zone=1}^{n} \left( \frac{Elec_{PZ} \times N_{E}}{N_{E}} \right) + \frac{G_{R} \times G_{E}}{N_{TEG1}}$</td>
<td>Utility consumption data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance Aspects</th>
<th>Building Objects</th>
<th>Performance Objectives</th>
<th>Performance Metrics</th>
<th>Formulae</th>
<th>Datum Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Consumption</td>
<td>Research group 1</td>
<td>Maintain cost within 5% of benchmark</td>
<td>Research group energy cost</td>
<td>$\sum_{zone=1}^{n} \left( \frac{Elec_{PZ} \times N_{E}}{N_{E}} \right) + \frac{G_{R} \times G_{E}}{N_{TEG1}}$</td>
<td>Utility consumption data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required HR data streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Required emission factor data</td>
</tr>
</tbody>
</table>

**Nomenclature**

- **Elec**<sub>PZ</sub>: Electricity consumption per zone (kWh)
- **N<sub>E</sub>**: Number of employees in the zone (People)
- **N<sub>TEG1</sub>**: Number of Group 1 employees in the zone (People)
- **N<sub>TG1</sub>**: Total number of employees from Group 1 (People)
- **G<sub>T</sub>**: Total gas consumption for the building (People)
- **G<sub>R</sub>**: Gas unit rate (€/m<sup>3</sup>)
- **G<sub>EF</sub>**: Gas emissions factor (People)
- **E<sub>EF</sub>**: Electricity emissions factor (People)
- **E<sub>R</sub>**: Electricity unit rate (€/kWh)

**Figure 6**: Scenario model using linked data and complex event processing defines energy consumption per research group and CO<sub>2</sub> emissions per research group.
implemented at DERI enables real-time access to the required information and displays relevant information to the building manager.

clearly illustrates how the DERI energy consumption disaggregates by organizational group. To maintain consistency with the scenario definition, also includes a comparison of measured and predicted energy consumption, cost, and CO₂ emissions. In this case, the benchmark figures represent historically normalized data, as a whole building energy simulation model has not been developed for this building.

Figure 7: Dashboard image comparing total building energy consumption against a benchmark and as broken down by research group.
Figure 8 complements analysis at the building level by analyzing how members of Research Group 1, who sit at fixed locations, consume energy throughout the building—information that is difficult to determine using traditional processes and resources. The software implementation supports a breakdown of energy consumption for all research groups within the DERI organization. With this information at hand, the building manager can communicate transparently with upper management and fulfill his assigned duties.
This effective proof-of-concept demonstration must consider a number of technical challenges before it can be applied at a large scale.

### 2.3 Technical Challenges

To meet the high-level challenges discussed in Section 2.2, the proposed solution needs to meet the following technical requirements:

- **Multi-domain Information:** In large-scale environments, it becomes difficult to agree on the same concepts for the same things, due to federated responsibility of event and data publication. The system will need to support heterogeneous use of vocabulary to describe events and associated data. For example, an energy sensor may use the term “usage” while another one might use the term “consumption” to describe the same activity of consuming energy.

- **Inclusion of Real-time Sensor Data:** To support decisions with minimum delay, the system has to handle real-time observations that come from physical and virtual sensors, and to be able to process such streaming data efficiently in order to draw higher-level and aggregated information from it. Inclusion of external data sources, weather data, and utility price data may result in data quality issues, such as accuracy and uniformity, that need taken into account.

### 3 Implementation
3.1 Technology Background

3.1.1 Linked Data

Semantic Web technologies and standards play an important role in simplifying access to existing and future data and enabling large quantities of data to be shared on the Web. The Resource Description Framework (RDF) standard provides a common interoperable format and model for data linking and sharing on the Web. Linked Data is a best-practice approach used to expose, share, and connect data on the Web based upon W3C standards. In contrast to documents, Linked Data is not aimed at human consumption, it is processed and queried by computers, similar to relational data stored in conventional databases. Linked data technology uses web standards (e.g., the Resource Description Framework (RDF)[29]) in conjunction with the four following basic principles for exposing, sharing, and connecting data:

- **Use a standardized way to identify objects**: The use of Uniform Resource Identifier (URI) [30] (similar to a URL) to identify things such as a person, a place, a product, an organization, or an event—or even concepts such as risk exposure or net profit—simplifies data reuse and integration.

