A framework for integrating syntax, semantics and pragmatics for computer-aided professional practice: With application of costing in construction industry

Abstract

Producing a bill of quantity is a knowledge-based, dynamic and collaborative process, and evolves with variances and current evidence. However, within the context of information system practice in BIM, knowledge of cost estimation has not been represented, nor has it been integrated into the processes based on BIM.

This paper intends to establish an innovative means of taking data from the BIM linked to a project, and using it to create the necessary items for a bill of quantity that will enable cost estimation to be undertaken for the project. Our framework is founded upon the belief that three components are necessary to gain a full awareness of the domain which is being computerised; the information type which is to be assessed for compatibility (syntax), the definition for the pricing domain (semantics), and the precise implementation environment for the standards being taken into account (pragmatics). In order to achieve this, a prototype is created that allows a cost item for the bill of quantity to be spontaneously generated, by means of the semantic web ontology and a forward chain algorithm. Within this paper, 'cost items' signify the elements included in a bill of quantity, including details of their description, quantity and price. As a means of authenticating the process being developed, the authors of this work effectively implemented it in the production of cost items. In addition, the items created were contrasted with those produced by specialists. For this reason, this innovative framework introduces the possibility of a new means of applying semantic web ontology and forward chain algorithm to construction professional practice resulting in automatic cost estimation. These key outcomes demonstrate that, decoupling the professional practice into three key components of syntax, semantics and pragmatics can provide tangible benefits to domain user.

Keywords: rule-based, semantic, ontology, IFC, BIM, cost estimation, computer-aided professional practice, expert system

1 Introduction

The ability to produce a standardised bill of quantities is a key issue for those undertaking cost estimation activities. Although the recent introduction of BIM can assist the process of cost estimation, difficulties remain where specialist knowledge, for instance measuring quantities that comply with standards and apply subjective unit price from expert experiences, is required to input the original quantities. Cheng et al. [1] state that this is also greatly impacted by the subjective decision making practices of estimators. While the use and accessibility of previous data is of great value, it is again subjective and regardless of the complexity of the tool, basic spreadsheet or cost modelling software, being used a degree of specialist input is needed. In fact, this process is used extensively as a means of overcoming the issues surrounding successful cost estimation [2,3].

This paper will focus on the key limitations relating to the production of BIM models created for cost estimation purposes. It is necessary to employ a broad and open-minded point of view as many factors must be acknowledged including current standards of practice, issues surrounding compatibility and those affecting subjective decisions. The detailed elements relating to standards of practice and expert knowledge lead to the necessity for an elaborate software development process. Specifically, this is required in order to tackle the complexities of sharing the requirements of meeting industry standards with software developers, in addition to confirming the functionality of the software created. The resulting outcome of this particular model is that it does not generally tackle details of pricing systems.

The overall aim of this study is to construct, review and confirm a common rule-based semantic specialist cost estimating process. Those undertaking the work believe two key outcomes will emerge from its development, specifically: the ability of domain experts to understand and improve the standards of practice and degree of awareness incorporated within open software architecture; the growth in understanding that methodical cost estimating closely corresponds with rule-based analysis and the framework to incorporate expert reasoning into information system. The outcome of this is the ability to substantiate the approach with increased levels of precision.

Subsequent to this introductory section, the background of cost estimation will be explored, reviewing the issues existing in this area as well as the previous work into potential intelligent solutions that has been carried out. Following this, section three will focus on the foundation of the framework and its associated method, while the next section will present the proposed framework and its elements. Then a case study detailing the results generated will be presented. Finally, the results will be examined and future developments suggested.

2 Backgrounds

The process of cost estimation incorporates a great deal of subjectivity and specialist input is generally required whether a basic spreadsheet or intricate cost modelling system is used. An estimate is a specialist's view of an expected future cost. While expert problem solving is not a specific cost estimation procedure, it is well understood and often used [2,3]. Specialists in the field generally reach estimated figures through the use of analogies and comparisons [4–10]. Studies undertaken by Sinclair et al. [11] reached the conclusion that existing methods of reasoned application were overly simplified throughout the construction industry. These methods regularly use existing items (with known specifics) to draw comparisons with new items (with unknown specifics). Once the details have been allocated to the item it then moves from being new to existing [12]. The National Aeronautics and Space Administration (NASA) utilises a valuable cost estimation model that is similar to this, which incorporates clear processes for undertaking estimation. This style of process does, however, provide only an abstract portrayal of estimation requirements. This is particularly clear when considering the Eurostat survey method, which cannot be analysed by computers. Quantity surveying, in relation to project cost estimation, requires the completion of systematic or analogical tasks. They undertake these processes, using historical data, to build a cost estimate. In general, cost estimation activities require the detection of an analogical link between the project in question and the previous work, associating the specifics of the project with their equivalent points, leading to a clear result. Mair et al. [13] state that an electronic form of Case-based reasoning (CBR) cycle allows the development of an altered and approved outcome. As a result, as highlighted by RSMeans [14] and The Building Cost Information Service (BCIS) [15] problems are solved by linking the outcomes of previous work to current challenges.

On the other hand, there has been very little investigation into the cognitive reasoning that a specialist puts in place when reaching a decision or the ways in which this can be linked to the activities of cost estimating. Kiziltas et al. [16] argue that these factors are so often not documented that the details are not available to other specialists. As a result, very few studies have been carried out into the process or attempts made to summarise the rationale behind it. Research in the construction engineering field often focuses on algorithmic features because organisations are generally more at ease with a statistical approach because organizations concern individual subjective input as a 'black box' and hasn't been captured in the system. [17–19]. Hughes [20] explains that researchers are not content with such an undesirable attitude towards specialist knowledge, although an effective means of reinforcing this has not yet been established [16,21,22].

Attempts have previously been made towards automating the cost estimation process within BIM, and for this reason a broad investigation into intelligent solutions for cost estimation is required [23–27]. It should also be noted that the reports discussed have been selected from a literature review of over 100 sources. These were identified using keywords such as semantics model, construction, design, building, built-environment, ontology, resource description framework (RDF), semantic web ontology language (OWL), and IFC (dated between 2002 and 2015).

The process of identifying the most appropriate work related items for expense estimation can be effectively carried out using OWL and SWRL [27]. The data collected is extracted from an IfcXML file, to which semantic analysis is performed to produce a range of working conditions. This class of file allows data regarding measurements, methodology and materials to be gathered, thereby offering a clear indication of the items required for expense estimations to be made.

Technological advances in information management in the construction industry tend to be heavily dependent on the functionality of the latest internet release [28]. This is often referred to as knowledge demonstration and allows for the possibility of major steps forward in innovative processes. A key feature of this approach is that it must allow for interrelated functionality between different software products, not excluding web-based and intelligent offerings. When focusing on the construction industry the software products that they must work with are those featuring the ISO-10303 Standard Exchange of Product Data (STEP). Moreover, the IFC was created, using the specific elements of STEP, as a building information model to be used in the construction industry. The semantic heavy model has led to the improvement of building information through the use of semantic technologies. Overall, the main aim is to develop the current availability of building information through the production of data that can be electronically processed.

OWL is often regarded as offering the best expertise representation language, specifically from a visual point of view, due to it being effective, well-known and widely supported. It should be highlighted that this study does not take into account offerings available in other languages, although it remains the case that information relating to engineering processes does not include a wide range of semantics and pragmatics [29].

It is vital that the model produced has the ability to build a degree of understanding while supporting individual input. For this reason, the usefulness of the model is largely focused on semantics, which allow for precision in the results produced and functionality of the model. Where they are unclear, the outputs are open to interpretation and unfortunately, the building models available tend to lack adequate semantic detail. An example of such can be seen on traditional CAD programmes, which provide sophisticated drawing options but lack a human perspective (e.g. simple geometric shapes). In other words, most traditional CAD drawings rely on user interpretation of constructed primitive geometric shapes and are not semantically marked-up with relationships and labels. This is a result of an absence of semantic details. In fact, the principle behind BIM was to construct models that make use of object-based images, to enable an environment to be detailed semantically. Furthermore, BIM offers the inclusion of an information source that construction industry personnel can access to locate specific product information. BIM has been largely supported by those in the construction industry since its introduction as a means of tackling many long-standing concerns.

