



## Linking building data in the cloud: Integrating cross-domain building data using linked data

Edward Curry<sup>a,\*</sup>, James O'Donnell<sup>b</sup>, Edward Corry<sup>c</sup>, Souleiman Hasan<sup>a</sup>, Marcus Keane<sup>c</sup>, Seán O'Riain<sup>a</sup>

<sup>a</sup>Digital Enterprise Research Institute, National University of Ireland, Galway, IDA Business Park, Lower Dangan, Galway, Ireland

<sup>b</sup>Simulation Research Group, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 90R3111, Berkeley, CA 94720, United States

<sup>c</sup>Informatics Research Unit for Sustainable Engineering, National University of Ireland, Galway, Ireland

### ARTICLE INFO

#### Article history:

Available online 16 November 2012

#### Keywords:

Linked data  
Data interoperability  
Building energy analysis  
Building management  
Building information model  
Data as a service

### ABSTRACT

Within the operational phase buildings are now producing more data than ever before, from energy usage, utility information, occupancy patterns, weather data, etc. In order to manage a building holistically it is important to use knowledge from across these information sources. However, many barriers exist to their interoperability and there is little interaction between these islands of information.

As part of moving building data to the cloud there is a critical need to reflect on the design of cloud-based data services and how they are designed from an interoperability perspective. If new cloud data services are designed in the same manner as traditional building management systems they will suffer from the data interoperability problems.

Linked data technology leverages the existing open protocols and W3C standards of the Web architecture for sharing structured data on the web. In this paper we propose the use of linked data as an enabling technology for cloud-based building data services. The objective of linking building data in the cloud is to create an integrated well-connected graph of relevant information for managing a building. This paper describes the fundamentals of the approach and demonstrates the concept within a Small Medium sized Enterprise (SME) with an owner-occupied office building.

© 2012 Elsevier Ltd. All rights reserved.

## 1. Introduction

The holistic management and maintenance of facilities is a multi-domain problem encompassing financial accounting, building maintenance, facility management, human resources, asset management and code compliance, affecting different stakeholders in different ways. Within the operational phase buildings are now producing more data than ever before, from energy usage, utility information, occupancy patterns, weather data, scheduling software, financial control, etc. There is usually little interaction between the islands of information in traditional Building Management Systems (BMS) and the broader domain of building data, making it difficult to assemble cross-domain information. The volume of information can quickly become overwhelming without a clear and focused methodology for information management. Within the wider-context of the organization other relevant information sources must also be utilized to optimize both the building and the operation of the organization itself. This information, which may also exist in silos, includes payroll, human

resources, production systems, ordering systems, resource-planning systems etc. In order to manage a building holistically it is important to use knowledge from across these sources. However, many barriers exist to their interoperability.

Cloud computing involves the delivery of computing services to end-users with the promise of faster provisioning, dynamic allocation, improved manageability, and reduced maintenance. Typically, cloud computing is delivered in a service model, such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). At the foundation of cloud computing is the broader concept of converged infrastructure, shared services, and pooling resources for economies of scale. An emerging trend is to move data to the cloud. Data as a Service (DaaS) is based on the concept that data can be provided on demand to the user regardless of geographic or organizational separation of the data provider and consumer. Moving information systems, such as building management systems, to the cloud may have many benefits in terms of cost and service provision. However, new cloud-based data services designed in the same manner as traditional building management systems will still suffer from data interoperability problems. We need to reconsider how we approach interoperability and embrace the concept of resource sharing within the design of new cloud-based data services.

\* Corresponding author. Tel.: +353 91 492973.

E-mail addresses: [ed.curry@deri.org](mailto:ed.curry@deri.org) (E. Curry), [JTODonnell@lbl.gov](mailto:JTODonnell@lbl.gov) (J. O'Donnell), [edwardcorry@nuigalway.ie](mailto:edwardcorry@nuigalway.ie) (E. Corry), [souleiman.hasan@deri.org](mailto:souleiman.hasan@deri.org) (S. Hasan), [marcus.keane@nuigalway.ie](mailto:marcus.keane@nuigalway.ie) (M. Keane), [sean.oriain@deri.org](mailto:sean.oriain@deri.org) (S. O'Riain).

Emerging from research into the Semantic Web, linked data leverages the existing open protocols and W3C standards of the Web architecture for sharing structured data on the web. Linked data proposes an approach for information interoperability based on creation of a global information space; also known as the Linked Data Cloud [1]. Because linked data is web-centric, it is an ideal technology for cloud-based data services. Linked data provides a mechanism through which information silos can exist in a homogeneous format. Representing building data, such as a Building Information Model (BIM), using the linked data format, will allow it to be combined with linked data from other relevant cross domain silos. In doing so, organizations can generate and extract additional value from current stand-alone repositories, across multiple domains. The resulting merged cross-domain linked data cloud provides a holistic view of the building's operations, which can have added value for domain stakeholders throughout the organization. This paper describes the fundamentals of the approach and demonstrates the concept with an owner-occupied office building.

Section 2 motivates the need for cross-domain information sharing of relevant building data, while highlighting the benefits and challenges. Linked data principles for publishing data on the web are introduced in section 3 along with a discussion on its interoperability benefits. Section 4 provides a detailed technical description of the linked data approach to interoperability and how it can be applied to linking building data in the cloud. Section 5 describes a case study of the proposed linked building data approach within an office building, with a specific focus on energy management. Conclusions and future work are discussed in section 6.

## 2. The case for cross-domain information sharing

Most building costs are experienced during the operation phase of the Building Lifecycle (BLC). Financially speaking a “life cost ratio” is described in words of “initial capital cost”, “operational cost” and “salary cost”. On average, a ratio of 1:10:200 is applicable to commercial buildings [2]. In other terms, operational and salary costs are 10 and 200 times greater than the capital cost.

Buildings are generating increasing volumes of data during their operational phase that can be leveraged to reduce the operational costs. However, the poor provision of operational information is leading to inefficient building performance. Significant costs [3] are incurred recreating, verifying and transforming building information, with studies [4] quantifying the costs associated with this lack of interoperability at \$15.8 billion in the U.S. capital facilities supply chain for 2002 alone. The research also showed that building owners and operators have to absorb two thirds of these costs.

### 2.1. Improving operational strategy

The holistic management and maintenance of facilities is a multi-domain problem with many stakeholders. The type of information needed by each stakeholder is different and varies in the level of detail required. The skill-set of each stakeholder is also markedly different and each has very different motivations for accessing information on the building.

Navigating through this mass of data in a coherent manner to derive information and tailor output to specified end users is a challenge. Different stakeholders need different views of the information [5]. For example, the financial controller is concerned with cost metrics, the human resource function is concerned with occupant comfort, whilst the owner is motivated by the overall efficiency of the building. There are a variety of measurement methodologies that can be utilized to quantify each of these, but

when considered independently, it is difficult to get a complete picture of the building.

There is a clear need to define a cross-domain operational strategy in a comprehensive and structured manner. A well-designed strategy will provide a roadmap through the vast array of cross-domain data generated during the operational phase, providing the relevant information view to each particular stakeholder.

### 2.2. Improving performance evaluation

There is a wide variety of standards and methodologies for the specification and measurement of building performance and most require far greater levels of interaction with data than currently takes place. There is a lack of quantitative feedback from buildings preventing the detection of problems [6]. Building optimization processes tend to use models that compare current performance to past performance, such as continuous commissioning [7] and the BuildingEQ project [8]. Building codes tend to assess building performance against normative tables based on statistical analysis of other buildings. Studies show that buildings do not operate as designed, even buildings that are actually designed to operate to a particular standard, such as LEED buildings [9].

Cross-domain performance evaluations are important to understand the performance and optimize the operation of the building to ensure it is meeting the needs of the organization/occupants. The scenario modeling concept [10] calls for a variety of information views across the building performance domains with a range of performance aspects, providing a comprehensive performance framework for the analysis of building operation. Scenario modeling proposes a methodology where building performance is assessed through measurable metrics, describing aspects of building operation. The creation of consistent, continuous and unambiguous building performance information is seen as a key enabler for building optimization. The inclusions of specific performance objectives from other domains, such as human resources (i.e. occupancy comfort) and enterprise resource planning (i.e. room utilization), within the scenario model would lead to a greater clarity for building managers on the performance of the facility in meeting organizational objectives [11].

### 2.3. The building information model

Building Information Modeling (BIM) is a new departure in the area of information management within the Architecture, Engineering and Construction (AEC) industry, allowing the accurate capture and maintenance of building information. A building information model is a fully object-oriented, integrated, parametric model of the building, and it can theoretically provide a full view of the entire data associated with the building, rather than the narrow subset of information associated with 2-D drawings. BIM allows for the definition of objects, with defined relationships and associated meta-data. BIM provides an effective framework for the capture of this information in a usable electronic format, aiding interoperability between domains in the building design, construction and management phases.

BIM can be seen as a central repository of building data, for use by all project stakeholders, across the project lifecycle. It is a relatively new concept in the domain and offer a mechanism through which a complete set of project data can exist in one easily accessible format for building managers [12]. However, within the wider-context of the organization the BIM is only one silo of information, other relevant information must also be utilized to optimize both the building and organization itself. This information, which may also exist in silos, includes payroll, human resources, production systems, ordering systems, resource-planning systems etc.

