Kris McGlinn, Christophe Debruyne, Lorraine McNerney, and Declan O'Sullivan. Integrating Ireland's geospatial information to provide authoritative building information models. In SEMANTICS, pages 57–64. ACM, 2017

Integrating Ireland's Geospatial Information to Provide Authoritative Building Information Models

Kris McGlinn
ADAPT, Trinity College Dublin
Dublin 2, Ireland
kris.mcglinn@scss.tcd.ie

Christophe Debruyne
ADAPT, Trinity College Dublin
Dublin 2, Ireland
christophe.debruyne@adaptcentre.ie

Lorraine McNerney
Ordnance Survey Ireland
Dublin 8, Ireland
lorraine.mcnerney@osi.ie

Declan O'Sullivan
ADAPT, Trinity College Dublin
Dublin 2, Ireland
declan.osullivan@adaptcentre.ie

ABSTRACT

Building Information Modelling (BIM) is a key enabler to support integration of building data within the buildings life cycle and is an important aspect to support a wide range of use cases, related to building navigation, control, sustainability, etc. Open BIM faces several challenges related to standardization, data interdependency, data access, and security. In addition to these technical challenges, there remains the barrier among BIM developers who wish to protect their intellectual property, as full 3D BIM development requires expertise and effort. This means that there is often limited availability of BIM models. In Ireland, the Ordnance Survey Ireland (OSi) has a substantial dataset which includes not only GIS data (polygon footprint, geodetic coordinate), but also additional building specific data (form and function). In this paper we demonstrate the use of an applied and tested methodology for uplifting GIS data (relational data) into RDF (GeoSPARQL and OSi ontology) and demonstrate how this data is used for interlinking to other building data with an initial, simple exploratory example, taken from DBpedia. By interlinking building data and making it available, new insights about buildings in Ireland can be made, currently not possible due to lack of availability of data. This is an important step towards the iterative integration of ever more complex BIM models into the wider web of data to support the aforementioned use cases. [1]

CCS CONCEPTS

• **Information systems → Resource Description Framework (RDF); Geographic information systems; Building Information Models**

KEYWORDS

GeoSPARQL, Linked Data, Interlinking, Information Modelling, Geographic Information System, Ordnance Survey Ireland

ACM Reference format:

1 INTRODUCTION

Access to reliable structured data plays a central role in supporting existing and future services for managing smart and sustainable buildings and cities [1]. ICT solutions are becomingly increasingly important for supporting new control and monitoring capabilities for managing buildings. Building Information Modelling (BIM) has been identified as a key enabler to support integration of building data not only within the buildings life cycle (BLC), which includes its design, construction, operation and re-design, as well as demolition/recycling [2], but also with other data sources, such as those related to geolocation, people and their behaviour, weather, energy, etc. [3]. Currently, there is a great need for BIM to support new services in the operational stages of the building [4]. Open BIM faces several challenges, related to; standardization, data interdependency, data access and security [5]. In addition to these technical challenges, there remains the barrier amongst BIM developers who wish to protect their intellectual property, as full 3D BIM development requires expertise and effort [6]. This means that there is often limited availability of BIM. The potential therefore to link to available open datasets, can provide a source of enrichment for the data within a BIM

In 2014, Ordnance Survey Ireland (OSi, Ireland's national mapping agency) delivered a newly developed spatial data storage model known as Prime2[2]. With Prime2, OSi moved from a traditional map-centric model towards an object-oriented model from which various types of mapping and data services can be produced. Prime2 currently holds information of over 45,000,000 spatial objects (road segments, buildings, fences, etc.), of which some have more than one representation. Prime2 not only stores GIS data (polygon foot print, geodetic coordinate), but also captures additional building data (such as address, form and function, and also provenance data related to changes made to the

[2] Prime2: Data Concepts and Data Model Overview. Tech. rep., Ordnance Survey Ireland (2014), http://www.osi.ie/wp-content/uploads/2015/04/Prime2-V-2.pdf

building over time). This data meets some basic requirements for open BIM. For example: how to handle multiple building representations, geometries and how to combine this with authoritative geospatial datasets.

In this paper, we demonstrate the use of an applied and tested methodology for uplifting the OSi data into RDF and explore how this data, using Linked Data [16], can be a basis for interlinking to other open building data to support a wide range of use cases, related to building navigation, control, sustainability, etc. The efficacy of the approach is also explored based upon a simple example using DBpedia, to demonstrate the interlinking of DBpedia data with OSi form and function data. The paper proposes that this approach to uplifting data can provide new insights and knowledge about buildings in Ireland and the integration of authoritative OSi geospatial data is an important step towards the iterative integration of ever more complex BIM models into the wider web of data.

2 BACKGROUND AND RELATED WORK

2.1 Geospatial Data in Ireland

The OSi aims to leverage user engagement with their geospatial information (and derived maps), as it has a legal weight in Ireland. One of the initiatives they launched is called GeoHive (http://www.geohive.ie/), allowing one easy access to publically available spatial data – but not as Linked Data (LD). Though OpenStreetMap, Google Maps, etc. allow people to easily engage with maps, the information provided by those are i) not authoritative, and ii) not always correct. One of the major discrepancies that can be observed between these services and the information provided by the OSi are the points that refer to buildings. Where the former usually uses the entrance as the point, the OSi uses a building's centroid as a reference next to keeping track on which street the main entrance can be found. The latter can thus give a better indication of the size or location of a building with respect to the surrounding streets, for example.