While OWL is the most commonly used form of knowledge representation language within built-environment, analysis of estimations activity highlights the use of four forms of semantic model [29]. Grzybek et al. [29] explain that issues arise due to an overall incapacity of the general knowledge engineering process and suggest a guideline of ontology development. We believe that there is an overall lack of association between semantics and pragmatics within languages. Furthermore, based on our interpretation, the World Wide Web Consortium (W3C) has addressed this issue in the provenance (PROV) Family of Documents as entity, artefact and process respectively [30]. Additionally, the use of OWL as a knowledge representation language for domain knowledge is common as it is the most well-used and known. It should be noted, however, that its use does not take into account any concerns about capacity with regard to languages. For this reason, a more extensive analysis of the outcomes of the estimating process is needed[24,27,31,32], see table 1 below:

.

Table 1 Comparison between Estimation Methods

Estimation method	**General Cost Estimation Process**	
	Cost Item Quantification	Unit Price Determine
Conventional Cost Estimation	Manual	Manual
BIM based Cost Estimation	Automatic*	Automatic*
Ontology Improved BIM and language used	1: describe cost item comply with standard; IFC-based ontology. 2: Identify working condition to select cost item; OWL	3: Identify project location to select labour cost; UML+OWL. 4: Identify construction condition to adjust unit price; ontology language not specified

Note: BIM based QTO and price identification does not comply with standard
1: [31]; 2: [27]; 3: [24]; 4: [32]

The findings of Ma, Wei and Zhang [31] explain the importance of maintaining a process that enables the transfer of data from IFC to OWL. This leads to the system being able to automatically catalogue the construction elements into expense items with associated costs, prior to producing a final cost summary. In this process, pricing details are not included. The reason for this, as explained by the researchers, is that the expense items have the ability to continually reflect market prices and for this reason item pricing is not required. On the other hand, OWL is used in the process of application while IFC is built upon the semantic process.

Lee et al. [27] explores how employing ontology classification capacity can derive the working method of cost item.

Abanda et al. [24] investigated the expenses associated with work activities in Cameroon. There it is common practice to use simple data engineering processes to build UML models for work expenditure, which are then inputted to OWL. Following this, semantic web rule language (SWRL) for OWL is applied, which results in a range of job locations being identified and allowing expenditure estimations to be made for each location.

The research findings of Staub et al. [32] have previously identified a range of building features that enable experts to make precise pricing estimations. The method of application involves establishing the features of the particular element, using examples of building product models (for example IFC models), identifying the exact specifications of the element and relating the knowledge of experts to this in a project specific way. In doing this, it is possible to assess or incorporate requirements for each element of the building and leads to expenses being altered accordingly. Staub-French et al. [23] also provide details of an ontology of building elements that assist specialists in the cost estimating process. This is executed by producing details of the particular element from images captured in building product models, such as IFC- based outputs, noting the specific design details of the element and applying specialist knowledge specific to the project as a means of evaluating or including certain building features. Subsequently, the costs associated with the project are amended dependent on these outcomes.

Within this analysis of the background, a number of factors relating to BIM-based cost estimation have been investigated. The potential for updated and advanced processes has been suggested, specifically relating to the ontology employed in the construction sector. To date, no studies have been undertaken that have successfully combined the relevant elements within one model. Although many of the key studies in this area note the importance of a semantic basis within the domain [33,34], it is understood that to provide a summary of all the necessary opinions a structure based on information is required, focusing on compatibility (syntax), a description of pricing analysis (semantics), and exact details of the circumstances surrounding the elements being assessed (pragmatics). Within the construction industry, this is particularly important as there is a growing trend towards reverse engineering of the semantics relating to the domain from the key information source, the IFCs [35]. There are however concerns regarding the suitability of this method, as IFCs have been produced for the purpose of storing information and therefore are not able to completely deliver the correct range of semantics required by construction management professionals. The assimilation of these three aspects is the most important advantage offered by our approach. It is hoped that it will continue to expand the reasoning capacity of the pricing models used alongside BIM, particularly in comparison to similar models outlined in this chapter, and that it will serve as an advanced progression. In our opinion, this new approach offers plenty of benefits to specialists who use BIM.

3 Semiotic Framework for an Integrated Solution

The study of signs, called semiotics, was independently developed by the logician and philosopher Charles Sanders Peirce and the linguist Ferdinand de Saussure. Stamper [36] furthered the field of semiotics with the inclusion of a philosophical, radical subjectivist stance, that redefined information as signs concerned with an individual's cognitive process. Stamper [37] develops a semiotics framework for information system analysis and design with the use of three additional layers in conjunction with the traditional three divisions of semiotics, namely; the physical, empirical, and social layers (see Figure 1).

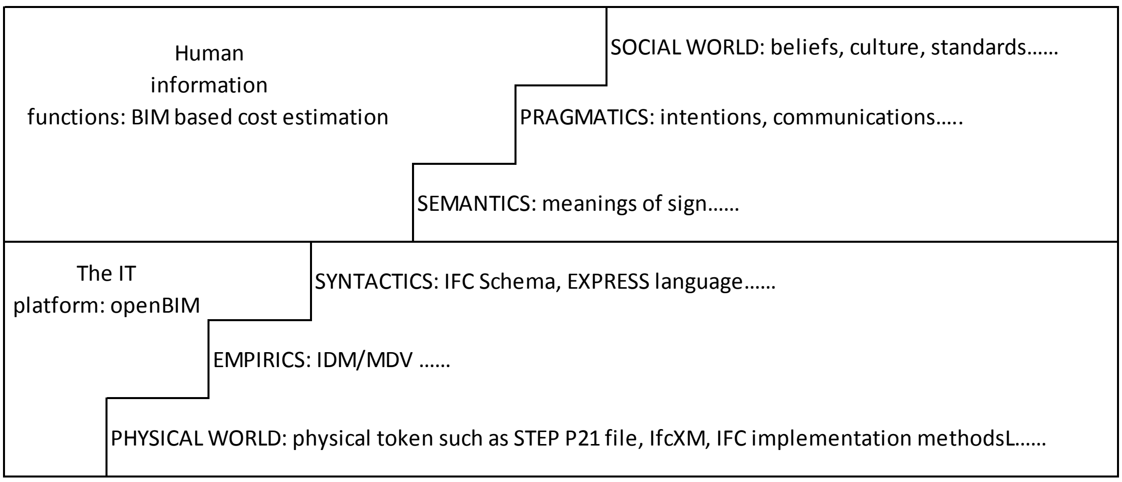


Figure 1 Semiotic Framework for BIM based Cost Estimation (adapted from Liu (2000))

Within the context of this research, there can be said to be two main 'pillars': human informational functions (the upper three layers), which refer to BIM-based construction cost estimation, and which are mainly concerned with signs; what signs are, and how the signs are to function within the communication as well as the IT platform. The IT platform (which forms the lower three layers) refers to open-BIM, which answers questions regarding the structuring of signs, the means by which signs are organized as well as the physical properties a sign has. More specifically, each layers has been depicted as follows:

3.1 Six Perspectives of BIM-based Cost Estimation

Cost estimation is a service provided by professional organisations. Previously, this kind of service was considered to be an approach intended to represent an objective reality regardless of its nature of subjective and human intervention. Today's building projects are more complicated than ever before, and are executed on a much larger scale. However, today's cost estimation methods fail to satisfy users' requirements since they tend to be technically sound but cumbersome in practice. Researchers have found that where cost estimation software has not yet been successful, this is down to a lack of knowledge about the cost estimation field. As mentioned earlier in this paper, those within the industry must adopt a fresh strategy that emphasises the human factor and its role in cost estimation. It has been established that whilst the IFC is the most commonly-adopted standard of field-specific language within the industry, it is not capable of representing one of the crucial elements of cost estimation: the complicated nature of organisational behaviour. For this reason, it is suggested that this topic is tackled through the application of normative elements. Whilst it is widely acknowledged that there is a semantic flaw between cost estimation software and organisation, those within the industry are still expected to find a way to deal with the costing process of unclear boundaries. As such, Staub et al. [39] argue that estimators will therefore be required to adopt this role based on their own conscious or subconscious inclinations. This can be problematic if the decision made by the estimator is not the best strategic decision for the organisation.

Physical Level: A sign studied at the physical level focuses upon the physical properties of that sign, including the sign's shape, size, source and destination—among other criteria—depending on the type of sign. From the physical perspective of databases, there are collections of physical tokens which can be stored and moved around for input, output or display [38]. Within the context of this research, models of BIM at physical level describe the range of physical tokens that may be available from physical components. Those physical tokens may refer to STEP P21 files and IfcXML files [40,41].