While BIM has shown undoubted potential in the design and construction phases of the BLC, the utilization of BIM in the operational phase is currently limited to the use of BIM as a static repository of information concerning building entities. However, significant potential benefits have been identified in projects where building information [13] has been utilized within the operational phase to provide an as-built BIM at the handover stage. Potentially, BIM could serve as an informational backbone for Computerized Maintenance Management Systems (CMMS), Computer Aided Facility Management (CAFM), and Integrated Workplace Management Systems (IWMS) used by facility managers [14].

It is logical to assume that far greater use can be made of a BIM, in particular making use of its relationship properties to provide a strong informational backbone to data analysis efforts throughout the operational phase. By integrating the BIM with standardized operational strategies, an operational BIM could serve as a domain silo in the management of diverse building data streams, emanating from the myriad of cross domain sources in the modern building, including weather, BMS, utility billing, occupancy and HR.

#### 2.4. Interoperability challenges

There are few if any methodologies available that are aimed at providing an analysis of building operation to all interested stakeholders. A subset of data contained within each domain would represent a subset of data required by a stakeholder in a different domain. Due to the critical lack of information interoperability, it is quite difficult to get a complete cross-domain view of a building in terms of interaction of data streams in a clear and structured manner.

The Industry Foundation Classes (IFC) standard promotes interoperability within the building and construction domain, and for BIMs in particular [15]. However, IFC by itself is not sufficient to enable interoperability with systems outside of the AEC domain. Not all building related information is currently described in the BIM process. Efforts to describe the exchange of data for a range of sub domains in the AEC industry are ongoing [8]. The Model View Definitions (MVD) process is being used to extend the transfer of information in a number of domains. The question of interoperability between the various domains in the AEC/O industries is one that has led the BuildingSMART Alliance to introduce a standardized information model in IFC to conform to international standard ISO/PAS 16739.

Moving these data sources to the cloud will not solve the interoperability problem by itself. The fundamental barriers to interoperability will remain. A number of challenges need to be addressed when integrating data from different sources, and a detailed analysis of integration challenges is provided by [16]. We classify the challenges into four groups: *object identifiers and separate schema, data mismatch, abstraction levels mismatch, data quality*.

- *Object identity and separate schema*: Structured data is available in multiple formats such as Comma-Separated Values (CSV), Extensible Markup Language (XML) and the record sets returned from relational databases. To integrate structured data it is necessary to first lift (map) that data to a common format. Having equivalent formats however does not guarantee consistency as the originating sources may state what is essentially the same fact differently. These differences exist at both the data description (schema) and actual data (individual object) levels. Consolidating information from different sources therefore requires methods for identifying and reconciling different representations at both the schema and object levels.
- *Data mismatch*: When creating an integrated, holistic view over multiple data sources a key challenge is the mismatch between sources. Data is often ambiguous – the same company might be

referred to using several variations (e.g. DERI, Digital Enterprise Research Institute). This ambiguity makes cross-linking structured data difficult. In addition, data expressed in human language is difficult to process via software programs.

- *Abstraction levels*: Data sources can describe data at incompatible levels of abstraction or classify their data according to taxonomies pertinent to a certain sector. For example, the IFC provides a taxonomic classification that may be incompatible with other common classification schemes. Since data may be modeled at different levels of abstraction (e.g. person, building, site, or country), data aggregated for an individual viewpoint may not match that from other viewpoints.
- *Data quality*: Data quality is a general challenge when integrating data. Data is often erroneous, and combining data can aggravate the problem. A low quality data source can have a devastating impact on the overall quality of a merged dataset.

### 3. Cross-domain data sharing with linked data

The web and cloud computing are intimately related technologies. Cloud computing services are typically accessed via the web, making it a critical enabler of cloud computing. However, the web has more to offer cloud computing than just a web-based user interface, it can play an important role to share and interoperate data. Web standards can simplify access to data and enable the sharing of large quantities of data 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 for exposing, sharing, and connecting data on the Web based upon W3C standards. Linked data has the following characteristics:

- *Open*: Linked data is accessible through an unlimited variety of applications because it is expressed in open, non-proprietary formats.
- *Modular*: Linked data can be combined (mashed-up) with any other pieces of linked data. No advance planning is required to integrate these data sources as long as they both use linked data standards.
- *Scalable*: It's easy to add and connect more linked data to existing linked data, even when the terms and definitions that are used change over time.

We propose that RDF and linked data provide an appropriate technology platform to enable the sharing of cross-domain information relevant to the operation of a building. We propose that as we move data to the cloud, linked data technology offers a viable medium for the sharing and reuse of data across silos. Whether it's the integration of multiple energy management systems, billing systems, building management systems, or spreadsheets, linked data offers a method of exposing, sharing, and connecting data in a manner that is reusable and not a one-off integration solution. The key concept, as depicted in Fig. 1, requires that all data relevant to the operation of the building be expressed in a common format (RDF), making it easily accessible by end users. In doing so building data such as BIMs are reusable outside the AEC/Facilities Management domain and can provide greater value across organizations.

#### 3.1. The linked open data cloud

The Web is moving from a medium for sharing documents to a medium that can also be used to share data. Fuelled by the *Open Data* initiative, the emerging "Web of Data" means easier access to data for consumers. Typically, open data is non-textual material such as maps, genomes, chemical compounds, mathematical and

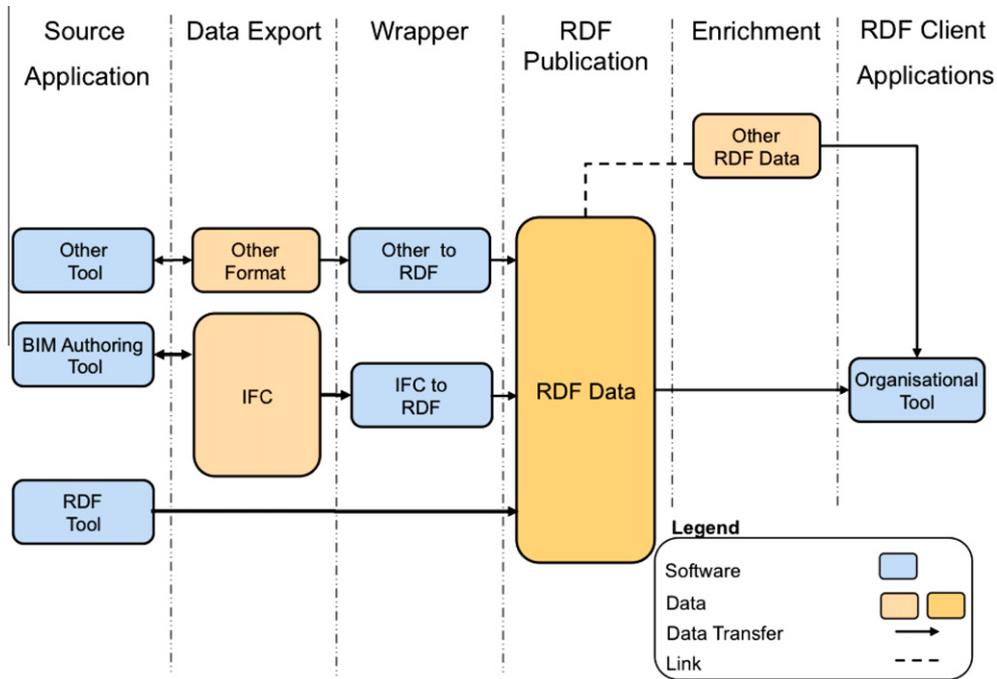


Fig. 1. An environm capable of leveraging IFC based BIM data for use with other applications through RDF data format.

scientific formulae, medical data and practice. Open data can also include generalized business news, product information and financial data [17] available from an assortment of sources. Demands for greater levels of transparency have resulted in Open Government

initiatives that have made available large numbers of sector, statistics, financial, and economic datasets for public consumption.

The Linked Open Data Cloud, see Fig. 2, represents a large number of interlinked RDF datasets that are being actively used by