2.2 DATA.GEOHIVE.IE

The OSi aims to adopt Linked Data as one means to publish its geospatial data. By doing so, it facilitates the exploration, adoption and use of OSi's authoritative geospatial datasets. In [7], we reported on `data.geohive.ie`, which publishes and serves Ireland's authoritative boundary datasets as Linked Data on the Web. Boundary datasets were chosen as they have been made available under a CC BY 4.0 license as part of OSi's Open Data Release. Figure 1 depicts the geometry of County Dublin plotted on one of OSi's base maps. The platform was designed to support two use cases; 1) providing different "resolutions" of administrative boundaries, and 2) providing the evolution of these boundaries as ordered by, for instance, Statutory Instruments. With "resolutions" we mean the level of detail in the geometries that represent the boundaries; the higher the resolution, the bigger the string representing the boundary and, as a consequence, the higher the overhead. The first use case is supported by extending GeoSPARQL [8] with dedicated named graphs for each resolution. For the second use case, we extended PROV-O [9] with concepts such as "Statutory Instrument" (as a subclass of prov:Entity) and "Boundary Change" (as a subclass of prov:Activity).

2.3 Open and Closed Geospatial Data

The National Mapping Agreement (NMA)[3], launched in January 2017, is an Irish agreement that gives government departments and public sector bodies unrestricted access to most of OSi's geospatial data. We note that the boundary data made available in the previous subsection was made available with an accessible license under OSi's open data release. With the NMA, on the other hand, one can request access to other datasets such as buildings and infrastructure. Stakeholders who fall under the NMA can request the OSi for access to the data, which are currently made available as dumps and shared over FTP. Others who wish to avail of OSi's data that do not fall under this agreement, e.g., commercial entities, need to interact with the OSi's commercialization team and pay a license fee for obtaining the data, also shared as dumps. One can thus see that this data is not open, yet how the OSi currently provides access to the data is not scalable.

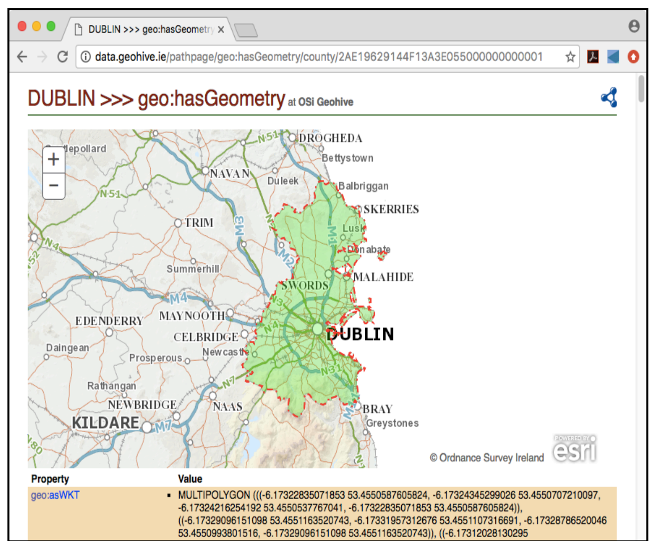

Figure 1: Plotting OSi's Polygon on OSi's base maps, part of the HTML served to users.

The ongoing collaboration between the OSi and ADAPT is currently investigating the use of semantic technologies to make available their data as so called "closed" Linked Data. The goal is to develop access control mechanisms allowing stakeholders to consume OSi's interlinked data in the way that they hold rights for. The outcome of this study, however, will be reported elsewhere. Irrespective of how it is accessed, the requirements for the buildings dataset, will be similar as those for the boundary dataset mentioned earlier: making available different representa-

[3] https://www.osi.ie/services/national-mapping-agreement/

tions (according to best practices in the domain, using standardized or well recognized ontologies), and the evolution of those based on PROV-O.

2.4 Building Information Modelling

Building Information Modelling (BIM) is a concept which has arisen to support the management of the vast amounts of data generated across a buildings life cycle [2]. BIM describes an integrated data model for storing all information relevant to the BLC. This can include a 3D model of an architectural design, electrical installations, fire protection, occupancy, energy consumption, costs, etc. A Building Information Model goes further than just providing consistent representation of objects; it also defines object parameters and relations to other objects. The use of BIM is active and growing in Ireland, with a 2016 Irish Digital Transition Survey reporting that 76% of respondents possess confidence in their organization's BIM skills and knowledge [10]. Typically, Irish BIM is looking to follow the UK process, which has had a strong drive to generate Level 2 BIM for all centrally-procured projects in England, Wales and Northern Ireland, but like England, challenges remain for SMEs who must weigh the known benefits against barriers, such as costs of software and training [11]. It is understood though, that a key requirement for BIM in Ireland is the adoption of existing standards and further standardization [12]. The availability of open BIM models are subject to these same issues in Ireland, as across the globe, i.e. that there are still many barriers to sharing BIM models, related to standardization, data interdependency, data access and security [5].

For the true potential of BIM to be realized, it is important that developers are given access to available, open and authoritative BIM. The supported use cases have the potential to demonstrate the benefits of making BIM data available to the wider community of developers in domains such as energy, building control, etc. Currently, GIS and other data sources which can be used to construct rudimentary BIM models based on location, and other attributes like address, are available openly, but are scattered between different services including the aforementioned data.gov, as well as DBpedia (http://wiki.dbpedia.org/). Linking these data sources with an authoritative dataset, will provide an important step toward making BIM available and the use of open standards is a necessary requirement to support this process. Standardisation of datasets is an important part of ensuring data interoperability.