Empirics: Empirics studies the properties of signs, based on a collection of signals or marks. The questions regarding a stream of signals from the sending end to the receiving end are investigated in the study at this level, irrespective of meaning and any problems encountered at the physical level; such as transmission, reception, coding and decoding, and channel capacity, among others. Within BIM, more specifically, IFC-based BIM studies, this level is to focus more upon the efficiency of utilising BIM, which corresponds to one core pillar of building SMART technologies, namely; IDM/MDV.

Syntactics: Syntactics studies the presentation of meanings based on reliable signal encoding methods to organise a collection of physical tokens at both the empirical and physical levels. Syntactics is primarily concerned with the composition of more complex signs from simpler ones. For example, an IfcEntity can be seen as a simple sign, while a building model is more complex, which is a combination of IfcEntities following language structure. Indeed, syntactics remain important when distinguishing information from data as information may be syntactically understood while data is not necessarily always understandable through such means.

Semantics: Semantics studies the meaning behind signs on the basis of a particular language structure understandable to the sign receiver. The semantics of a sign connects the sign itself to the entity, object or concept that that sign represents— a sign denotes a denotatum. Between the objective truth and the reality within any individual mind there will always be a semantic gap; and people need to interpret a sign in order to comprehend the truth. A sign may potentially bridge the gap once it possesses some meaning as a result of being mapped onto objects within reality. However, within a social environment there can be said to be no unique reality—as even a consensus is temporal and may not last 'forever'. For instance, an IfcCostItem within the definition of IFC schema ensures that "An IfcCostItem can be used to represent the cost of goods and services, the execution of works by a process, lifecycle cost and more." [42]. From the perspective of a quantity surveyor, the use of the standardised measurement model version 7 (SMM7), is composed of a first division, second division, and a third division; according to specific standards. Meanwhile, by employing new rules of measurement (NRM) [43], the cost item is seen to be composed of building elements, construction products, product properties and working methods. Therefore, it is important to acknowledge that a sign's meaning is, in part, based upon the individual who interprets that sign. Consensus is established when necessary and is subsequently shared by a group of people as their reality, as it shall be questioned, criticised and modified over time.

Pragmatics: Pragmatics studies the use of signs that possess particular intentions and/or the study of the "purposeful use of signs" [38]. Pragmatics is concerned with the relationship between signs and an individual's behaviour. For the same purpose, behaviour may vary according to the actor's personal experience, value system, and expectations. This personal possession of knowledge and experience can be called "pragmatic information". A scenario in which two individuals are communicating with one another using their own pragmatic information can be said to be normative practice. This communication not only includes the delivery of meanings, but also the pragmatic information itself—the exchanging of values and expectations among others. However, the pragmatics of signs does not make sense unless it is studied with the support of semantics. Indeed, pragmatics can be seen as a communicative shell which contains a core; their meaning. Another important view is that the conflicts caused by communication within social group originate with the differences within pragmatic information, which may potentially render communication ineffective or inefficient. Indeed it may prove costly to improve communication to a satisfactory level.

Social Level: The Social Level studies changes that occur within the real world due to the effect of communication. These changes may be seen in culture, custom, welfare, knowledge, attitudes or behaviours (among other facets). Within human societies, change is usually made under the governance of a particular community, such as a team, a company, an industry, or a nation. These communities accept and establish a set of social norms or conventions that govern members' thinking and behaviour. Consequently, communication too should be governed, and governed in accordance with a particular pattern. Thus, communication and change interact with one other to promote the development of the social world.

Although Stamper and Liu [44] include all six layers within their communication framework, the focus of organizational semiotics still rests upon the three upper layers, namely, semantics, and pragmatics, as well as the social world level that are to study the human information functions and its relationship with IT platforms. The lower three layers are to be studied initially—in order to illustrate the IT functions—after which the upper three layers will be studied. It is interesting to note that some researchers have started to consider the IFC model as a semantically 'rich' model

[45–47], however, others claim that the IFC is, in fact, insufficiently semantic [48–50]. This conflict will be addressed and considered as part of the framework of this study. IFC as an object-oriented model language is capable of carrying this information for all stakeholders within construction, however, a purely technological perspective is insufficient to bring meaningful changes to the industry as a whole. More specifically, it should consider existing working practices and business processes if it is to be successfully implemented [51].

The IFC does not have the formal specifications related to the various entities and relationships regarding reasoning mechanisms from the perspective of the organisation. As previously expounded upon and discussed; without an integrated process of cost estimation and a systematic framework—at both the semantic and pragmatic levels—patterns of behaviour are not fully recognised or defined. With regard to the new framework, introduced to BIM-based cost estimation, IFC lacks semantic, pragmatic and social form of organisational perspectives of which there are two aspects of knowledge. These are 'know-what' and 'know-how' [52].

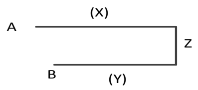
3.2 Semantic analysis

Together with the development of information systems, this research adopt semiotics that try to comprehend cost estimation from an informational, systematic perspective. By capturing this knowledge, both reasoning and inference steps can be used to improve the development of the information system. This research incorporates a suite of semiotics tools: semantic analysis (SA, produce ontology chart) and norm analysis (NA) [38]. This permits the identification of conceptual interests along with the ontological dependency between semantic units, and having established the ontology chart, it will then proceed to ascertain the semantic relations—hence formalising these relationships to model the behaviour of informational systems design [53].

The SA is the process of conceptualising a business organisation. It is the catalyst through which the behaviour of the organisation is analysed, and captured within the ontology model [38]. As previously outlined, it was initially found that the logic and inference steps can only be understood by experts, which means that their behaviour patterns are not recognised. Therefore, SA is suitable to perform this task. The semantic primitives which appear in the model represent the possible patterns of actions of a complex agent, or 'affordances' as they are known, within the NORMA model, deriving from NORM and Affordance and devised as a language for specifying norms and affordances as systems analysis and requirement specification, see table 1 NORMA syntax .

Table 2 Table of NORMA Syntax

Rectangle	___ employs ___	Relationship
Ellipse	nation	Concept
@	_ @ →	Knowledge denote
Ax	A—x	x is an affordance of A
Ax.y	A—x—•— y	y is a part of x; they are all afforded by A
Ax#y	A—x—#y	A affords x which has a determiner y
A((a:b:c:d)→ f	A— f a b c d	a,b,c,d are specifics of f

(A#x,B#y)z		A with role name x, and B with role name y jointly afford z

4 A conceptualized framework for computer-aided professional practice in construction industry

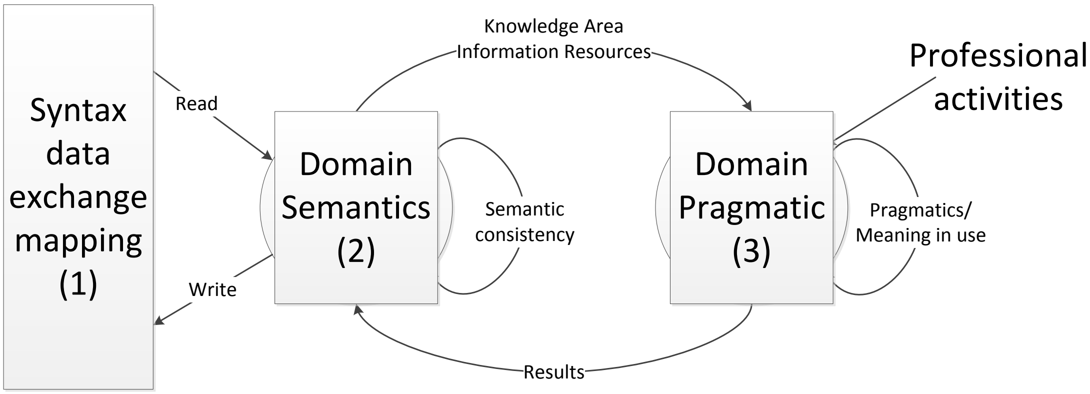


Figure 2 A framework for costing professional services under BIM environment (modified from Xu et al [52])

Semantic information is at the heart of the planned framework for providing automated cost estimating data, supported by additional detail regarding the work requirements of a traditional quantity surveying practice. The primary aim associated with the creation of this framework is that the output from problem solving specialists should be closely related to those concerned with the domain.