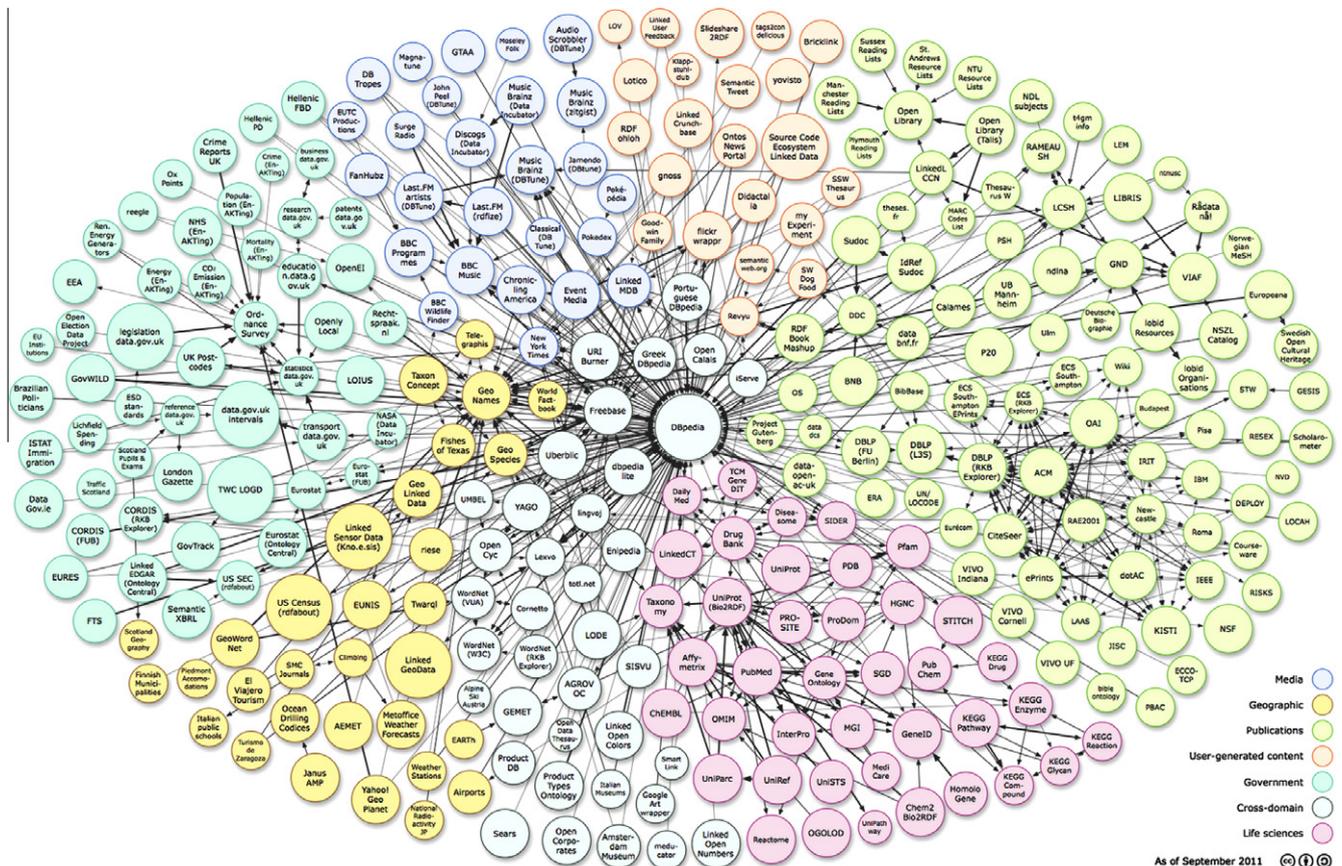


Fig. 2. The linked open data cloud [19].

industry, government and scientific communities [18]. The linked data cloud has been growing considerably in the past years and provides a foundation upon which applications leveraging that data can be built on. As of September 2011, it consists of more than 300 open data sets with more than 35 billion facts, interlinked by 500 million typed links. The cloud is doubling in size every 10 months.

### 3.2. Linked data principles

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. RDF) in conjunction with four basic principles for exposing, sharing and connecting data. These principles are:

1. *Use URIs as names for things*: the use of Uniform Resource Identifiers (URI) (similar to URLs) to identify things such as a person, a place, a product, a organization, a building, a room, an event or even concepts such as risk exposure or net profit, simplifies reuse and integration of data.
2. *Use HTTP URIs so that people can look up those names*: URIs are used to retrieve data about objects using standard web protocols. For a person this could be their 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, etc.
3. *When someone looks up a URI, provide useful information using the standards*: when someone looks up (dereferences) a URI to retrieve data, they are provided with information using a standardized format. Ideally in Semantic Web standards such as RDF.
4. *Including links to other relevant URIs so that people can discover more things*: retrieved data may link to other data sources, thus creating a data network e.g. data about a product may link to all the components it is made of, which may link to their supplier.

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 (relational databases, data warehouses, etc.). The objective is to expose the data within existing systems, but only link the data when it needs to be shared.

### 3.3. Benefits for information interoperability

Information integration projects typically focus on one-off point-to-point integration solutions between two or more systems in a customized but inflexible and ultimately non-reusable manner. This limits both the information flow and its oversight between systems to those that have been integrated. The fundamental concept of linked data is that data is created with the mindset that it will be shared and reused by others. Linked data is created in a fashion that allows it to be easily linked to at the information-level (data) not the infrastructure-level (system) by focusing more on the conceptual similarities (shared understanding) between the information. Users are also more comfortable working with data expressed at the information level.

- It allows separate systems designed independently to be later joined/linked at the edges.
- It allows interoperability to be added incrementally when needed and where cost-effective.
- It allows any data to be expressed in a mixture of vocabularies.

Linked data technology can transform the way data is shared and consumed, by freeing data to be quickly and easily linked as reusable data sets. It can also improve the availability and timeliness of critical data to facilitate better decision-making.

## 4. Linking building data in the cloud

As part of moving building data to the cloud there is a critical opportunity to reflect on the design of cloud-based data services and how they are designed from an interoperability perspective. In this paper we propose the use of linked data as a flexible mechanism for data publishing and integration. The main components of the approach are (1) Universal Resource Identifiers (URIs) to name things, (2) RDF for representing data, (3) Linked data principles, and (4) common vocabularies to establish and share understanding. Each of these components is now discussed using examples from the building domain.

### 4.1. Universal resource identifiers

In order to avoid situations where multiple data sources create identical identifiers for denoting different things, coordination between participants is required to name things unambiguously. The web with its established infrastructure is used to help create this naming scheme. URIs and more recently Internationalized Resource Identifiers (IRIs) are used to define identifiers for resource names. A resource can be a file (such as a *hypertext document*, a *digital image*, or a *digital video*), a real-world entity (such as a *person*, a *company*, or a *device*), or an abstract concept (such as *energy conservation* or relationships such as *consumed by*). URIs provides a basic mechanism to globally identify resources and form the basic mechanism used to associate disjoint pieces of data. URIs can be retrieved (dereferenced) via HTTP (Web browser or application). Example URIs for identifying entities are provided in Table 1.

### 4.2. Resource description framework

The Resource Description Framework (RDF) [20] is the basic machine-readable representational format used on the Semantic Web to represent information. RDF is a general method for encoding graph-based data that does not follow a predictable structure. RDF is schema-less and self-describing, meaning that the labels of the graph describe the data itself. Data and facts are specified as statements and are expressed as atomic constructs of a subject, predicate, and object, also known as a triple. The statement “*Edward Curry is the Occupant of Room 202e*” is expressed in triple format as:

**8. Subject** – “*Edward Curry*”

**9. Predicate** – “*is the Occupant of*”

**10. Object** – “*Room 202e*”

**Table 1**  
Example universal resource identifiers for entities.

Entity	URI
Galway	<a href="http://www.data.nytimes.com/galway_ireland_geo">http://www.data.nytimes.com/galway_ireland_geo</a> <a href="http://www.dbpedia.org/resource/Galway">http://www.dbpedia.org/resource/Galway</a> <a href="http://www.rdf.freebase.com/ns/en.galway">http://www.rdf.freebase.com/ns/en.galway</a>
Barack Obama	<a href="http://www.dbpedia.org/resource/Barack_Obama">http://www.dbpedia.org/resource/Barack_Obama</a> <a href="http://www.data.nytimes.com/obama_barack_per">http://www.data.nytimes.com/obama_barack_per</a>
DERI Building	<a href="http://www.lab.linkeddata.deri.ie/2010/deri-rooms#building">http://www.lab.linkeddata.deri.ie/2010/deri-rooms#building</a>
Room 202e, DERI Building	<a href="http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e">http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e</a>
Edward Curry	<a href="http://www.deri.ie/about/team/member/edward_curry#me">http://www.deri.ie/about/team/member/edward_curry#me</a>

RDF has been designed for use in decentralized web-scale graph data models. For this reason the statement parts need to be identified so that they can be easily reused. As noted in Table 1, <http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e> denotes “Room 202e” within the “DERI Building”, while the URI [http://www.deri.ie/about/team/member/edward\\_curry#me](http://www.deri.ie/about/team/member/edward_curry#me) denotes “Edward Curry”. RDF uses these URIs for identification, and expressing the previous triple statement in RDF becomes:

11. **Subject** – [http://www.deri.ie/about/team/member/edward\\_curry#me](http://www.deri.ie/about/team/member/edward_curry#me)
12. **Predicate** – “is the Occupant of”
13. **Object** – <http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e>

Relationships are also identified with a URI, the “is the Occupant of” relationship can be identified with the URI <http://www.vocab.deri.ie/rooms#occupant> which results in a final triple of:

[http://www.deri.ie/about/team/member/edward\\_curry#me](http://www.deri.ie/about/team/member/edward_curry#me)  
<http://www.vocab.deri.ie/rooms#occupant>  
<http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e>

Multiple triples can be joined together to build up a graph of information. Fig. 3 provides an example of a graph segment that models a description of a Room: its name, occupant, containing building, and height dimension.

Within the graph, subjects and objects are nodes, while a predicate is an arc. The DERI building is identified by “<http://www.lab.linkeddata.deri.ie/2010/deri-rooms#building>”, and we represent the height of the room within RDF using an IFC ontology [21].

#### 4.3. Publishing with principles

Any data provider, such as BMS, ERP, and BIM, wanting to make their data sharable in the cloud should publish data following the linked data principles:

1. Use **URIs** to identify building objects, i.e. <http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e> to identify “Room 202e” in the DERI building.
2. Use **HTTP URIs** so that building objects can be referred to and looked up (“**dereferenced**”) by people and **user agents**. The URI for Room 202e can be looked up in a browser or software.
3. Provide useful information about the building object when its URI is dereferenced, using standard formats such as **RDF/XML**. A lookup on the URI using an RDF-capable agent returns <http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e.rdf>, a RDF document containing data about “Room 202e”.
4. Include **links** to related URIs to improve discovery of relevant information on the cloud. The data returned contains links to other data sources, for example to

the occupant of the room [http://www.deri.ie/about/team/member/edward\\_curry#me](http://www.deri.ie/about/team/member/edward_curry#me), which can then be dereferenced to provide further information.