Within the Architecture, Engineering and Construction (AEC) community the leading standard around the concept of BIM is Industry Foundation Classes (IFC), developed by buildingSmart [13]. IFC is also the only BIM currently an ISO PAS standard (ISO 2013), and so it remains a primary candidate for BIM. IFC is a non-proprietary data model which addresses several core data domains required for building AEC processes (architecture, structural analysis, control, etc.), enabling information to be passed between different stakeholders across the BLC. IFC has seen major government clients in the UK, Norway, and Finland, as well as a growing commitment in China [14] and the US [15]. In practice IFC has yet to make the impact expected of it in the AEC communities. One major barrier to the use of IFC is its complexity. Often it is difficult for new software developers, unfamiliar with the schema, to do simple things such as declare a building and assign a geolocation. To meet the requirements of the IFC standard, a complex set of relationship must be met which assign IfcBuilding to IfcSite, and also maintain a GeometricContext at the project level to have a global coordinate system for the buildings constituent entities. For a software developer who wants to declare a building with a geodetic location, or an address, or some other simple properties, this complexity can be off putting, resulting in the development of ad hoc models with alignments to IFC support an afterthought.

The Linked Building Data on the Web community group[4], which is working towards becoming a W3C working group (a draft charter can be found here[5]) is currently developing a Building Topology Ontology (BOT)[6] which aims to reduce the complexity of the reference model, by linking all building related ontologies to a very simple reference ontology which describes only the most fundamental properties of a building in terms of its topology. The intention here is to support linking to other ontologies when details are required for other aspects of the building, such as those related to geolocation, building products, automation and control, HVAC and energy as well as those data not traditionally within the scope of BIM, for example weather and energy tariffs. To support the process of linking, ontologies such as ifcOWL and BOT make use of Linked Data. In this paper, we demonstrate how to publish OSi authoritative geometric data with alignments to ifcOWL and other upcoming ontologies, such as BOT, using Linked Data principles.

2.5 Linked Data and BIM Linking

Linked Data (LD) is an approach to expose, share, and connect related data, which was not previously linked, on the Web [16]. RDF and textual (HTML) content do not just live next to each other on the Web of Data, but are also indirectly connected to each other. In modern AEC, data related to different domains such as building geometry and topology data, sensor data, behaviour data, geo data, are generated and consumed across BLC stages. The combination of BIM and LD has the potential to meet the requirements for storing and sharing those data. However, those data have to be represented as or at least tagged using RDF.

With the development of ifcOWL [17], now an official standard in buildingSMART [18] which transforms the well-established IFC standard defined in EXPRESS schema into OWL, querying using SPARQL has become an active area of research, with example queries to meet specific use cases are already being explored [19]. These developments are opening up the potential for linking ifcOWL, and other BIM ontologies, with other domains. Examples of openly available ontologies which explore domains of particular relevance to buildings are the Smart Ap-

[4] https://www.w3.org/community/lbd/
[5] https://w3c-lbd-cg.github.io/lbd/charter/
[6] https://w3c-lbd-cg.github.io/lbd/bot/2016/11/index-en.html

pliances Reference ontology (SAREF) for smart appliances and devices, in particular with reference to energy domains [20], which is an ETSI standard [21]. SAREF was built upon a comprehensive review of several related ontologies, e.g. FIEMSER [22] and DogOnt [23] and has alignments to the IoT communication standards OneM2M [24]. SAREF also explores links to IFC through the SAREF4BLDG extension. Another candidate for linking to BIM is the Semantic Sensor Network (SSN) for sensor devices domain [25]. SSN is also a W3C standard. Adapt4EE Occupancy Ontology for behaviour domain [26] and also KnoHolEM and Serum-iB which cover multiple domains [27][28]. An approach has also been developed to transform GbXML into OWL [29] in the building energy simulation domain, and there is an openly available version of CityGML as OWL [30]. From this snapshot of ontologies in the building domain, it can be seen that there is no shortage to satisfy a wide range of potential use cases in the building domain. Therefore methodologies for Linked Data generation to transform existing resources into Linked Data together with linking to authoritative building data, like that provided by the OSi, can provide a sound basis for interlinking these ontologies.

Several research projects have and are looking at the issue of linking Geospatial data with BIM. In [31] the integration of IFC for the preconstruction stage of a building, to support site planning, in terms of localization of materials and services appropriate for optimized productivity of a particular construction project, was examined. Other research has investigated the conversion of standards like IFC directly into CityGML [32] [33] [34].

As discussed previously, the OSi's authoritative geospatial data has been converted into GeoSPARQL[7] and made available through GeoHive (data.geohive.ie). GeoSPARQL is an Open Geospatial Consortium (OGC) standard which not only defines a vocabulary for representing geospatial data on the Semantic Web, but also specifies an extension to the SPARQL query language for processing that geospatial data. This builds upon existing work described in work by [35] in the geospatial domain. Some initial research has also examined GeoSPARQL for providing location information for buildings [36], although their use of owl:sameAs relations may result in incorrect reasoning over Linked Data datasets and is generally not recommended when linking [37]. In the remainder of this paper, we explore the conversion of the OSi building dataset into RDF, which includes location and other properties, such as form and function. We also explore the linking of this with other datasets for providing additional semantics (ifcOWL, BOT, GeoSPARQL), as a means to create an authoritative basis for linking BIM datasets.