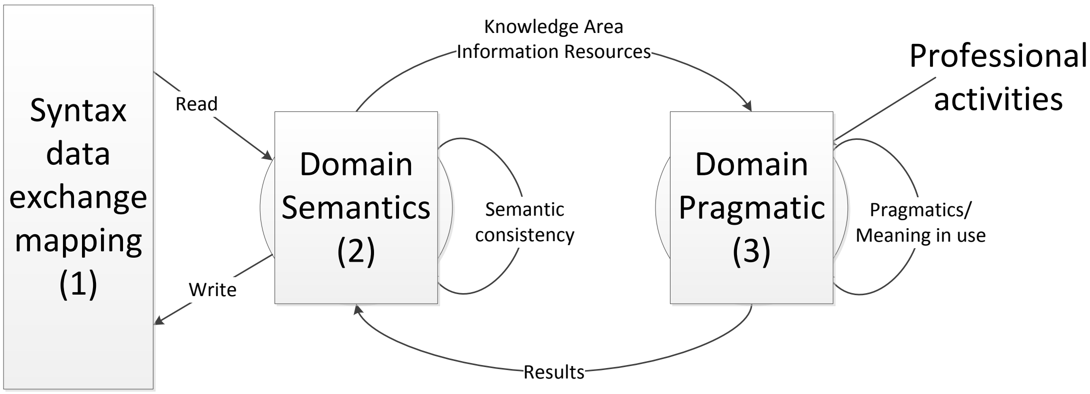


Figure 2 represents the three key areas of development within the framework construction, explicitly how those focused aspects being connected thus can provide tangible benefits to information system development: (1) the data structure being mapped based on entity in domain semantics; (2) the domain semantics result in entities captured with explicit knowledge input; (3) the domain pragmatic result in processes captured.

Syntax results in semantic data exchange: This element is largely concerned with the device that uses semantic mapping to access the data file, through the use of semantic analysis. It identifies links between the categories in the data format and the populated semantic fields. Additionally, the connection between each of the elements needs to be recorded. Traditionally, the intricacy of this process has led to the information required being entered using a routine method. The outcome of this is that the BIM coordinator is required to produce a set of process driven activities as well as the semantic mapping.

Semantic of professional practice concepts: The cost details on which this system is based are built around a set of interconnected ontologies that are either prepared in advance or created within the knowledge engineering process. These key details are classified using semantic analysis (SA), a systematic system engineering process. They are then developed further using pre-constructed terminology box (T-Box) ontologies, which represents the schema or taxonomy of the domain at hand, such as, core domain ontology, code of practice ontology, specialist knowledge ontology, or database ontology. The advantages associated with this method are explored in the following section of this paper.

Pragmatic activities with reasoning engine: This element is associated with the delivery of guidelines for the specialist problem solving process. As previously stated, the key concept of this framework is its ability to connect the three approaches, in contrast with models currently available, making use of semantic web ontology such as SWRL. Issues can arise however when incorporating SWRL in that it can become uncontrollable as the system grows and new techniques are employed. For this reason, this framework will incorporate systems out-with those of logical programming. It also includes variable problem solving methods that will be carried out by a specialist in order to obtain the required outcome. There is no traditional clarification available for the potential outcomes of these processes. As previously mentioned, the researchers believe that successful cost estimation can only occur using a framework incorporating three main elements: (1) a comprehension of the data format (syntax); (2) the significance of the costing domain (semantics); and (3) the exact application parameters of the particular items (pragmatics).

In addition to the above the intended framework offers a diverse method at a conceptual level and provides direction as to how the model can be put to use within a BIM environment. The framework demonstrates that the IFC data formation is analysed for available building data and supports the delivery of cost estimations. An evaluation of the semiotics framework based on the IFC standards [54,55] confirmed this, along with the cost estimation domain standards [56], providing details of a semantic based model for building information that does not support specialist input [49]. For this reason, IFC holds a place as the key source of information for construction projects, while domain ontology has been introduced as a provision for specialist knowledge.

The proposed framework will employ a number of software items:

- A user interface that enables specialists to extract semantic information from the available code of practice papers and company information through the use of an ontology chart, leading to the creation of cost estimation ontology.
- A first order logic rule creator that will generate rules from the metadata provided in the detailed code of practice papers.
- A semantic translator that, using established costing processes, serves the purpose of translating the information gathered from a standard file and producing an assertions box (A-Box) ontology, which describes the attributes of instances (or individuals), the roles between instances, and other assertions

about instances regarding their class membership with the TBox concepts, from the semantics of the current ontology.

The subsequent section will provide further information regarding the methods used to implement the framework within the case study. Initially, the process of collating the costing information from the company and code of practice is explored. This is followed by an analysis of how this information is then recorded on a standard data model and detailed code of practice documents. Finally, the means by which the data gathered from these processes is used in the production of automated cost estimation.

5 Domain application and case study

As a means of authenticating this framework, a case study has been carried out within the area of cost estimation, incorporating the expertise of industry specialists, in order to detail how the framework achieves its main goals:

- Enabling the domain experts to map their organisational procedures based on the organisational structure and individual knowledge usage;
- Enabling data transfer through the use of the IFC open standard, while continuing to permit information from other sources to be mapped;
- Enabling domain specialists from the cost estimation sector to interrogate the outcomes using the processes of NRM;
- Demonstrating that the cost estimation created is precise in comparison to the manual version;
- Demonstrating how this process has assimilated cost estimation in comparison with previous studies.

The following section will initially provide details of the QS company and the project requiring cost estimation being undertaken. It will then explain the way the framework was utilised and the outcomes achieved.

5.1 Background

All the fundamental semantic elements required to complete construction project cost estimations are provided through semantic analysis and ontological charts. Intriguingly, the process of undertaking an investigative cost estimation is triggered by the client making a request for these details. The ontological chart used depicts the organisational structure of a standard project-centred construction project (as shown in Figure 3). This structure is key to the needs of this research as it provides the options necessary for multi-national project teams located in various sites. A project team can be created by a client organisation that hopes to build in a certain location and who hires a design company for the purposes of designing the job. On the other hand, it can be located in a different area, with the consulting organisation. These circumstances demand a variable transfer of information, regardless of the site, to allow for the physical transfer of files.

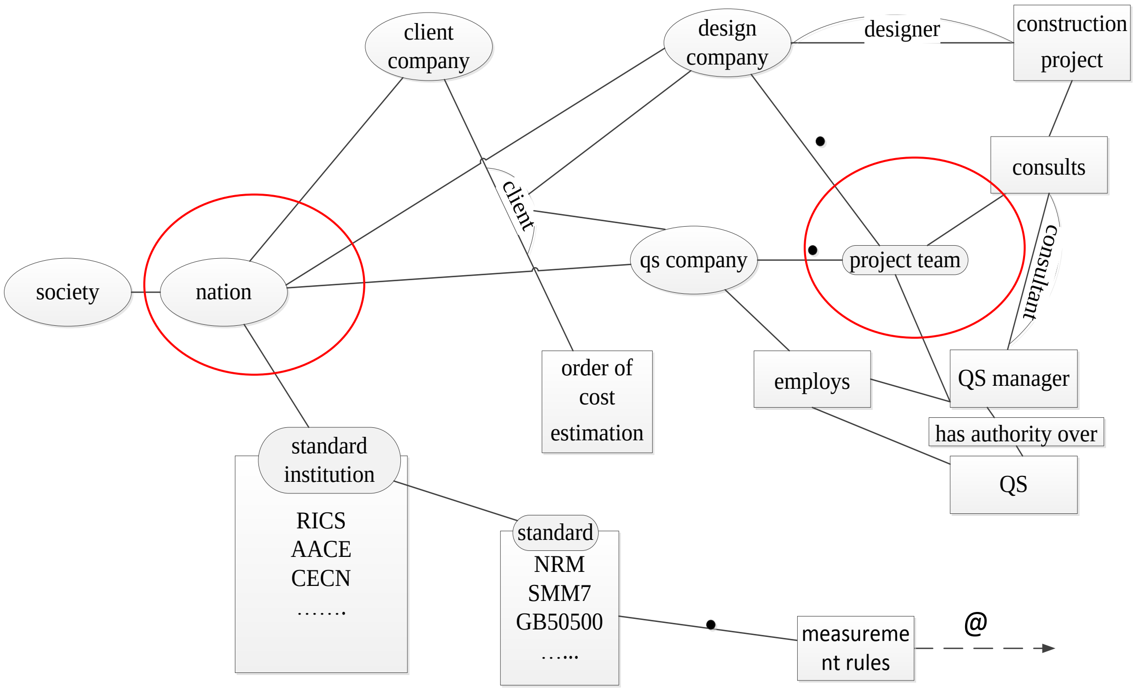


Figure 3 The Project Team based Structure of Studied Cost Consulting Company

The figure above displays the selection for codes of practice that are available. In this situation, the RICS new rules of measurement (NRM) will be utilised. This details clear direction on the requirements needed for the production of bills of quantities and quantified schedules of work for building works taking place. Additionally, it also offers details on the preparation of standard or bespoke schedules of rates. The RICS new rules of measurement: Bill of Quantities for Works Procurement also stipulates the requirements for measuring and defining building work and represents an example of good practice [43].