- **Use a standardized way to get data about objects**: URIs are used to retrieve data about objects using standard web protocols. For a person, this could be his or her organization and job classification; for an event, this may be its location, time, and attendance; for a product, this may be its specification, availability, price, or some other feature.

- **Use a standardized way to represent data**: When someone looks up a URI to retrieve data, provide it using standardized formats, ideally in Semantic Web standards RDF and SPARQL [31].

- **Use a standardized way to interlink information**: Retrieved data may link to other data sources, thus creating a data network (e.g., data about a product may link to all components it is made of, which may link to all suppliers).

Linked data technology can be accommodated with minimal disruption to existing information infrastructure, as a complementary technology for data sharing, and should not be seen as a replacement for current IT infrastructure (e.g., relational databases, data warehouses). The objective is to expose the data within existing systems, but only link the data when its needs to be shared.

3.1.2 Complex Event Processing

To process information flows, systems that are dedicated to high rate information flow processing are used. Data Stream Management Systems (DSMS) and Complex Event Processing (CEP) systems are two that have been adopted by commercial systems over the last few years [32]. In Complex Event Processing [33], the situations of interest are expressed in the form of event patterns and stored in the CEP engine. New information items can participate in the evaluation of the pattern if they are relevant. When a pattern is matched, a new higher-level event is generated and can participate in further processing, or could be forwarded to an event consumer like a dashboard or a business process management tool.

The bridge between Complex Event Processing and Linked Data is achieved with the use of Semantic Sensor networks (SSN). SSN were proposed [34] to optimize the benefits of current standards in the field of sensor networks [35]. The SSN ontology [34] has been designed to represent common concepts in a sensing environment. It handles metadata information such as sensor types and platforms, as well as classes, to describe real-time sensor observations. The use of sensor networks that respect linked data principles [36] when publishing data forms a solid basis for a real-time energy platform.

3.2 Linked Dataspaces for Energy Intelligence
This paper proposes that Linked Data and Complex Event Processing can be the basis for an interoperability approach that would help to overcome technical and conceptual barriers to cross-domain scenario modeling. The approach has been implemented using the Linked dataspace for Energy Intelligence developed at DERI. Figure 9 shows the proposed placement of the Linked dataspace for Energy Intelligence (LEI) [37,38]. LEI serves as a platform that applications can be built upon. The approach can support interoperability with linked data providing common syntactic and access protocols.

Figure 9: Linked dataspace for Energy Intelligence (LEI) enables cross silo data access for a diverse range of uses.

The main components of the architecture are:

- **Adapters**: At the bottom of the architecture are the existing operational legacy information systems. Adapters perform the “RDFization” process, which transforms multiple formats and legacy data and lifts it to the dataspace.

- **Linked Data**: The result of the linked data wrappers is a cloud of interlinked resources that reflect virtual or actual entities with links to relevant knowledge and contextual information from across all the information systems that have exposed linked data. Each entity within the cloud has a dereferenceable URI that returns data in a machine-readable format, describing the resource identified.

- **Support Services**: Dataspace support services are designed to simplify the consumption of the linked data cloud by encapsulating common services for reuse. Some example support services used in this work are:
 Entity Management Service, to improve data quality and inter-linkage between entity data scattered among legacy systems. The EMS can leverage automatic data integration algorithms that are supported by humans for collaborative data management [39].

A Complex Event Processing engine [40], to assess situations of interest that are encoded as event\action rules. Real-time information from sensors networks are also supported via the Semantic Sensor Network Ontology.

Data Catalogue and Provenance service [41], to query the catalogue about data sources with specific attributes such as freshness and publisher, and track the data back to their origin.