Data published adhering to these simple principles leads to a unified mechanism for data access and inter-linkage. Linked data provides a graph-based knowledge representation formalism, in conjunction with global resource identifiers, which facilitates data integration on an enterprise and even worldwide scale. The key point is that URIs uniquely identifies data and relationships within triples using RDF syntax. Since URIs are unique they can be used to integrate data by matching on the URIs, and because URIs are HTTP based the Web infrastructure can be leveraged to access and integrate data distributed across silos.

#### 4.4. Vocabularies

Establishing a shared understanding between data providers and consumers of what schema-level identifiers mean requires a method to describe schema-level constructs. Vocabulary descriptions provide a shared conceptual model using schema-level constructs of classes and properties that are annotated with meta-data, such as human-readable labels and descriptions. They can support simple data models to knowledge representation constructs rooted in logic that allow for mathematical specification of meaning and mechanical inference (such as subclass and sub-property relations). The main vocabularies used are RDF, RDF Schema (<http://www.w3.org/TR/rdf-schema/>) and the Web Ontology Language (<http://www.w3.org/2004/OWL/>).

Consider the partial representation of the data model describing a room in Fig. 3. Observe the abbreviated prefixes of *rm*, *foaf*, and *ifc*. Each denotes a particular domain vocabulary that provides distinct descriptions and definitions of the schema/namespaces that they represent. Namespaces allow for dereferencing class and property definition, so they can be re-used or extended by an other vocabulary. “*rm*” refers to the Building and Rooms schema; “*foaf*” to the Friend of a Friend vocabulary that describes people; and “*ifc*” to the Industry Foundation Class vocabulary [21]. Looking at the data model, *rm:Room* stipulates that the definition for *Room* is derived from the rooms (“*rm*”) namespace, <http://www.vocab.deri.ie/rooms#Room>. By dereferencing *rm:Room* a description can be retrieved, it will establish that *rm:Room* is a class, of which there can be instances such as “Room 202e”. This can be stated in RDF as:

<http://www.lab.linkeddata.deri.ie/2010/deri-540rooms#r202erdf:typem:Room>

While different sources may use different class URIs for the same class grouping, these can later be flexibly mapped at the schema level. The description can contain useful information such as: sub-classes, useful for reasoning; text descriptions, useful for explaining the purpose of the class to any user or consumer of the class (*rm:Room*) and labels (“Room”) that are used for results presentation. In addition to classes, properties can also be defined and dereferenced.

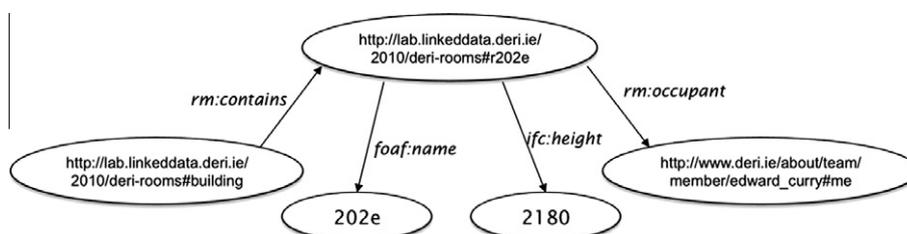


Fig. 3. An RDF graph describing Room 202e in the DERI building.

### 5. Case study: DERI building energy management

In order to support the position made in this paper a case study of a linked building data cloud has been developed for the Digital Enterprise Research Institute located in the IDA Business Park adjacent to the campus of the National University of Ireland, Galway. The objective of the case study is to demonstrate how building data and other related data can be integrated within a linked data cloud. Specifically the case study targeted the energy management of the building. The purpose of a building energy management system is to gather data related to energy consumption and to present it in an actionable manner so that minimal effort is needed to move from the presented knowledge to energy-saving decisions. The system was utilized by users from multiple domains thus it needed to have a simple user interface to support non-building domain experts. This section discusses the building, legacy information systems, system architecture, and the vocabularies in the use case. It also describes how the vocabularies were linked to break down the silos, and the energy management applications and services built using the resulting linked building data cloud.

#### 5.1. Building description

DERI is a premier research institute with approximately 130 research students and staff with a worldwide reputation in its area. It is based in a dedicated building with 2190 sq. m of space, comprising 22 unit offices, 160 open plan workspaces, 1 large 80-seat conference room with audio visual and video conferencing facilities, 4 meeting rooms, 3 kitchens, 1 air conditioned data centre with

backup generator, 1 sensor network laboratory, a 30 person café, and Ireland’s National Museum of Computing History.

As with a typical modern office building there are many sources of power consumption such as Heating, Ventilation and Air Conditioning (HVAC) systems, lights, and electronic devices. The building provides a first-class technical infrastructure (networks, presentation technology, and meeting places) to its researchers. In addition, the University Computer Services Group provides significant network support to DERI (e.g., a Gigabit HEAnet connection) and the technical administrative team within DERI provides the institute’s computer infrastructure technical support.

#### 5.2. Legacy information systems

The DERI building was built in the 1990s and does not employ a building management system or an energy management system. The DERI building has been retrofitted with energy sensors to monitor the consumption of power within the building. In total there are 31 fixed energy consumption sensors covering office space, café, data centre, kitchens, conference and meeting rooms, computing museum along with five mobile sensors for devices, light and heaters’ energy consumption as well as motion detection.

As typically in an organization DERI has a number of information systems that run its operations, including finance and enterprise resource planning. Some limited building-related data was available, including a post-construction BIM, and basic energy meter recording software. These enterprise systems can help in identifying energy wastage and non-ecological actions of the organizations.

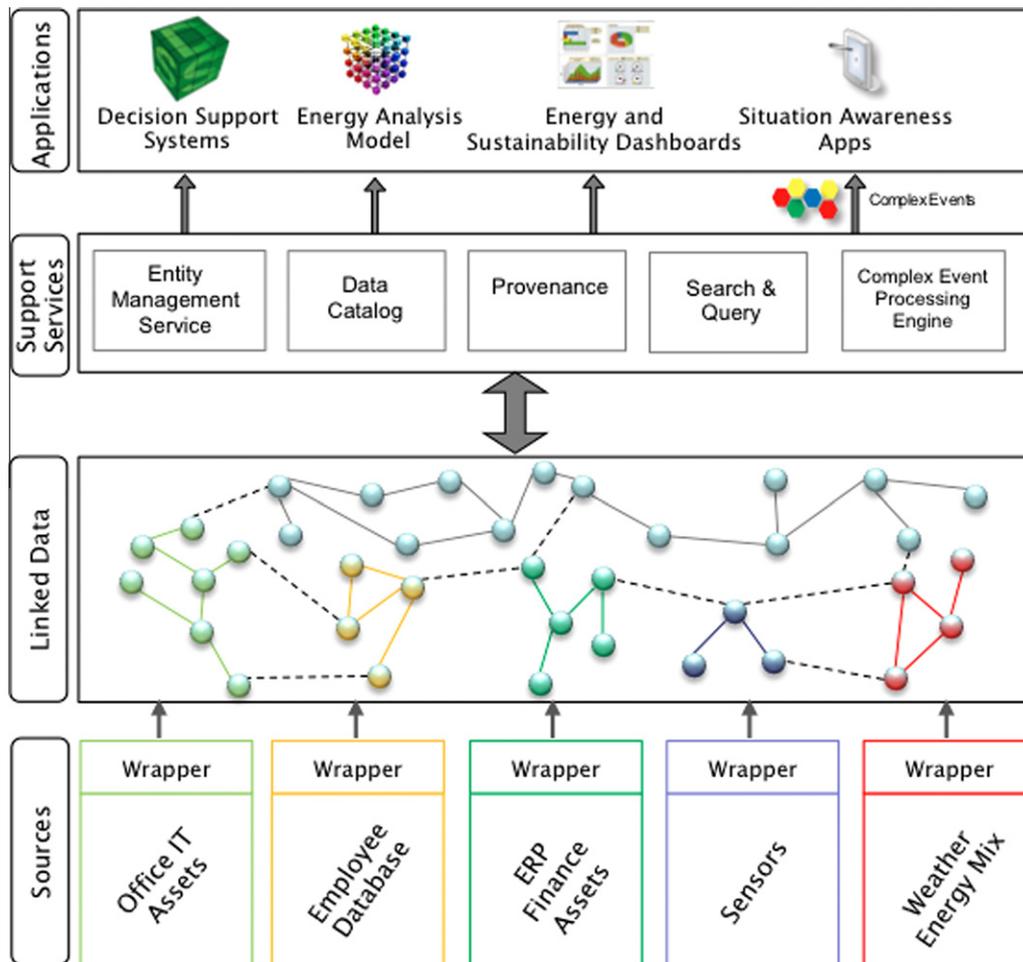


Fig. 4. Linked dataspace for Energy Intelligence (LEI) [24].

### 5.3. Architecture

The linked building data cloud for the DERI building use case has been implemented using the Linked dataspace for Energy Intelligence (LEI) developed by the Green and Sustainable IT group at DERI. LEI follows the dataspace approach to data management that recognizes in large-scale integration scenarios, involving thousands of data sources, it is difficult and expensive to obtain an upfront unifying schema across all sources [22]. Dataspace shifts the emphasis to providing support for the co-existence of heterogeneous data that does not require an upfront investment into a unifying schema.