5.2 Detailing the framework with industry application

5.2.1 Syntax results in semantic data exchange

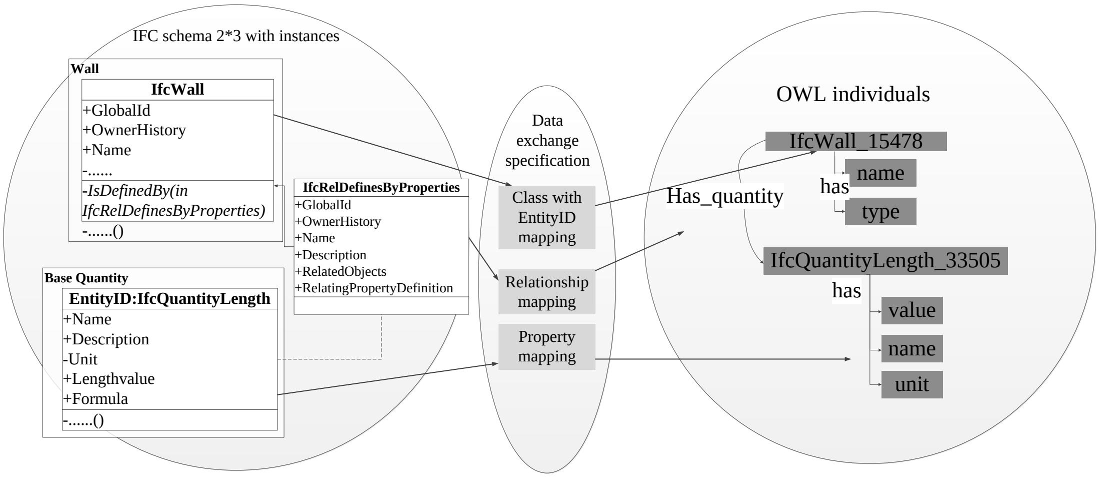


Figure 4 Data mapping engine

Following completion of the process of semantic analysis of the methods employed using metadata and the construction of the ontology, the subsequent task involves charting the semantics within the ontology into a specific format. This is undertaken in order to allow information to be transferred between the physical IFC document and OWL, and is assisted by the semantics detailed in the preceding phase. This step is carried out through the use of a data exchange engine that has been created.

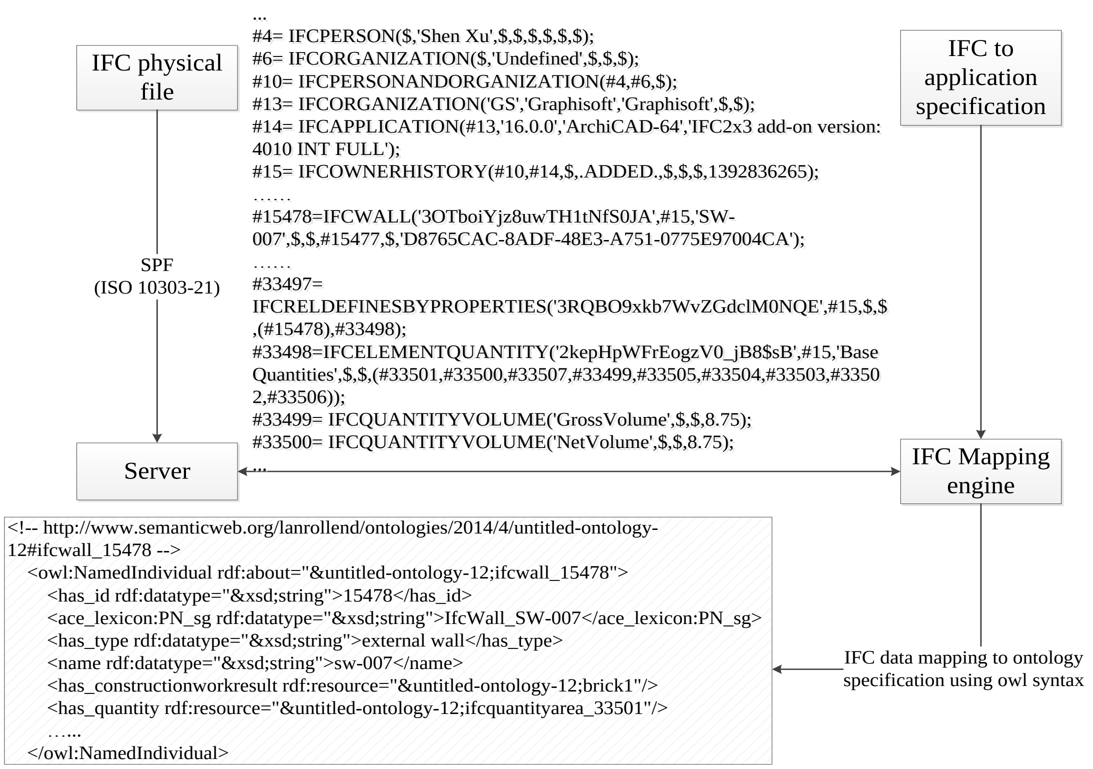


Figure 5 The architecture of data mapping engine

Figure 5 demonstrates the engine design that has been created for the planned framework. The engine functions on data sourced within IFC 2*3 documents [55], which is connected to the researchers belief that IFC is a suitable provider of building semantics [29,52,57,58]. For this reason, the narrative is largely concerned with the transfer of information and employing the features of the IFC instead of recording semantic elements between each category.

The nature of this engine is that it operates on specific items within a tangible document, an IFC physical file that comprises the following: ID, EN and EA. Under these circumstances ID provides an identity to the IFC item being considered, EN refers to the item's name and EA details the item specifics along with values. These elements are sorted and entered into the A-Box of OWL for use at a later date.

Figure 4 provides details of a transfer relating to the requirements of a wall item and a quantity item, which are then recorded under the relevant category for each item. Most commonly three specific activities are performed: recording the category; recording the property and recording the relationship. Recording the category requires the relationship that exists between a category in the OWL and an item with a certain ID within the data file to be detailed. The purpose behind the inclusion of the item ID is to address the unique name condition of the OWL. By recording the property details, the one to one relationships that exists between the particular characteristics of the item and the data held within the OWL are identified.

Additionally, figure 5 demonstrates the way in which relationships are recorded. The relationships have been clarified using semantic analysis, as they are ambiguous when presented in data format. At present, the recordings display the relationships within the data file in terms of their inverse attributes, connection and property details.

It requires a mapping mechanism which can convert the IFC physical file into Web Ontology, see figure 5. Briefly, it will employ specification and mapping engine. There are multiple choices of mapping engine implementation, for instance IFC API, IFC DLL and IFC Open Shell etc. [59]. Discussion of this mechanism implementation is outside the scope of this research.