3 ANALYSIS AND REQUIREMENTS FOR OSI LINKING TO BIM

Figure 2 provides a high level overview of the process for converting OSi Prime2 tabular data into RDF, publishing it and making it queryable to applications using SPARQL. This process has been demonstrated in previous research which explored the enrichment and use of OSi geospatial data based on location [39], which is a useful method for identifying whether two buildings are the same across datasets [38]. Here we apply the same methodology to OSi building data, exploring alignments with standard vocabularies for describing buildings, as well as exploring existing web of data datasets to support iterative BIM development. In order to publish OSi data using the aforementioned available ontologies (IFC, BOT and GeoSPARQL) and make them available for interlinking we first set out to analyse the available data store Prime2 to identify relevant concepts and alignments.

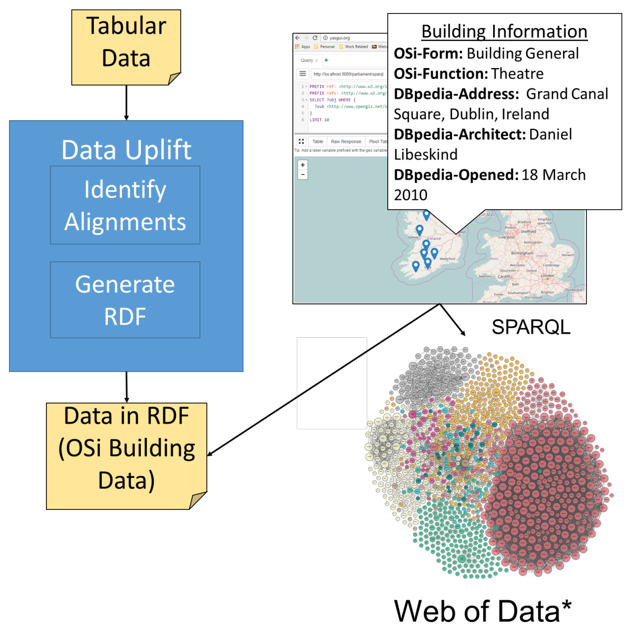

Figure 2: Process for Uplifting OSi Tabular Data, and example of integrated OSi and DBpedia data – (*) Linking Open Data cloud diagram 2017, by Abele, A. McCrae, J. Buitelaar,P. Jentzsch, A and Cyganiak, R. http://lod-cloud.net/

3.1 Analysis of OSi PRIME2 Building Data

The analysis of Prime2 building data structure is presented in Figure 3. The Prime2 **Building** is made up of several core concepts. The first we discuss here is **Placement**, which is related to the geolocation of the building. This is represented by a 2D point derived from determining the median of the polygon which represent the 2D footprint of the building. The second is **Geometry**. This is a geometric object (aforementioned polygon). In Prime2, this may also be a 3D object when an additional height property is included, thus conforming to LOD1 in CityGML. In Prime2, a geometry is encoded as a SDO geometric object[8]. Each point of the polygon is represented by its own geodetic coordinate according to a particular reference system, for example Irish Transverse Mercator (ITM), or WGS84 [39].

[7] http://www.opengeospatial.org/standards/geosparql

[8] https://docs.oracle.com/cd/B10501_01/appdev.920/a96630/sdo_objrelschema.html

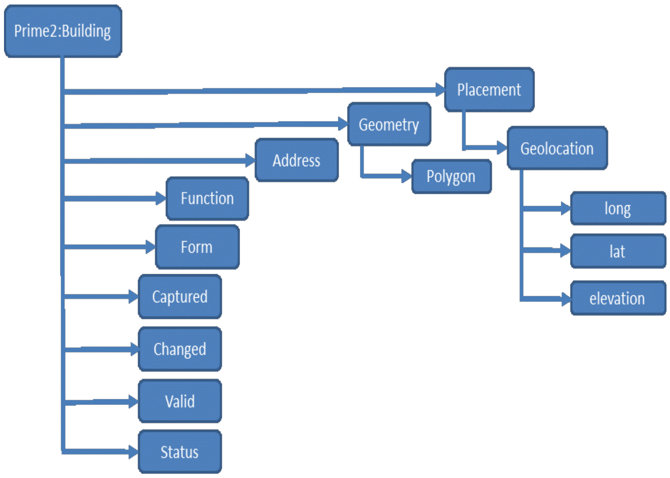

Figure 3: Overview of building concepts in OSi, Prime2

There is also a value which indicates the resolution of the data, e.g. 20 meters, 50 meters or 100 meters. The third concept is an **Address**. Currently, the Prime2 database references an address database using a geo_id, which is an integer. The fourth and fifth concepts are building specific. These are the **Form** and **Function** of the building. Form represents the physical form of the building, for example whether it is a "Building General" or "Barracks", and is an enumerated list of these type of values. Function, similarly is an enumerated list of values, but represents the use of the building, e.g. "Bank" or "Army Barracks". Some building forms have only one use, for example "Airport Terminal". Others, like building general can have multiple functions, e.g. airport building, bakery, courthouse, etc. The remaining concepts are related to who **Captured** the data (e.g. OSi), how the data was **Changed** (e.g. Re-engineered), and who **Validated** this, e.g. the OSi. **Status** represents the current status of the building, e.g. 'In Use', and is related to its BLC stage.

These enumerated values also have a direct numeric value representation. Captured, Changed and Valid can be assigned to geometric objects, to the form, to the function, etc. Some other concepts of note which we mention now but are not included in the diagram, are the capture specification, which indicates the specification used to store the data, e.g. Prime2. Next we explore our mappings of the Prime2 data model with existing ontologies, starting with previous work, already published [39], which examined GeoSPARQL to manage the geometric representation of the building.