The LEI platform has been applied in a number of energy management scenarios including cloud computing [23] and enterprise energy management [24]. In this section we show how LEI can be applied to the management of building energy. Fig. 4 shows the placement of LEI [25] [24] as an independent layer placed above the existing systems layer. The approach supports interoperability with linked data providing a common syntactic and access protocols. The main components of the LEI architecture are the wrapper on existing systems, the linked dataspace consisting of a linked data cloud & support services, and the resulting applications. The remainder of this section will now discuss each of these components in more detail.

#### 5.3.1. Sources

At the bottom of the architecture are the existing operational legacy information systems. Linked data wrappers perform the “RDFization” process, which transforms multiple formats and legacy data to linked data. Converter tools are available to automate the conversion process [26] according to the relevant RDFs and/or OWL vocabularies.

The DERI building use case comprises 14 existing operational legacy systems ranging from spreadsheet-based building resource planning to human resource systems and wireless sensor networks. Many of these sources of data are reused from a different

**Table 2**

Summary of legacy information systems.

Information system	Type	Information
Building information model	ArchiCAD BIM	Physical dimensions of building
Building resource allocation	Spreadsheet	Assignment of building resources to people and groups within DERI
Office IT assets database	SharePoint	Office IT equipment assignment
Employee database	SharePoint	Person detail, group affiliations and current work assignments
Energy sensor network	Enistic sensor platform	Real-time sensor reading from power meters within building

project on Sustainable Reporting [27], because they were made available as linked data they are easily reused within the linked building data cloud. A summary of the main data sources is provided in Table 2. In addition to the legacy data sources, 30+ real-time power sensors monitor the power consumption within the building. The sensor readings are adapted to RDF using the Jena framework [28], the Semantic Sensor Network Ontology [29], and enriched based on the linked building data cloud.

#### 5.3.2. Linked building data cloud

The result of linked data wrappers over data sources is a linked building data cloud rich with knowledge and semantics about building related entities. Each entity within the cloud has a dereferenceable URI that returns data in a machine-readable format describing the resource identified. This mechanism makes it possible for a user (or software agent) to “follow their nose” to find out more information related to the identified resources. This is illustrated in Fig. 5, where the URI for “Room 202e” is dereferenced to show the RDF returned. The linked building data cloud and the process used to link data from different systems within it are covered in detail in the next section.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rooms: <http://vocab.deri.ie/rooms#> .
@prefix foaf: <http://xmlns.com/foaf/0.1#> .
@prefix deri-energy-ont: <http://energy.deri.ie/ontology#> .
@prefix deri-energy: <http://energy.deri.ie#> .
@prefix ssn: <http://purl.oclc.org/NET/ssnx/ssn#> .
@prefix deri-rooms: <http://lab.linkeddata.deri.ie/2010/deri-rooms#> .

deri-rooms:building      rdf:type rooms:Building;
                        rooms:contains deri-rooms:1st-floor.
deri-rooms:1st-floor     rdf:type rooms:Floor;
                        rooms:contains deri-rooms:1st-floor-south-wing.
deri-rooms:1st-floor-south-wing  rdf:type rooms:Wing;
                        rooms:contains deri-rooms:r202e.
deri-rooms:r202e        rdf:type rooms:Desk;
                        rooms:occupant <http://www.deri.ie/about/team/member/edward_curry#me>.
<http://www.deri.ie/about/team/member/edward_curry#me>  rdf:type foaf:Person;
                                                                deri-energy-ont:has_a deri-energy:macbook15698;
deri-energy:macbook15698  rdf:type deri-energy-ont:Laptop;
                        ssn:observedBy deri-energy:sensor123.
deri-energy:sensor123    rdf:type ssn:Sensor;
                        ssn:observes deri-energy-ont:PowerConsumption.
deri-energy:observation1236  rdf:type ssn:Observation;
                        ssn:observationResult deri-energy:observationResult1236;
                        ssn:featureOfInterest deri-energy:macbook15698;
                        ssn:observedBy deri-energy:sensor123.
deri-energy:observationResult1236  rdf:type ssn:SensorOutput;
                        ssn:hasValue deri-energy:vale1236.
deri-energy:vale1236     deri-energy-ont:hasQuantityValue 31.5;
                        deri-energy-ont:hasQuantityUnitOfMeasurement deri-energy-ont:kWh.
```

**Fig. 5.** Results of a dereferenced URI for the Room 202e.

### 5.3.3. Support services

The next level contains the dataspace support services that are designed to simplify the consumption of the linked data cloud by encapsulating common services for reuse. Some example support services used in the use case are:

- **Entity Management Service (EMS)** to improve data quality and inter-linkage between entity data scattered among legacy systems. The EMS provides a consolidated view of an entity from across the legacy systems using a single URI, or multiple URIs along with *owl:SameAs* relationships. The EMS leverages automatic entity consolidation algorithms that are supported by humans for collaborative data management [30].
- **Complex Event Processing (CEP) engine** [31] is used 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 [29]. When an event pattern is detected, one or more rules are activated and further events may be generated for applications such as dashboards. CEP engines may also be enabled by a reasoner such as ETALIS [38] and can use reasoning in order to match RDF events and generate complex events. For example, a CEP rule that expects energy consumption for instances of the class “*en:Device*” would match energy consumption events for the class “*en:Laptop*” given that the ontology states that “*en:Laptop rdfs:subClassOf en:Device*”.
- **Data Catalogue and Provenance service** [32] to track data sources with specific attributes such as freshness and publisher.
- **Search and Query services** [33] allow users to interact with the dataspace using structured or natural language interfaces.

### 5.3.4. Applications

At the top of the architecture are the applications that consume the resulting data and events from the linked building data cloud. Applications come in many forms such as dashboards to graphically present data in a user-friendly manner and building energy analysis tools to gain a deeper understanding of the energy consumption behaviors in the building to enable effective decision support.

## 5.4. Vocabulary used in the linked building data cloud

The data in each of the existing legacy systems are described using a different vocabulary. This section describes the key vocabularies used from the building management perspective. The discussion uses a minimal representation of these vocabularies to allow discussion focus on interoperability concerns. It should be noted that in many cases more expressive knowledge representation such as ontologies could be used to enable reasoning behaviors [34].

The example queries are expressed using the SPARQL query language. SPARQL is an official W3C recommendation that is able to retrieve and manipulate data stored in RDF. SPARQL allows for a query to consist of triple patterns, conjunctions, disjunctions, and optional patterns. Queries are run against a SPARQL *endpoint*, and SPARQL queries can be translated to/from other query languages (i.e. SQL and XQuery).

The remainder of this section details the vocabularies used and how they are linked together to allow for a holistic queryable information overview of the building's operations.

### 5.4.1. People, groups, and devices

People, Groups and Devices are tracked using two vocabularies 1) FOAF, a simple vocabulary for describing people and groups (prefix *foaf:*) and 2) DERI Energy: covering devices and laptops (prefix *en:*). As illustrated in Fig. 6, these vocabularies track the

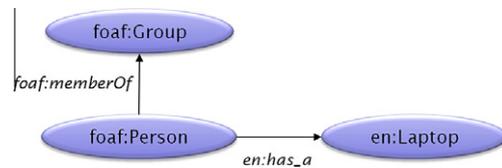


Fig. 6. People, groups, and devices graph.

people in the DERI building, what groups they are associated with, and the devices they use. The vocabulary is used by the DERI administration and IT staff to perform queries on group membership, and device allocation.

Using the vocabularies the following question can be asked: *Who are the members of the Green IT group?* This question can be represented in the SPARQL query language as:

```
SELECT ?member
WHERE
{
  ?member rdf:type foaf:Person.
  ?member foaf:memberOf <http://www.deri.ie/
    research/units#dgsit>
}
```

The vocabularies support other types of questions including:

- What laptop does Edward Curry have?
- What laptops does the Green IT group use?
- Who is using a MacBook Pro?

### 5.4.2. Building and rooms

Buildings and Rooms (<http://www.vocab.deri.ie/rooms>), illustrated in Fig. 7, is a simple vocabulary for describing the rooms in a building (prefix *rm:*). The vocabulary is used by the DERI administration staff within the resource allocation process to assign desks to people in the building. The vocabulary can be used to perform queries on the structure of the building, as well as the location of people.

The following question can be asked using the vocabulary: *How many floors are in the Building?* This question can be represented in the SPARQL query language as:

```
SELECT count(?floor)
WHERE
{
  ?floor rdf:type rm:Floor.
}
```

The vocabulary can also be used to ask:

- How many desks are located on the first floor?
- Who is the occupant of Room 202e?
- Where does Edward Curry sit?
- How many people are in the building?

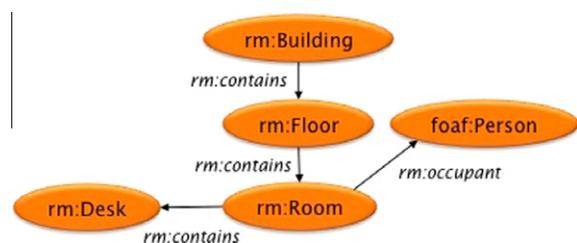


Fig. 7. Building and rooms vocabulary graph.

**Table 3**  
IFC to RDF mapping.

IFC entity	RDF entity
IFC Building	Building
IFCRelAggrgatss	Contains
IFCBuildingStorey	Floor
IFCRelAggregates	Contains
IFCSpace	Room
IFCRelContainsInSpatialStructure	Contains
IFC Furnishing Element	Desk
IFCRelDefinesByType	
IFCFurnitureType.Name = Table	
IFCRelContainsInSpatialStructure	Occupant
IFC Person	Person

This vocabulary is analogous to geometry information obtained from a BIM. A manual mapping of IFC to RDF was done based on IFC files exported from the ArchiCAD BIM authoring tool to ifcOWL [21]. This mapping, detailed in Table 3, represents a subset of the spatial representation provided by BIM through IFCBuilding-IFCBuildingStorey-IFCSpace hierarchy and IFCFurnishingElement and IFCPerson are related to IFCSpace.

#### 5.4.3. Energy sensor network

The Energy Sensor Network installed within the DERI building is described using the Semantic Sensor Network (<http://www.purl.org/NET/ssnx/ssn>) vocabulary for sensors and observations (prefix *ssn:*) together with the DERI Energy vocabulary. Both vocabularies are used to track energy usage within the DERI building; the key concepts are illustrated in Fig. 8. The vocabulary can be used to perform queries on sensors installed in the building and connected to devices.

The following question can be asked: *What is the total energy consumption of Room 202e?* This question can be represented in the SPARQL query language as:

```
SELECT sum(?consumption)
WHERE{
deri-rooms:r202e rdf:type rm:Room.
?device deri-energy-ont:hasLocation deri-rooms:r202e.
?device rdf:type deri-energy-ont:Device.
?device ssn:observedBy ?sensor.
?sensor rdf:type ssn:Sensor.
?sensor ssn:observes deri-energy-ont:PowerConsumption.
?observation ssn:featureOfInterest ?device.
?observation rdf:type ssn:Observation.
?observation ssn:observationResult ?result.
?result rdf:type ssn:SensorOutput.
?result ssn:hasValue ?value.
?value deri-energy-ont:hasQuantityValue?consumption.
?value deri-energy-ont:hasQuantityUnitOfMeasurement deri-energy-ont:kWh.
}
```

Other more granular and exploratory queries possible with the vocabulary include:

- What is the energy consumption of a specific laptop?
- What is the average energy consumption of a MacBook Pro vs. a Lenovo?
- What is the energy consumption of the lights in Room 202e?
- What is the total energy consumption of heaters in the building?

#### 5.5. Linking the vocabularies

The objective of linking building data in the cloud is to create an integrated well-connected graph of relevant information for managing a building. Linked data links at the information-level (data) not the infrastructure-level (system) by focusing more on the conceptual similarities (shared understanding) between information. LEI follows an entity-centric approach that focuses on the concepts that exist within the systems, for example, business entities like employees, products, customers, rooms, buildings, etc. Entities are enriched with data from multiple systems. This results in a cloud of interlinked entities with links to relevant knowledge and contextual information from across all the information systems that have exposed linked data. New data sources and links may be added in an incremental manner to facilitate a *pay-as-you-go* methodology for data integration allowing the incremental enrichment of entity profiles. Agreement on data schemas can be achieved in a number of ways, from minimal agreement on a small number of simple vocabularies, to the use of agreed upon domain ontologies. The most appropriate approach will depend on the type of data.

An important step in merging data is the issue of identity and uniquely identifying resources, which is essential for integrating data across sources. Matching identifiers across sources can lead to problematic data integration and can cause data to remain fragmented across multiple instances. Fusing identifiers is especially important for applications relying on the resulting data [35]. The problem of entity consolidation has received significant attention in the database community as record linkage, instance fusion, and duplicate identification. Entity consolidation remains an active research area as evidenced by the literature. A full discussion of the possible approaches are outside the scope of this article, for further information we direct the reader to [36].

The EMS plays an important role within LEI by merging datasets where different URIs exist for identifying the same entity. Currently there is limited agreement on the use of common URIs for the same instances across sources; it is more common for the same entity to be identified with multiple URIs. Where more than one URI exists for identifying a resource, associations between resources identifiers are needed to integrate the data. To facilitate the merging of data from multiple sources, OWL constructs such as *“sameAs”* supports entity consolidation. Expressing equivalences using *owl:sameAs* can be used to state that different resources are actually the same. When resources are determined to be the same by these associations, information about them from different sources can be merged. Take for example the three vocabularies described in this section. The key to merging these data sources is to identify the resources that appear in multiple sources and create *owl:sameAs* relationships that map their equivalence as illustrated in Fig. 9.

The following triple maps the identifiers for *Room 202e* between the Building and Rooms and DERI Energy sources using the *“owl:sameAs”* predicated:

```
http://www.lab.linkeddata.deri.ie/2010/deri-rooms#r202e
owl:sameAs
http://www.energy.deri.ie#Room202e
```

Within the vocabularies presented the overlapping resources are People, Rooms and Laptops, which can be mapped as illustrated in Fig. 10.

This merged graph can be used to perform queries across all three graphs. The following types of questions can be asked:

- What is the total energy consumption of rooms on the first floor?

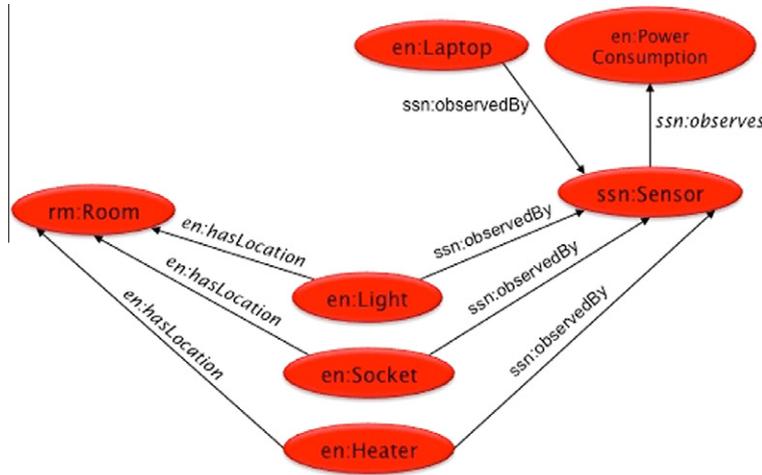


Fig. 8. Energy sensor network vocabulary graph.

- What is Edward Curry’s total energy consumption? (Laptop, lights, heats, etc.)
- What is the energy consumption of the laptops for the Green IT group?
- What is the total consumption of members of the Green IT group on the first floor?
- What is the laptop energy usage of members of the Green IT group on the first floor?
- What is the average energy cost of offices with 2 occupants?

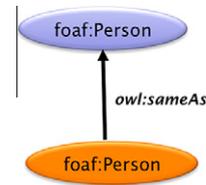


Fig. 9. SameAs relationships between rooms and people vocabulary graphs.

tion that examines the environmental impacts of the DERI organization as a whole.

5.6. Applications and services using the linked building data cloud

Within DERI a number of applications utilize the linked building data cloud. This section discusses some of these applications, including a real-time service for energy awareness within the DERI building, an energy-tracking dashboard for understanding the causes of building energy usage, and the Sustainable DERI applica-

5.6.1. Energy situation awareness service

With the large volume of information available from both legacy systems and real-time sensors it is becoming increasingly more difficult for users to understand and interpret the data to make decisions. In order to assist users to understand the data there is a need to create situation awareness (SA) systems where the multitude of data can be synthesized into a single holistic picture of

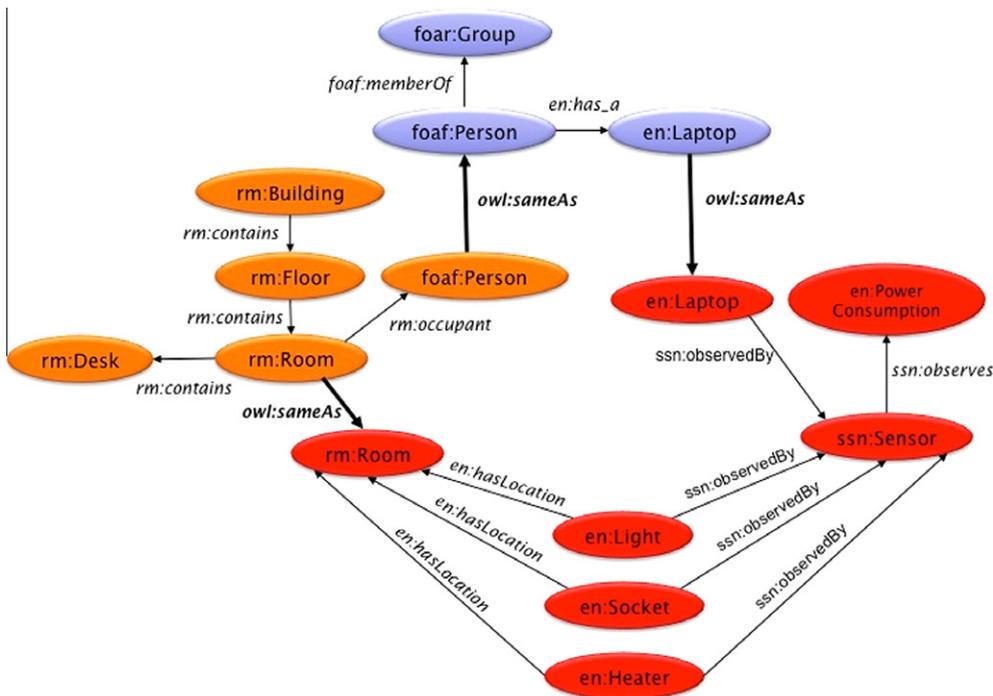


Fig. 10. Merged vocabulary graph.

entities and their relationships within an environment. With this picture in place, users can now use it to understand the environment and make appropriate decisions.

The Energy Situation Awareness Service is a performance evaluation for energy usage within the DERI building. The service follows the scenario modeling concept [10] and uses both energy performance objectives and performance objectives from other related domains, such as human resources (i.e. occupancy comfort) and enterprise resource planning (i.e. room utilization), to understand the performance of the facility in meeting business objectives.

The service is implemented using the CEP support service to define situations of interest within an energy management context [31]. The CEP engine process the low-level energy sensor readings, evaluating them for a pattern match. When a pattern is matched, a new higher-level energy situation awareness event (e.g. *ExcessiveEnergyUsage*) is generated. These energy situation awareness events are forwarded to human users via emails or SMS, they can also be forwarded via messaging systems [37] to applications such as decision support systems or dashboard UIs. An example energy situational awareness widget is provided in the next section.

### 5.6.2. Building energy explorer

The Building Energy Explorer dashboard is a central user interface that makes extensive use of the merged data within the linked building data cloud. The dashboard visualizes relevant linked building data together with energy consumption sensor data and presents it in an actionable manner that requires minimal effort

for users to leverage the knowledge within energy-related decision-making. The objective of the dashboard is to help users identify energy leaks and non-ecological actions within the DERI building by linking actual energy usage data to the entity(ies) responsible for the energy usage. This linking of information adds the necessary context to understand the causes of energy usage, which can then be utilized to achieve reductions in energy consumption and cost saving.

The main screen of the dashboard is presented in Fig. 11, with in box (1) data from the rooms, people, and group vocabulary can be seen; it is used to add context to the energy consumption readings. In (2) historical usage along with real-time instant measures from the energy sensors are shown, along with a breakout for consumption type (lights, heat, sockets). The interface also displays the output of the Energy Situation Awareness Service via a widget in (3). This particular widget performs energy situation assessment by comparing the accumulative consumption with historical usage data and usage targets to detect high usage situations. In the widget, two bars are used to show the daily-accumulated energy usage in comparison with the monthly average to highlight any deviations in the consumption pattern. Other widgets can be utilized based on different performance criteria including cost, occupancy comfort, and average energy usage per occupant.

### 5.6.3. Sustainable DERI

There is a growing need for organizations to develop sustainable business practices to reduce their impacts on the environment

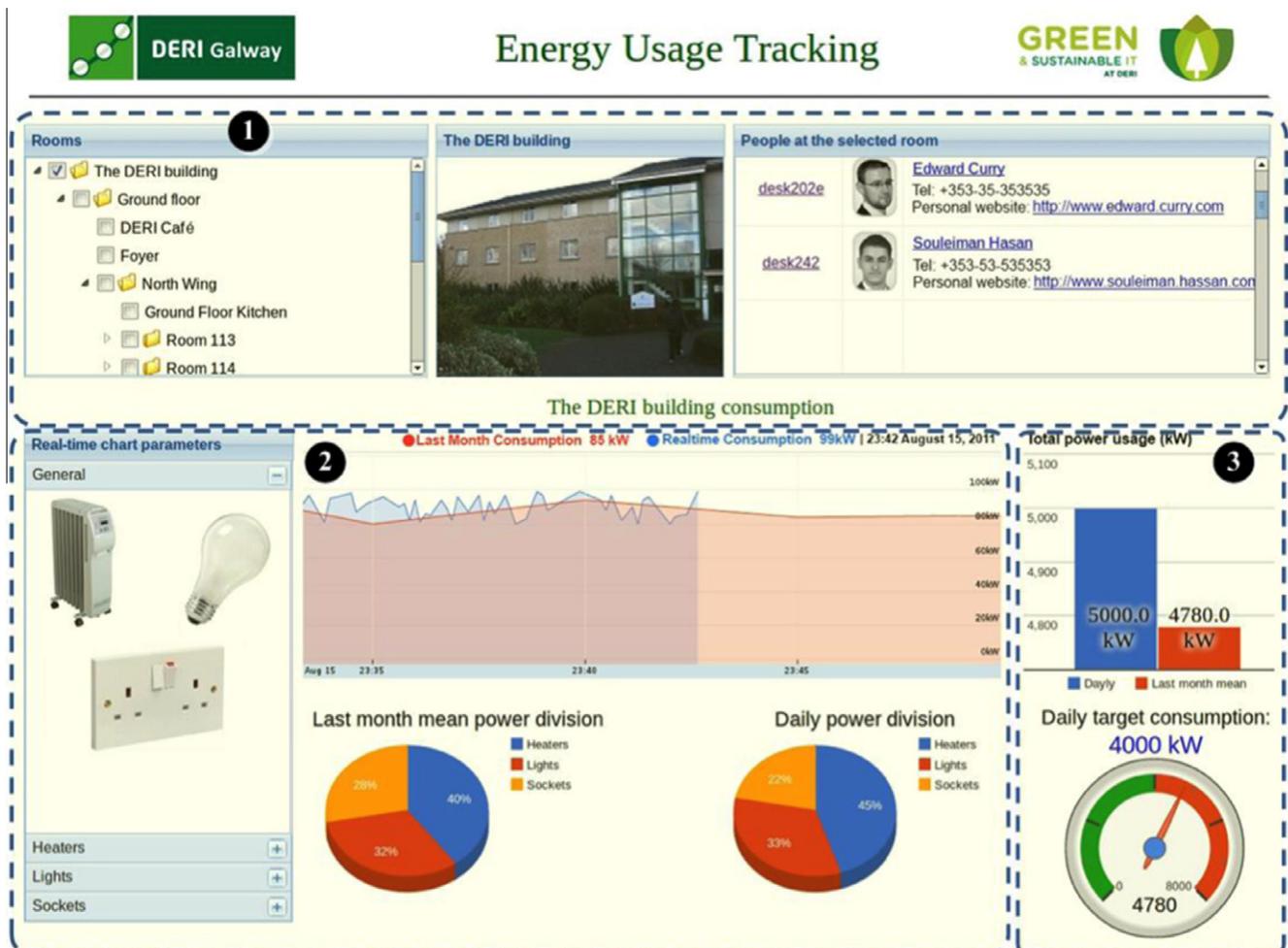


Fig. 11. Building energy explorer dashboard.

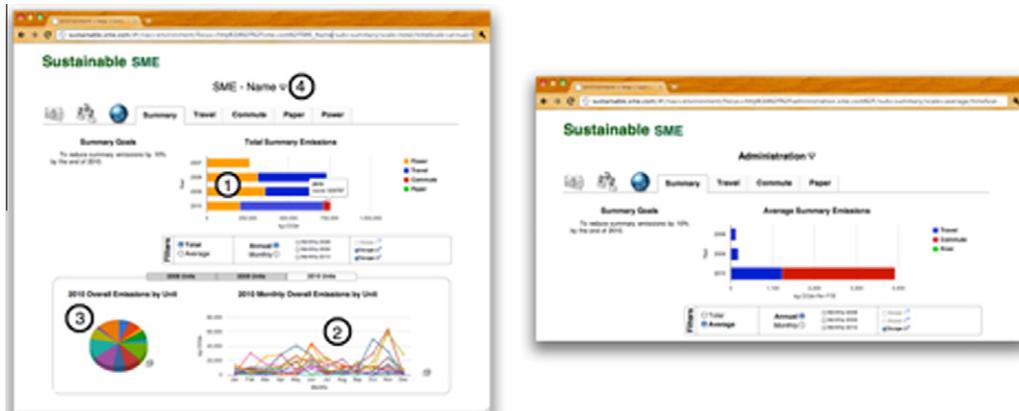


Fig. 12. Two screenshots of sustainable DERI.

with many organizations seeing sustainability as a key business issue [38]. Within organization Green Information Systems (Green IS) are needed to assist in forming a cohesive representation of the environmental impact of business operations at both micro- and macro-levels [39]. The Sustainable DERI [27] application provides near-real time information on the environmental impacts of DERI business activities.

The DERI building and its energy usage form a significant part of DERI's environmental footprint. The data in Sustainable DERI is published as linked data and is used in the linked building data cloud. Since relevant building data (i.e. power consumption) has also been published using linked data it can be easily linked to and consumed by Sustainable DERI. This allows Sustainable DERI to have detailed information on facilities usage and associated power usage. The power information is merged with data from other information systems which track environmental concerns, including business travel, commute, and print usage. The resulting integrated data allows stakeholders within DERI to understand their overall environmental impact, allowing them to incorporate sustainable business concerns within their day-to-day activities and decisions. The tool also allows senior management to understand the macro-level sustainable business concerns of the organization.

Sustainable DERI includes a tool to communicate this information via a multi-level model that makes use of the linked data cloud. In each mode of the interface, illustrated in Fig. 12, the user can see the relative impact of individual KPIs (circle 1), see trends within and between KPIs over time (circles 1 and 2). These dashboards can assist in decision making via four specific areas:

1. Real-time direct action items that employees can respond to while in the planning stages of emission-producing activities.
2. Daily direct action items that employees can respond to after reviewing their previous day's activities.
3. Monthly direct action items that employees can respond to by understanding the extrapolated effects of their daily activities.
4. Monthly indirect action items that employees can respond to by understanding the effect of their day-to-day activities as averaged over the enterprise.

### 5.7. Case study summary and benefits

The goal of the case study was to prove that linked data technology is an effective mechanism for improving the interoperability of cross-domain data sources when they are moved to a cloud-based data service. The case study connected over 14 existing systems and was used to deliver a number of building energy management

tools for end users from different domains across the organization. For each of these tools the linked building data cloud provided the linkages required to integrate the necessary operational legacy information system, and real-time data streams, in a straightforward and flexible manner.

## 6. Conclusions and future work

The holistic management and maintenance of facilities is a multi-domain problem encompassing financial accounting, building maintenance, facility management, human resources, asset management and code compliance, affecting different stakeholders in different ways. In order to manage a building holistically it is important to use knowledge from across these information sources. However, many barriers exist to their interoperability. The move to cloud computing offers many potential benefits. Data as a Service (DaaS) is based on the concept that data can be provided on demand to the user regardless of geographic or organizational separation of provider and consumer. However, if we design these new cloud-based information services in the same manner as our traditional systems they will still suffer from the data interoperability problems we currently face. Linked data technology leverages existing open protocols and W3C standards of the Web architecture for sharing structured data on the web.