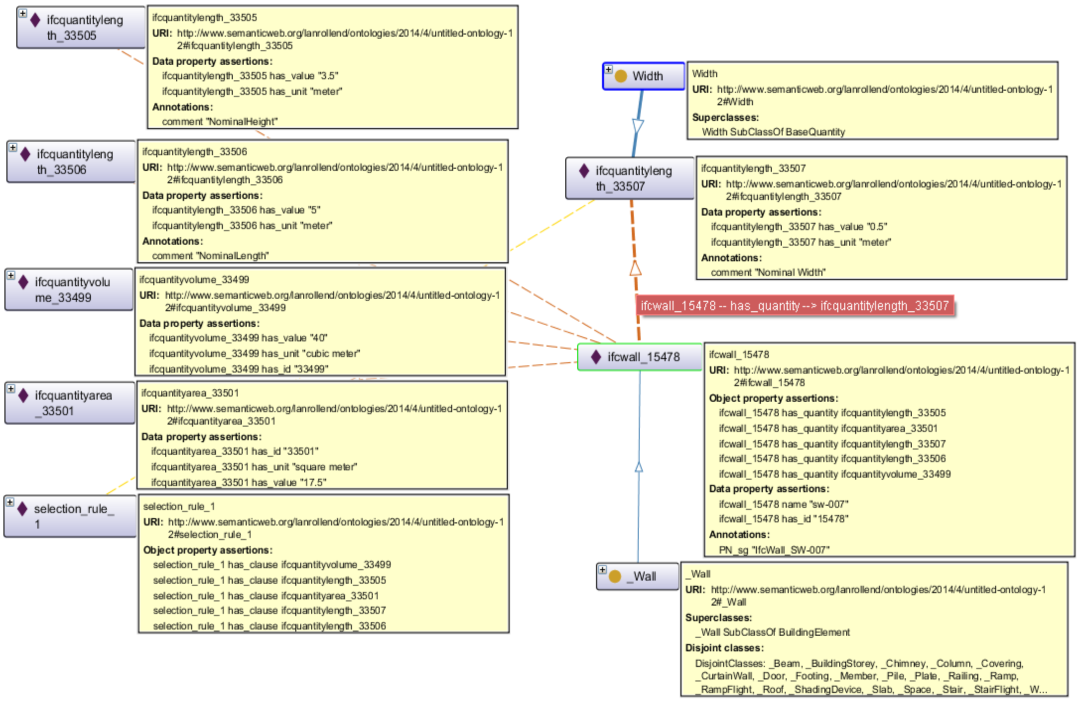


Figure 6 Example of Analytical Cost Estimation Ontology with IFC Instances

The instances from the IFC physical file have been organised according to the IFC triplets (ID, EN, EA); moreover, as is required by the unique name assumption in Protege, IFC ID is being inputted into the instances' name as shown in the Figure below. Figure 6 states that IFC instances: ifcwall_15478 *has_quantity* ifcquantitylength_33507 while ifcquantitylength_33507 has type Width. Furthermore, the data property assertions states that: ifcquantitylength_33507 *has_value* "0.5", *has_unit* "meter".

5.2.2 Semantic of professional practice concepts

This section will describe the overall process of extracting concepts, to elaborate the SA [38] method of semantic data from analysing organization, and how this method results in the population of our costing concepts. The SA method includes the problem definition, semantic units classification and ontological dependency checking. Ontological dependency can be defined as one item (y) existing only if another item (x) also exists. This provides an assurance of the abilities of the cost estimation process. Each of the concepts are linked through their ontological dependency, which is a required factor in Semantic Analysis. Liu [38] explains that there are two key purposes for selecting organisational semiotics. The first of these relates to issues surrounding the various ontology languages. This is due to a belief that a lack of knowledge associated with the process of cost estimation will lead to expertise being mislaid by the system as this is generally sustained by the individual.

The semiotic ladder provides complete details of syntax, semantic and pragmatic elements, as well as a social perspective. Stamper [60] explains that semantic analysis enables the incorporation of semantics and pragmatics into a fully functioning process. For this reason, the overall process can be dissected to provide assistance for specific activities, such as, the identification, selection, description and adjustment of information throughout the process of cost estimation. The second key reason for utilising organisational semiotics is the necessity for an automated approach to estimation resulting from the increasingly technology led environment in which the work is carried out. There is a requirement for human input within this system and the authors of this study strongly support the inclusion of compatible and process driven, investigative strategies that are used in information system engineering and other linked activities. It is of key importance to note that this does not promote the replacement of individuals with automated processes, rather than they can be brought together within an organisation.

Through the process of categorising semantic units, it is possible to expand the information gathered. As an example, the 'building element' highlighted through semantic analysis, is an incomplete knowledge that needs to be explored further. In order to do this it can be linked to prevailing ontologies, detailed as follows:

Core domain ontology: the core ontology is a high level concept of the main elements within a specific code of practice domain. As a rule, ontologies for a number of domains have already been established, which leads to the possibility of their inclusion within the framework. **Error! Reference source not found.** demonstrates the exploration of the building element within the core ontology.

Code of practice ontology: this is created as a means of obtaining rules from files. When the process is completed, the T-Box ontology generated will signify the semantics included in the code of practice and the link to the semantics within the core ontology. The outcome of this is the collation of a complete range of semantics demonstrating the novel approach that a specific selection of semantics can have upon the core semantics of the domain in question. These factors will be explored further in the following section.

Expert knowledge ontology: this signifies the semantics associated with the elements that guide construction costs and the process involved in defining the particular cost of each item. Within construction there are a number of variables, for example the specific measurements of an element [21,61–63], the formation of the element [39,61,63,64], climatic influences [65–69], congestions factors on site [62,68–70], equipment size [68,71], temporary requirements [68,70,71], and the make-up of the construction team [68,70], although the one study that entails every element of intelligent solutions is an item-based cost estimation model [72]. Within the proposed system an item-based ontology [32] was employed in order to provide greater detail within the specialist knowledge available, as shown in **Error! Reference source not found.**.

Database ontology: the database ontology charts the costing database semantics that are used in the transfer of information from the database to the model. The semantic

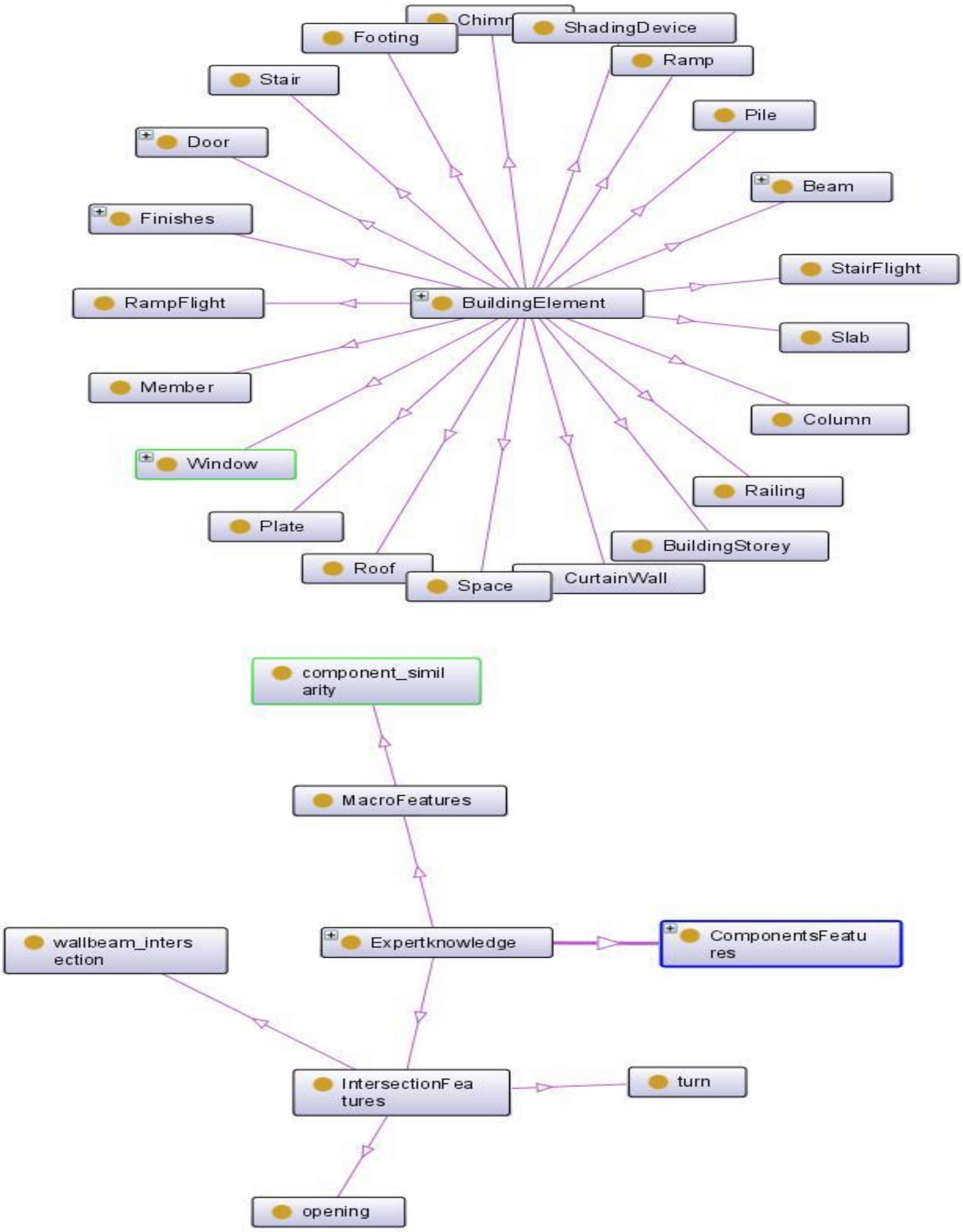


Figure 7 Further expansion of Incomplete Knowledge (building element and specialist knowledge)

details are gathered following the supplementary analysis of cost items, allowing the ontology to be generated with information delivered from the database.