3.2 Mappings to Support BIM Integration

Figure 4 gives an overview of mappings to existing standards to support BIM generation based upon the authoritative geospatial data in Prime2. Number 1 in Figure 4 presents the OSi concept of Building, Address and Form (or Function) which are created within the context of this work [41]. Building is aligned to Placement and Geometry through the use of GeoSPARQL. Address, Form and Function are aligned with IFC. GeoSPARQL is mapped to Placement and Geometry through the use of a sub-class relationship (see Figure 4, Number 2). Sub-class is indicated in Figure 4 by a line with a circle, e.g. osi:Building is a subclass of geos:Feature (geos = GeoSPARQL). GeoSPARQL is used to capture the geospatial coordinate of the building, as well as the polygon shape of the floor print, using Well-Known-Text (WKT) representations of polygons.

It should be noted, that Prime2 geometries are never more complex than LOD1, and these 3D objects can be represented in WKT. For more complex building geometries (LOD2 and above), WKT may not be sufficient and alternative methods need to be found. IFC can represent complex geometries, linked through the Object property relationship osi:hasIfcOwlRepresentation (in Figure 4), but how appropriate these methods are for RDF based geometry representations is an open research question [40], and not explored here. For Prime2, GeoSPARQL is sufficient to represent the available Prime2 geometries. For other Building related concepts such as Form and Function, we need to consider other standards. An important standard for managed BIM data is the Industry Foundation Classes (IFC), and therefore, it is important to be able to support interlinking of OSi data with IFC, if BIM integration is to be achieved. IFC is a complex data model, originally developed to address the Architecture, Construction and Engineering (AEC) domains with respect to buildings, in particular, the design and construction of buildings. Identifying appropriate alignments between Prime2 and IFC is therefore a non-trivial task. Our analysis has identified the need for the following high level IFC concepts (entities in the IFC EXPRESS schema); IfcProject, IfcSite and IfcBuilding (see Figure 5).

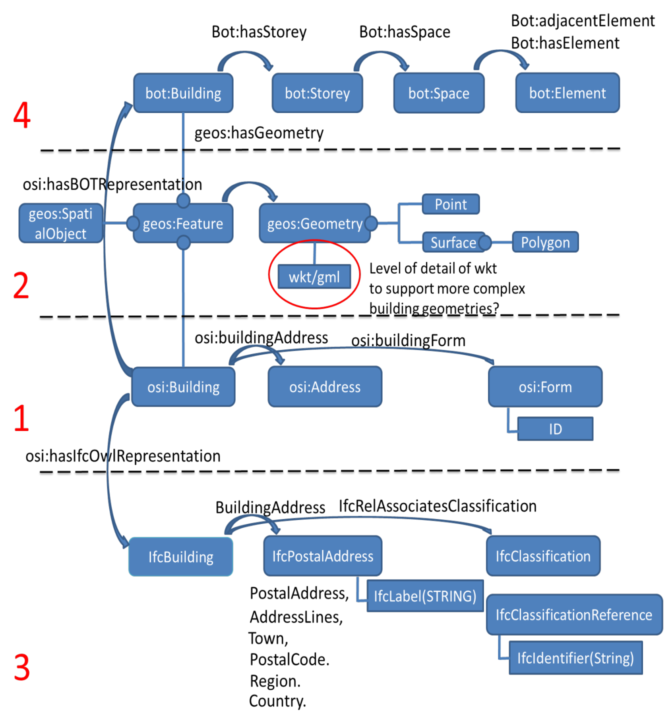

Figure 4: Mappings: OSi and GeoSPARQL, IFC and BOT

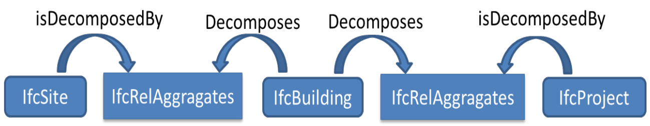

Figure 5 Mapping: OSi with IFC for Geolocation

IFC has a complex set of relations which must be met for organising composition of data (e.g. IfcRelAggregates[9]). We will ignore explaining these relationships here and instead focus on some of the more relevant entities which can be mapped to Prime2. Both IfcSite and IfcBuilding provide the capability to record postal addresses. Figure 4 demonstrates the use of IfcPostalAddress and the different properties that can be modelled using IFC entities. This can be used for recording Prime2 postal addresses, which are an important aspect for locating a building. IFC also has the capability to link to a classification system. Therefore, Form and Function can be modelled within IFC using IfcClassification through the IfcRelAssociatesClassification. An IfcClassification identifies a classification system along with a referencing classification key or id. It is therefore sufficient for Form and Function.

IFC can also be used to record information regarding geolocation, although this requires a large number of relations to be satisfied. Firstly, an IfcSite must be defined, as it is the only concept which provides a geodetic coordinate for the site. IfcBuilding provides capabilities to record the location of the building, relative to IfcSite through the IfcLocalPlacement, RelativePlacement relationship. IfcSite makes use of RefLatititude and RefLongitude relationship, which both point to IfcCompoundPlaneAngleMeasure concepts, which are lists of integers, for storing geodetic coordinates. IfcSite can also be used to record the elevation height of the Site using RefElevation and IfcLengthMeasure. IfcProject is then required to orientate the buildings geometry (IfcGeometricRepresentationContext). As can be seen, providing geolocation for a building using IFC is a complex process, and therefore, where possible, GeoSPARQL can provide a much more straightforward way of providing the Prime2 geolocation.