Search and Query services [42], to allow users to interact with the dataspace using structured or natural language interfaces.

Applications: At the top of the architecture are the consuming applications such as decision support applications and dashboards. Scenario models are defined at this level to gain a deeper understanding of the energy consumption behaviors in the building.

3.2.1 CEP Rules for Cross-Domain Scenario Modeling

The following code snippet implements the key concepts of the zone energy-sharing scenario model detailed in Figure 6. Together, the snippets represent a sample of the engineering implementation for the scenario model defined in Section 2.2. The first snippet details the event that is created when a new energy meter reading is received.

Implementation of “Zone Energy Sharing” Event for processing energy meter values

```
INSERT INTO ZoneEnergySharingEvent
SELECT
  EnergyConsumtionEvent.ZoneURI,
  GetZoneEnergySharing(EnergyConsumtionEvent.ZoneURI,
                      EnergyConsumtionEvent.Consumption)
FROM
  EnergyConsumtionEvent
```

The second snippet details the implementation of the energy allocation formulae defined in Figure 6 that apportions the energy usage for a zone to the groups that occupy that zone.

Function to apportion the energy consumption of a zone to individual groups

```
HashMap<URI, Double> GetZoneEnergySharing(ZoneURI, Consumption) {
    //For each GroupURI that has presence in ZoneURI
    //Add to the result map a pair of GroupURI and
    //Consumption*(proportion of GroupURI presence in ZoneURI)
    //Return the result map

    HashMap<URI, Zone> Zones=State.getZones();
    Zone zone=Zones.get(ZoneURI);
    HashMap<URI, Double> resultMap=new HashMap<URI, Double>();

    foreach(URI groupURI : zone.getOccupantsCount().getKeySet()){
        if(zone.getOccupantsCount().get(groupURI)>0) {
            Double groupConsumption =
```
4 Conclusion and Future Work

This paper illustrated the value of transferring data between different parts of organizations for use in holistic environmental and energy management activities. The combination of scenario modeling, linked data, and complex event processing significantly enhances the information available to building managers, especially when the information communicated is in a format specifically designed for building managers. With access to additional information that can in turn be processed by predefined algorithms, this solution removes the need for subjective or expert intervention with respect to holistic performance analysis. The technique minimizes the time spent by building managers on performance analysis and building performance optimization. In fact, building managers, empowered by additional cost and operational information, can communicate savings to other organizational stakeholders with greater certainty, and this elevates the role of building manager in the context of organizational objectives, especially if organizations practice continuous improvement as defined by the best practice “plan-do-check-act” cycle, the key component of ISO 50001 [43].

Future building managers should have access to all possible data that they require from an organization. It is imperative that these data are processed and communicated in a format suitable for building managers. Such data and processing include simulation-based modeling and control, expert algorithms for Fault Detection and Diagnosis (FDD), or functionality to leverage other types of simulation, such as Daylighting or Computational Fluid Dynamics (CFD). In order to include new data silos in the interoperable linked data environment, unique adapters are required for each silo. Bidirectional benefit is a key driver for the development of such adapters. For example, the HR domain could leverage occupancy data from the demonstration in Section 2.2.

Once established, Linked Data systems are easily scalable for campuses and portfolios of buildings, but should not be limited to assisting only the building manager. Further work could focus on stakeholder integration, where other stakeholders, particularly designers, commissioners, and financial controllers would benefit significantly from cross-silo data access. In conclusion, the scenario modeling technique is not limited to the energy domain, but is an extensible method of evaluating building related data. As more data streams become available in different building domains, the technique can be extended to incorporate these. The linked data approach is built on the premise of connecting diverse data streams. As these streams become more widespread in the building area, the possibilities for building wide and urban scale data platforms driving myriad applications becomes possible.

5 Acknowledgements

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technology, Building Technologies Program of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

This work has been funded by the Irish Research Council.

This work has been funded by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion-2).

6 References