5.2.3 Pragmatic activates with reasoning engine

Within this section the complete process associated with obtaining rules from current code of practice documents will be defined, explaining the NA [38] approach to extracting detail from documents, and how this was adapted to meet the authors' needs, leading to entry of information into their clauses. The method of obtaining quantifiable rules from the code of practice documents has been achieved by using the rules existing within the norm specification. To accomplish this task the NA approach, which was formerly used as a means of evaluating the information held within a company, was adapted. This approach was employed for three main purposes [73] namely: its ability to present clear information; the simplicity with which existing information can be amended or expanded; and the ease with which a tool can be created in the future that will display information, which can be easily understood.

Norm specification: Norm specification is at the centre of the framework being developed. Most of the rules that exist within a business environment can be classified as behavioural standards that provide expectations for the way people should conduct themselves, comparable with deontic operators, which specify whether an activity is compulsory, acceptable or forbidden [74]. This arrangement can also be attributed to the conditions of behavioural norms, as suggested by Liu [38]: Whatever the requirements of the situations are, the agent is the deontic operator of the activity.

Table 3 provides an example of the norm analysis that has been exposed by semantic analysis. The norm analysis discovered that all the construction items linked to building elements need to be identified. The quantity surveyor is required to include all the details relating to the building element as a means of categorising them into cost items, as shown in Table 4 Norm specification example and Table 3 Summary of norm tags. (It should be noted that N 1.2 will be detailed in the next section, exploring how this discovery was identified, as it does not form part of the norm analysis).

Table 3 Norm analysis example

Trigger decompose		
Input:	trigger description:	Output:
1. Building elements 2. Construction work result	Elements need to be further deconstructed; the building element has been attached to construction work, resulting in IFC property criteria. According to these standards, the construction product should list all the properties of the construction work results.	construction product;
Determiner		
Example		
A wall is designed as adobe brick, 102mm*102mm*305mm, cement mortar used, and located in a bath room.	The deconstruction needs to link all essential information to the construction products.	Wall, adobe brick, 102mm*102mm*305mm, cement mortar, bath room.
Knowledge: The principle which states that deconstruction of elements is to be grouped by different material, different types, and different properties.		

Table 4 Norm specification example

N1.1: whenever [quantifying **building element**]
if [**construction work results**] [**is**] [**not null**]
then the [quantity surveyor]
is [obliged]
to [list all the properties]

N1.3:	whenever [quantifying **building element**] if [**material and type**] [**is**] [**not same**] then the [quantity surveyor] is [obliged] to [classify construction product into different categories]

Table 5 Summary of norm tags

whenever [quantifying] [**building element**]
if [**property**] [**comparison**] [**value**]
then [quantity surveyor]
is [obliged]
to do **[action]**

Addition of meta-data: refers to the inclusion of tags stemming from norm analysis. The addition of these tags to a document will lead the specialist to include addition metadata items.

- The building component, i.e. Door, Wall, Window etc.
- The property i.e. type, material, location etc.
- The comparison i.e. equal to (=), greater than (>), less than (<) etc.
- The value
- The action

The two main sources of information that have a semantic impact are the building component and the property. When entering data for a particular tag the specialist would follow these steps: (1) choose a building component from those within the core ontology and (2) stipulate a property from either a pre-existing selection of properties or enter a new one. This procedure has been demonstrated in Table 6 – Adding metadata to code of practice document.

Table 6 Adding metadata to code of practice document

18 Tile and slate roof and wall coverings Plain tiling Interlocking tiling Fibre cement slating Natural slating Natural or artificial stone slating Timber of bituminous felt shingles Any other type of tile, slate, slab or block roof or wall covering								
Item or work to be measured	Unit		Level one		Level two		Level three	
Roof coverings Wall covering	m2	1 2	Pitch stated Vertical	1	Underlays and battens	1 2	Curved: radii stated. Conical: maximum and minimum radii stated	
Boundary work; location and method of forming described	m	1	Dimensioned description stating net girth		Abutments. Eaves. Ridges. Verges. Valleys. Hips. Vertical angles		Horizontal Sloping Raking. Vertical. Curved: radius stated. Stepped Pre formed	
Note:	Coverings are deemed to include underlays, battens and work in forming voids ≤ 1m2 No deduction is made for voids ≤ 1m2							
	Boundary work to voids is only measured where the void exceeds 1.00m2							

Condition on Building elements		Property	Comparison	Value	Assertion
Decompose	Wall Covering	work result, location, type,	=	Not null	List

		material			
Decompose	Wall Covering	voids	>	1 m2	Create
Decompose	Boundary work	work result, location, type, material	=	wall covering work result, and location	List

These stages of work impact the underlying pragmatics of the code of practice information. When a norm is stipulated, it will be recorded as a relationship within the code of practice ontology, after which a sub-class relationship will be created linking activities with sub-activities. This is a vital step in producing a full awareness of the semantic structure and pragmatics within the codes of practice. A subsequent result of this method is that a grasp of the structure of the code of practice is created, while the categories within the cost engine are also filled with data.

In conclusion, there are 13 steps in total for cost estimation [75], each step should repeat this process and 14 potential scenario have been engineered. The norm specifications created are being used as semantic tags to improve the code of practice documents and obtain categories for the cost engine. It is interesting to note that each step is supposed to correspond to one scenario, however during the norm analysis and interviewing with an expert, when an expert is identifying the item in the database, it occasionally happens that they cannot find such a data item and quotation from the market is required. Thus it is important to include this scenario in the step identifying N5, see Figure 8, Knowledge-based cost estimation application process. This is not intended to be an exhaustive list of potential scenarios however it is adequate for the purposes of the case study.