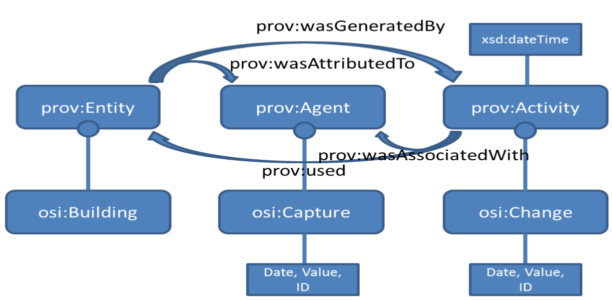

Figure 6 Mapping: OSi and PROV-O

[9]http://www.buildingsmart-tech.org/ifc/IFC2x4/rc2/html/schema/ifckernel/lexical/ifcrelaggregates.htm

There is also the capability within IFC to use IfcOwnerHistory, giving some data with respect to who "captured" the building data and when. We do not recommend this be used to record the Prime2 Captured concept, as IfcOwnerHistory only records the last made change. Instead, we recommend PROV-O, the W3C Recommendation for representing provenance, which is a strong candidate for handling the Captured, Changed and Valid aspects of Prime2 (Figure 6). This approach is being actively explored in related research, but is not directly addressed here.

To ease the process of interlinking with the OSi, we also explore mappings to the Building Topology Ontology (BOT) through the osi:hasBOTRepresentation object property. BOT has the advantage of being relatively simple to understand and use, and as such, to publish building related data. It is also in the early stages of development, and therefore open to extension with relevant Prime2 concepts. The main advantage of BOT though is that it is very lightweight and can be easily linked to other ontologies, providing a simple reference ontology to describe the topology of the building so that other specialized ontologies can handle other aspects of the building (e.g. SAREF[10] for smart appliances, SSN for sensors, etc.). BOT also links quite well with GeoSPARQL, by declaring bot:Building as a sub-class of geos:Feature.

4 PUBLISHING AND QUERYING OSI AND DBPEDIA BUILDING DATA

Once the data alignments have been identified the next step of Data Uplift is to generate the RDF (see Figure 2). Here the methodology developed in [39] is employed, which makes use of the R2RML mapping language, and define mappings to existing vocabularies. Listing 1 gives an example of a mapping of the Prime2 Building to the class Feature in GeoSPARQL, IfcBuilding in ifcOWL, and the OSi concept of Building, which is to be added to the publicly available OSi ontology. Here a triple graph called <#Building> is defined. Also defined is another graph for geometric data called <#BuildingPoint> and <#BuildingGeometry> based on the same principles as those for predicate object map for WKT points, also defined in [39].

R2RML Java libraries[11] are then used to convert the tabular data, which are taken from the Prime2 database. This RDF is then loaded into a triplestore that supports GeoSPARQL. GeoSPARQL provides geospatial functions, such as *nearby*, which takes a point and returns all points within a certain distance, e.g. 2 kilometres. This is used to locate a Feature in OSi, which is nearby a Place defined in DBpedia. To test the approach, a CONSTRUCT query was written (Listing 2) which takes all places from DBpedia in the Republic of Ireland and maps these to GeoSPARQL. This query is run on the DBpedia client (https://dbpedia.org/sparql) and results in only one RDF triple for the Board Gáis Energy Theatre in Ireland (which has a longitude and latitude value).

[10] https://sites.google.com/site/smartappliancesproject/ontologies/reference-ontology
[11] https://opengogs.adaptcentre.ie/debruync/r2rml/

```
@prefix rr: <http://www.w3.org/ns/r2rml#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix geos: <http://www.opengis.net/ont/geosparql#> .
@prefix ifcowl: <http://ifcowl.openbimstandards.org/IFC2X3_TC1#> .
@prefix osi: <http://ontologies.geohive.ie/osi#> .

<#Building>
  rr:logicalTable [ rr:sqlQuery """SELECT BUILDING_PNT.* FROM BUILD-
ING_PNT ORDER BY GUID""" ] ;

  rr:subjectMap [
    rr:template "http://data.geohive.ie/resource/building/{GUID}" ;
    rr:class geos:Feature ;
    rr:class osi:Building ;
    rr:class ifcowl:IfcBuilding ;
] ;
# rest of mapping omitted for brevity
```

Listing 1: R2RML Mapping

```
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX geos: <http://www.opengis.net/ont/geosparql#>
PREFIX dbo: <http://dbpedia.org/ontology/>
CONSTRUCT {
  ?s geos:hasGeometry [ geos:asWKT ?point ] .
}
WHERE {
  ?s a dbo:Place.
  ?s dbo:locationCountry <http://dbpedia.org/resource/Republic_of_Ireland>.
  ?s geo:lat ?la .
  ?s geo:long ?lo .
  bind(STRDT(concat("POINT(", ?la , " ", ?lo, ")"), geos:wktLiteral) as ?point)
  }
```

Listing 2: SPARQl Query to map DBpedia to GeoSPARQL

It should be noted, that an issue exists in Virtuoso (used by DBpedia) for the query in Listing 2 as Virtuoso does not fully support OpenGIS geometries, so any (valid) geometry is cast to virtrdf:Geometry on the fly. Therefore, it was necessary to adjust the resulting triple by hand. This is a known bug which will be addressed in future versions of Virtuoso. The resulting RDF triples are then loaded into a GeoSPARQL-enabled triplestore along with a subset of the OSi Prime database which can also be found on GeoHive (http://data.geohive.ie/dumps/totals-100m.ttl). Subsets of both OSi and DBpedia are used here to demonstrate the efficacy of the approach, in particular, to explore the geospatial functionality supported by the mapping process. As can be seen from the query, it returns all features (geos:hasGeometry) close to the DBpedia Place longitude and latitude, which we use explicitly here. Feature is used, and not specifically buildings, due to a current lack of geolocation data on buildings in Ireland available on DBpedia.

```
PREFIX geos: <http://www.opengis.net/ont/geosparql#>
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
SELECT * WHERE{
  ?loc geos:hasGeometry ?feature .
  ?feature geof:nearby(53.3442497253418 -6.240039825439453 2
<http://qudt.org/vocab/unit#Kilometer>).
}
```