In this paper we proposed the use of linked data as a technology for cloud-based data services. Linked data provides a mechanism through which all silos can exist in a homogeneous format. Most importantly, linked data principles identify common elements between silos and where possible interlink silos. The approach was demonstrated in an owner-occupied office building using the Linked dataspace for Energy Intelligence. The resulting merged cross-domain linked building data cloud provides a holistic view of the buildings operations for stakeholders across the organizations. A number of energy management applications were built using the merged data including a real-time energy awareness system and an audit-style energy tracking system.

Future work will focus in a number of areas including, the use of more expressive knowledge representations for reasoning within the linked building data cloud, the development of appropriate support services for linked dataspace (in particular the use of approximate semantic matching techniques [40]), and the interoperability of software services as RESTful services. We will also continue to investigate the definition of more sophisticated scenario models using the cross-domain data within the linked building data cloud.

## Acknowledgement

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

## References

- [1] T. Heath, C. Bizer, Linked data: evolving the web into a global data space, *Synthesis Lectures on the Semantic Web: Theory and Technology* 1 (1) (2011) 1–136.
- [2] D. Clements-Croome, *Intelligent Buildings: Design, Management and Operation*, Thomas Telford Services Ltd., 2004.
- [3] P. Teicholz, Labor productivity declines in the construction industry: causes and remedies, *AECbytes Viewpoint* 4 (14) (2004) 2004.
- [4] M.P. Gallaher, R.E. Chapman, Building, F. R. L. (US), O. of Applied Economics, N. I. of Standards, T. (US), I. Technology, E. Office, Cost Analysis of Inadequate Interoperability in the US Capital Facilities Industry. US Dept. of Commerce, Technology Administration, National Institute of Standards and Technology, 2004.
- [5] BuildingSMART, Industry Foundation Classes (IFC) Homepage, 2011.
- [6] R.J. Hitchcock, Standardized Building Performance Metrics, Lawrence Berkeley National Laboratory, 2003.
- [7] D. Claridge, M. Liu, W.D. Turner, Commissioning of Existing Buildings-State of Technology and It's Implementation, in: *International Symposium on HVAC Commissioning*, Kyoto, 2003.
- [8] C. Neumann, D. Jacob, Guidelines for the Evaluation of Building Performance, Building EQ, Project Report, 2008.
- [9] C. Turner, M. Frankel, U.S.G.B. Council, Energy performance of LEED for new construction buildings. New Buildings Institute Washington, DC, 2008.
- [10] J. O'Donnell, Specification of Optimum Holistic Building Environmental and Energy Performance Information to Support Informed Decision Making, Doctorate, University College Cork, Ireland, University College Cork, Ireland, 2009.
- [11] E. Corry, M. Keane, J. O'Donnell, A. Costa, Systematic development of an operational bim utilising simulation and performance data in building operation, in: *IBPSA Building Simulation 2011*, 2011.
- [12] C.M. Eastman, P. Teicholz, R. Sacks, K. Liston, *BIM Handbook*, Wiley Online Library, 2008.
- [13] K.K. Fallon, M.E. Palmer, *General Buildings Information Handover Guide: Principles, Methodology and Case Studies (NISTIR 7417)*, August, 2007.
- [14] D.A. Jordani, BIM and FM: The portal to lifecycle facility management, *Journal of Building Information Modeling* (2010) 13–16.
- [15] BuildingSMART, IFC Solutions Factory, The Model View Definition Site, 2011.
- [16] W. Shen, Q. Hao, H. Mak, J. Neelamkavil, H. Xie, J. Dickinson, R. Thomas, A. Pardasani, H. Xue, Systems integration and collaboration in architecture, engineering, construction, and facilities management: a review, *Advanced Engineering Informatics* 24 (2) (2010) 196–207.
- [17] S. O'Riain, E. Curry, A. Harth, XBRL and open data for global financial ecosystems: a linked data approach, *International Journal of Accounting Information Systems* 13 (2) (2012) 141–162.
- [18] E. Curry, A. Freitas, S. O'Riain, The role of community-driven data curation for enterprises, in: D. Wood (Ed.), *Linking Enterprise Data*, Springer, Boston, MA, US, 2010, pp. 25–47.
- [19] R. Cyganiak, A. Jentzsch, "Linked Open Data Cloud Diagram," 2011.
- [20] F. Manola, E. Miller, B. McBride, RDF primer, *W3C Recommendation* 10 (2004) 1–107.
- [21] J. Beetz, J. Van Leeuwen, B. De Vries, IfcOWL: a case of transforming EXPRESS schemas into ontologies, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 23 (01) (2009) 89–101.
- [22] M. Franklin, A. Halevy, D. Maier, From databases to dataspace: a new abstraction for information management, *Sigmod Record* 34 (4) (2005) 27–33.
- [23] E. Curry, S. Hasan, M. White, H. Melvin, An environmental chargeback for data center and cloud computing consumers, in: *First International Workshop on Energy-Efficient Data Centers*, 2012, pp. 117–128.
- [24] E. Curry, S. Hasan, S. O'Riain, Enterprise energy management using a linked dataspace for energy intelligence, in: *The Second IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT 2012)*, 2012.
- [25] E. Curry, System of systems information interoperability using a linked dataspace, in: *IEEE 7th International Conference on System of Systems Engineering (SOSE 2012)*, 2012, pp. 101–106.
- [26] D. Raggett, How can we exploit XBRL and semantic Web technologies to realize the opportunities? in: *19th XBRL International Conference*, Paris, 2009.
- [27] E. Curry, S. Hasan, U. ul Hassan, M. Herstand, S. O'Riain, An entity-centric approach to green information systems, in: *Proceedings of the 19th European Conference on Information Systems (ECIS 2011)*, 2011.
- [28] B. McBride, Jena: a semantic web toolkit, *Internet Computing, IEEE* 6 (6) (2002) 55–59.
- [29] L. Lefort, C. Henson, K. Taylor, P. Barnaghi, Semantic Sensor Network XG Final Report, W3C Incubator Group Report (28 June 2011), 2011.
- [30] U. ul Hassan, S. O'Riain, E. Curry, Leveraging matching dependencies for guided user feedback in linked data applications, in: *9th International Workshop on Information Integration on the Web (IIWeb2012)*, 2012.
- [31] S. Hasan, E. Curry, M. Banduk, S. O'Riain, Toward situation awareness for the semantic sensor web: complex event processing with dynamic linked data enrichment, in: *4th International Workshop on Semantic Sensor Networks 2011 (SSN11)*, 2011, pp. 69–81.
- [32] A. Freitas, T. Knap, S. O'Riain, E. Curry, W3P: building an OPM based provenance model for the web, *Future Generation Computer Systems* 27 (6) (2011) 766–774.
- [33] A. Freitas, J. Oliveira, S. O'Riain, E. Curry, J. Pereira da Silva, Querying Linked Data using Semantic Relatedness: A Vocabulary Independent Approach, in: *Proceedings of the 16th International Conference on Applications of Natural Language to Information Systems, NLDB 2011*, 2011, vol. 6716, pp. 40–51.
- [34] P. Pauwels, D. Van Deursen, R. Verstraeten, J. De Roo, R. De Meyer, R. de Walle, J. Van Campenhout, A semantic rule checking environment for building performance checking, *Automation in Construction* (2010).
- [35] A. Freitas, E. Curry, J.G. Oliveira, S. O'Riain, Querying heterogeneous datasets on the linked data web: challenges, approaches, and trends, *IEEE Internet Computing* 16 (1) (2012) 24–33.
- [36] M. Michalowski, S. Thakkar, C. A. Knoblock, Exploiting secondary sources for automatic object consolidation, in: *Proceeding of 2003 KDD Workshop on Data Cleaning, Record Linkage, and Object Consolidation*, 2003, pp. 34–36.
- [37] E. Curry, Message-oriented middleware, in: Q.H. Mahmoud (Ed.), *Middleware for Communications*, John Wiley and Sons, Inc., 2004, pp. 1–28.
- [38] E. Curry, B. Guyon, C. Sheridan, B. Donnellan, Developing an sustainable IT capability: lessons from Intel's Journey, *MIS Quarterly Executive* 11 (2) (2012) 61–74.
- [39] E. Curry, B. Donnellan, Sustainable Information Systems and Green Metrics, in: S. Murugesan, G.R. Gangadharan (Eds.), *Harnessing Green IT: Principles and Practices*, John Wiley & Sons, Inc., 2012, pp. 167–198.
- [40] S. Hasan, S. O'Riain, E. Curry, Approximate semantic matching of heterogeneous events, in: *6th ACM International Conference on Distributed Event-Based Systems (DEBS 2012)*, 2012, pp. 252–263.