Listing 3: SPARQL Query to locate all Features near by the Location of DBpedias Board Gáis Energy Theatre

Nonetheless, the use of geolocation here demonstrates that this approach can support linking of datasets describing features, such as buildings, based upon their geolocation. This can therefore be a powerful tool for supporting a basis for interlinking BIM datasets, based upon authoritative OSi geometry models. Figure 2 provide an example output using YASGUI (yasgui.org) which supports visualisation of GeoSPARQL geometries, along with some additional functionality to display labels. Our example displays data taken from both the OSi and DBpedia.

5 CONCLUSIONS

This paper presented the application of a methodology for uplifting authoritative non-RDF Ordnance Survey Ireland geometric building data to support iterative development of RDF based Building Information Models. It explored the creation of links between the OSi Prime2 data and an existing BIM standard, ifcOWL, to support interoperability with existing AEC processes. The geometric data is published using GeoSPARQL, which also supports additional geospatial functions within SPARQL queries. The goal is to provide a basis for sharing BIM data in the Irish context, built upon a foundation of authoritative geometric OSi data, along with additional building attributes, such as address, form and function. The addition of this geospatial building data would allow one to analyse, link, explore, and even build data analysis applications on top of several datasets using Semantic Web technologies. The feasibility of the approach is demonstrated through the integration of DBpedia data with the OSi data, as demonstrated in Figure 2. We chose these datasets due to the lack of available open BIM datasets in Ireland. We therefore believe it is of key importance to create this authoritative basis for developing integrated data on buildings, which can be linked too. OSi can provide just such a data hub.

Future work will explore linking the data with some sample IFC OWL models to further validate the approach. Also to be explored are the use of flat geometries to describe building floor plans (on a storey by storey basis), which can then be published and shared to support indoor navigation, control and energy management. Of particular interest also is the integration of more complex geometries, as GeoSPARQL currently does not support 3D geometries (see Figure 4, red circle).

Also, as not all OSi building data is currently open, and depends on a license, further work will also investigate the possibility to support queries for data, which return an indication that the data exists, alongside licensing information. This way, people can search available BIM and be made aware of its existence, which while not the ideal of "open", is preferable to the alternative, which is that it is not possible to be made aware of what data exists, without explicitly requesting this information from OSi. In other words, future work consists in investigating ways to manage access control to so-called "closed" Linked Data. Finally, the provenance aspects of the steps proposed in the method as additional metadata to improve the transparency, traceability and reproducibility of data enrichment exercises are being actively explored and will be applied also to manage the changing nature of building data in the OSi database.

ACKNOWLEDGMENTS

Within Ordnance Survey Ireland, we would like to thank Éamonn Clinton for his expertise. The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

REFERENCES

[1] M. Bilal, L. O. Oyedele, J. Qadir, K. Munir, S. O. Ajayi, O. O. Akinade, H. A. Owolabi, H. A. Alaka, and M. Pasha, "Big Data in the construction industry: A review of present status, opportunities, and future trends," *Adv. Eng. Informatics*, vol. 30, no. 3, pp. 500–521, Aug. 2016.

[2] C. Eastman, P. Teicholz, R. Sacks, and K. Liston, *BIM Handbook: A guide to Building Information Modeling for owners, managers, designers, engineers and contractors.* 2011.

[3] P. Pauwels, S. Zhang, and Y.-C. Lee, "Semantic web technologies in AEC industry: A literature overview," *Autom. Constr.*, vol. 73, pp. 145–165, Jan. 2017.

[4] R. Volk, J. Stengel, and F. Schultmann, "Building Information Modeling (BIM) for existing buildings — Literature review and future needs," *Autom. Constr.*, vol. 38, pp. 109–127, Mar. 2014.

[5] K. Afsari, C. E. and D. Shelden, K. Afsari, and C. E. and D. Shelden, "Cloud-based BIM Data Transmission: Current Status and Challenges," *ISARC Proc.*, pp. 1073–1080, 2016.

[6] A. Ghaffarianhoseini, J. Tookey, A. Ghaffarianhoseini, N. Naismith, S. Azhar, O. Efimova, and K. Raahemifar, "Building Information Modelling (BIM) uptake: Clear benefits, understanding its implementation, risks and challenges," *Renew. Sustain. Energy Rev.*, Dec. 2016.

[7] C. Debruyne, É. Clinton, L. McNerney, and A. Nautiyal, "Serving Ireland's Geospatial Information as Linked Data," *ceur-ws.org.*

[8] M. Perry and J. Herring, "OGC GeoSPARQL-A geographic query language for RDF data," *OGC Implement. Stand. ref OGC*, 2012.

[9] T. Lebo, S. Sahoo, D. McGuinness, and K. Belhajjame, "PROV-O: the prov ontology. w3c recommendation, 30 April 2013," *World Wide Web*, 2013.

[10] B. McAuley, A. Hore, and R. West, "BIM Innovation Capability Programme - Irish BIM Study," *Articles*, 2016.

[11] A. Galiano Garrigós, T. Kouider, and J. Scott, "The Big BIM battle: BIM adoption in the UK for large and small companies," in *Healthy Buildings: Innovation, Design & Technology*, 2016, pp. 53–66.

[12] B. McAuley, A. Hore, and R. West, "BICP Global BIM Study - Lessons for Ireland's BIM Programme," Jan. 2017.

[13] buildingSMART, "BuildingSMART Home Page," 2014. [Online]. Available: http://www.buildingsmart.org/.

[14] R. Howard and B.-C. Björk, "Building information modelling – Experts' views on standardisation and industry deployment," *Adv. Eng. Informatics*, vol. 22, no. 2, pp. 271–280, Apr. 2008.

[15] McGraw Hill Construction, *Smart Market Report: The Business Value of BIM in North America: Multi-Year Trend Analysis and User Ratings (2007-2012).* 2012.

[16] T. Heath and C. Bizer, "Linked Data: Evolving the Web into a Global Data Space," *Synth. Lect. Semant. Web Theory Technol.*, vol. 1, no. 1, pp. 1–136, Feb. 2011.

[17] W. Terkaj and A. Šojić, "Ontology-based representation of IFC EXPRESS rules: An enhancement of the ifcOWL ontology," *Autom. Constr.*, vol. 57, pp. 188–201, Sep. 2015.

[18] "ifcOWL ontology file added for IFC4_ADD1 | Linked Building Data Community Group." [Online]. Available: https://www.w3.org/community/lbd/2014/12/12/ifcowl-ontology-file-added-for-ifc4_add1/. [Accessed: 31-Aug-2015].

[19] C. Zhang and J. Beetz, "Querying linked building data using SPARQL with functional extensions," 2016.

[20] L. Daniele, F. den Hartog, and J. Roes, "Created in close interaction with the industry: the smart appliances reference (SAREF) ontology," *Int. Work. Form.*, 2015.

[21] Smartm, "TR 103 411 - V1.1.1 - SmartM2M; Smart Appliances; SAREF extension investigation," 2017.

[22] "FIEMSER ontology - Smart Appliances Project." [Online]. Available: https://sites.google.com/site/smartappliancesproject/ontologies/fiemser-ontology. [Accessed: 22-May-2017].

[23] D. Bonino and F. Corno, "DogOnt - Ontology Modeling for Intelligent Domotic Environments," Springer, Berlin, Heidelberg, 2008, pp. 790–803.

[24] J. Swetina, G. Lu, P. Jacobs, F. Ennesser, and J. Song, "Toward a standardized common M2M service layer platform: Introduction to oneM2M," *IEEE Wirel. Commun.*, vol. 21, no. 3, pp. 20–26, Jun. 2014.

[25] M. Compton, P. Barnaghi, and L. Bermudez, "The SSN ontology of the W3C semantic sensor network incubator group," *Web Semant. Sci.*, 2012.

[26] D. Ioannidis, D. Tzovaras, and C. Malavazos, "Occupancy and Business Modelling," 2012.

[27] H. Wicaksono, P. Dobreva, P. Häfner, and S. Rogalski, *Methodology to Develop Ontological Building Information Model for Energy Management System in Building Operational Phase*, vol. 454. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015.

[28] H. Wicaksono, K. Aleksandrov, and S. Rogalski, *An intelligent system for improving energy efficiency in building using ontology and building automation systems.* 2012.

[29] M. Kofler and W. Kastner, "Towards an ontology representing building physics parameters for increased energy efficiency in smart home operation," *Proc. 2nd Cent. Eur.*, 2013.

[30] ISI, "A direct translation of the CityGML XMLSchema to OWL." [Online]. Available: http://cui.unige.ch/isi/onto/. [Accessed: 22-May-2017].

[31] E. P. Karan and J. Irizarry, "Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services," *Autom. Constr.*, vol. 53, pp. 1–12, 2015.

[32] S. Donkers, H. Ledoux, J. Zhao, and J. Stoter, "Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings," *Trans. GIS*, 2015.

[33] Y. Deng, J. C. P. Cheng, and C. Anumba, "Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison," *Autom. Constr.*, vol. 67, pp. 1–21, 2016.

[34] L. Zhu and H. Mao, "Building Interior Model Interpolation between IFC and CityGML," *Open Constr. Build. Technol.*, 2015.

[35] S. Auer, J. Lehmann, and S. Hellmann, "Linkedgeodata: Adding a spatial dimension to the web of data," *Semant. Web-ISWC 2009*, 2009.

[36] Y. Bai, "Integrating GIS and BIM for community building energy design," 2017.

[37] H. Halpin, P. Hayes, J. McCusker, and D. Mcguinness, "When owl: sameas isn't the same: An analysis of identity in linked data," *Semant. Web–*, 2010.

[38] T. Kauppinen, R. Henriksson, R. Sinkkilä, and R. Lindroos, "Ontology-based Disambiguation of Spatiotemporal Locations.," *IRSW*, 2008.

[39] C. Debruyne, K. McGlinn, L. McNerney, and D. O'Sullivan, "A lightweight approach to explore, enrich and use data with a geospatial dimension with semantic web technologies," in *Proceedings of the Fourth International ACM Workshop on Managing and Mining Enriched Geo-Spatial Data - GeoRich '17*, 2017, pp. 1–6.

[40] P. Pauwels, T. Krijnen, W. Terkaj, and J. Beetz, "Enhancing the ifcOWL ontology with an alternative representation for geometric data," *Autom. Constr.*, vol. 80, pp. 77–94, 2017.

[41] Christophe Debruyne, Alan Meehan, Éamonn Clinton, Lorraine McNerney, Atul Nautiyal, Peter Lavin and Declan O'Sullivan. Ireland's Authoritative Geospatial Linked Data. In ISWC 2017, 2017