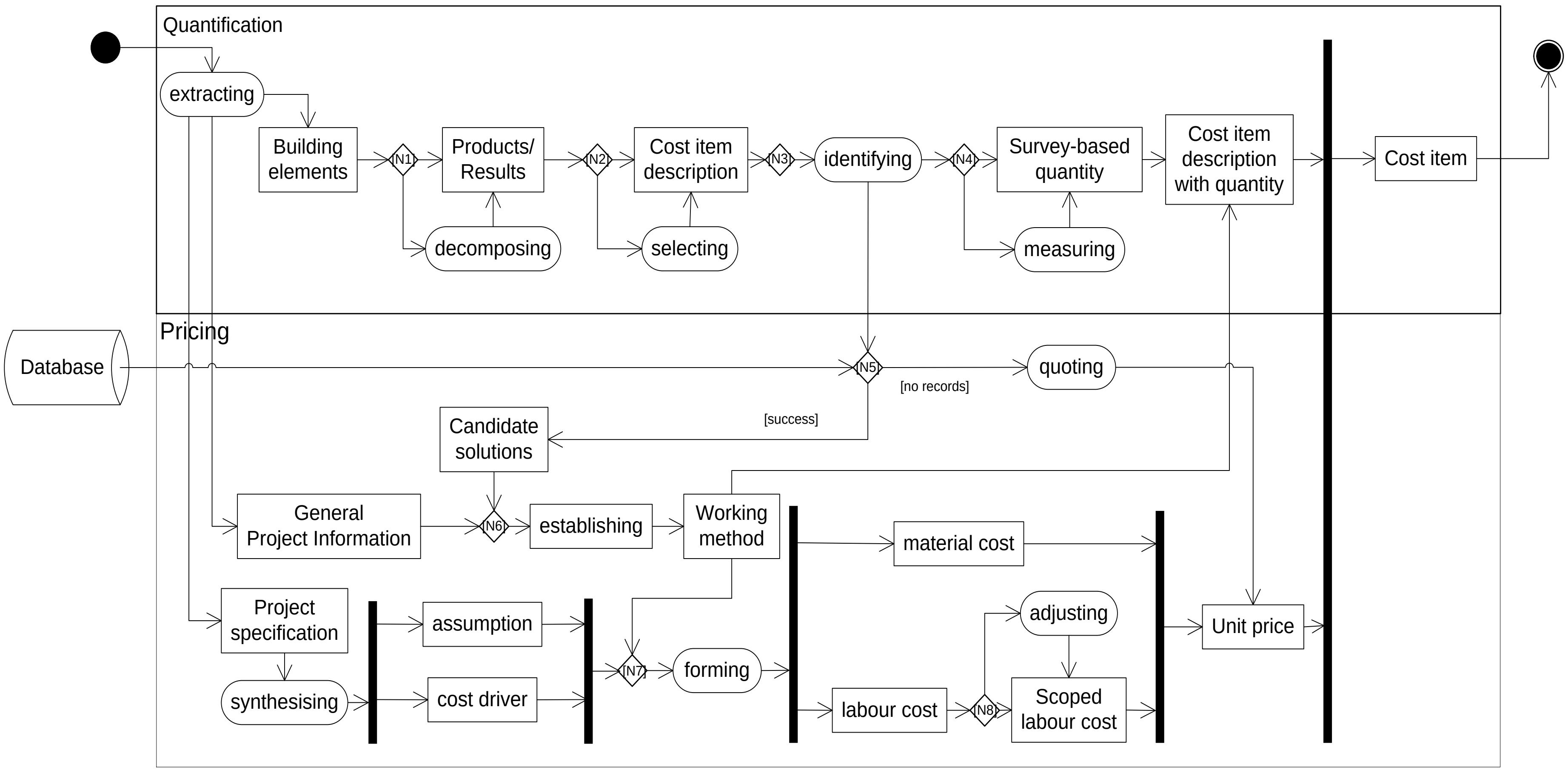


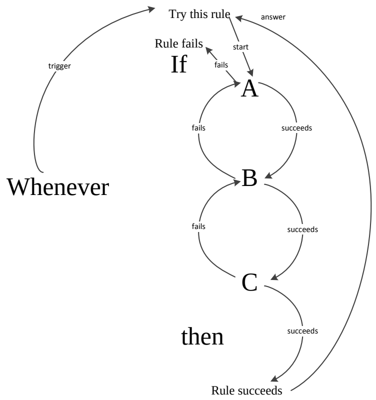
Figure 8 Knowledge-based Cost Estimation Application Process

Generating the clauses:

The norm specification tags attributed to the code of practice include a detailed rational structure. Adapting this structure to present a format that can be electronically read requires it to be translated into Horn clauses [76]. All the categories provided in this section relate to Prolog clauses syntax [77].

The procedure involved in translating the populated tags into Horn clauses is carried out by employing a range of logical formulae (rule operand) derived from the norm [38]. Typically, each tag is translated into a clause, which is then linked into a series of rules in Prolog, which in turn drive the outcome. As an example, by defining the link between the building element and the specific construction items, it is possible to produce a new cost item for a particular building component, such as the wall covering for the boundary work described above.

Figure *9* relates to the rational formula that is used in defining the consequences. The heading 'Whenever' identifies the state of the building component as a certain point, which it will then compare with the details in each of the clauses, after which the action will be performed. This will involve each clause being proven to be true or false. Where they are all true, the action will take place. A key advantage of this is that a record of data is collated, which can be referred to when specific information is absent.



Controlled Process in Prolog	Whenever	Clauses Specification
decomposing	(Buildingelement)	wall(Buildingelement), has_type(Buildingelement,Y), has_material(Y,A), layerthickness(Y,B), has_working_method(Y,C), has_method(C,D), member(A,[brick,adobe_brick]) → E=[A,B,D], assert(has_name(constructionproduct,E)), assert(decompose(Buildingelement,E)), assert(walls(constructionproduct).

Figure 9 Rule logic

5.3 Deployment and Implementation of application

To implement the framework in order to allow its authentication within the case study, a prototype of the application was created to provide cost estimation details. The framework was introduced in three stages: (1) metadata was obtain from code of practice documents, which related to cost estimation and new rules of measurement [43]; (2) semantic details were recorded; and (3) the knowledge-based cost estimation application process was introduced.

Semantic analysis: to accomplish this outcome of the framework the Order of Cost Estimation [56] and key procedures with regard to cost estimation [11,12,78–80] have been considered. The outcome has led to 123 units of metadata.

Semantic data transfer: to accomplish this outcome of the framework the semantic items used within the code of practice and the language of the IFCs have been recorded, as shown in Table 7.

Table 7 Semantic units mapping

Documents	Total	IFC	Number of	Number of

	entity	attributes	relationships
Order of Cost Estimation	22	27	8
Bill of Quantities	42	59	22

Implementation of procedures: the process then requires the introduction of the procedures that were highlighted by the semantic analysis process. These procedures relate to the activities carried out by the specialist and each step is linked to a particular knowledge set. Additionally, as a means of validating each of these 13 steps, a feedback form was created to provide assurance.

The next activity to be undertaken in the creation of the system is to authenticate the costing elements. This can be completed for official ontologies by using a reasoning engine. Protege offers this service through built in reasoners, for example FACT++, Hermit and Pellet [81–83]. The consistency of each element was checked and an inconsistent element was produced and identified using the reasoner.

The last stage required for this case study is confirmation that the rules have been created. This was achieved by translating norms using a logic programme. The engine developed has been created in such a way as to cater for the largest number of codes of practices possible and to allow the following actions to be performed: the ability to interrogate and amend the semantic model; the ability to extrapolate non-existing building components; and the ability to carry out cost estimation following the analytical process.

As a means of confirming the abilities of the cost estimation system, it is to be tested out on a genuine building project. This will be achieved by linking the cost estimation system to a specific version of IFC. For the purposes of the case study, a tool has been applied to the IFC to enable the IFC physical file version 2*3 to be translated into a resource description framework (RDF) format [84]. A sample of the translated IFC physical file is shown in Figure 10.

```
<!-- #ifcwall_15478 -->
  <owl:NamedIndividual rdf:about="&untitled-ontology-12;ifcwall_15478">
    <has_id rdf:datatype="&xsd;string">15478</has_id>
    <ace_lexicon:PN_sg rdf:datatype="&xsd;string">IfcWall_SW-007</ace_lexicon:PN_sg>
    <has_type rdf:datatype="&xsd;string">external wall</has_type>
    <name rdf:datatype="&xsd;string">sw-007</name>
    <has_constructionworkresult rdf:resource="&untitled-ontology-12;brick1"/>
    <has_quantity rdf:resource="&untitled-ontology-12;ifcquantityarea_33501"/>
    <has_quantity rdf:resource="&untitled-ontology-12;ifcquantitylength_33505"/>
    <has_quantity rdf:resource="&untitled-ontology-12;ifcquantitylength_33506"/>
    <has_quantity rdf:resource="&untitled-ontology-12;ifcquantitylength_33507"/>
    <has_quantity rdf:resource="&untitled-ontology-12;ifcquantityvolume_33499"/>
    <has_spacefunction rdf:resource="&untitled-ontology-12;ifcspace_1234568"/>
  </owl:NamedIndividual>
```

Figure 10 IFC Entity represented by Web Ontology

5.4 Results

Following completion of the development of the costing model, it was trialled on an actual building project. An IFC model for the creation of a residential project was employed in this instance and for the purposes of the case study, the building was a two storey villa.

The strategy that has been employed in the course of this assessment is to utilise the developed system alongside the traditional manual cost estimating system, using the standard procedures currently in existence within the industry. This enables the researchers to see a clear comparison between the outputs.

Overall the new rules of measurement (NRM) entails more than 350 rules; that being said this analysis focused on those concerned with wall coverings as a means of providing an example. Figure 11 shows the key rules associated with understanding the non-existing building components of walls.

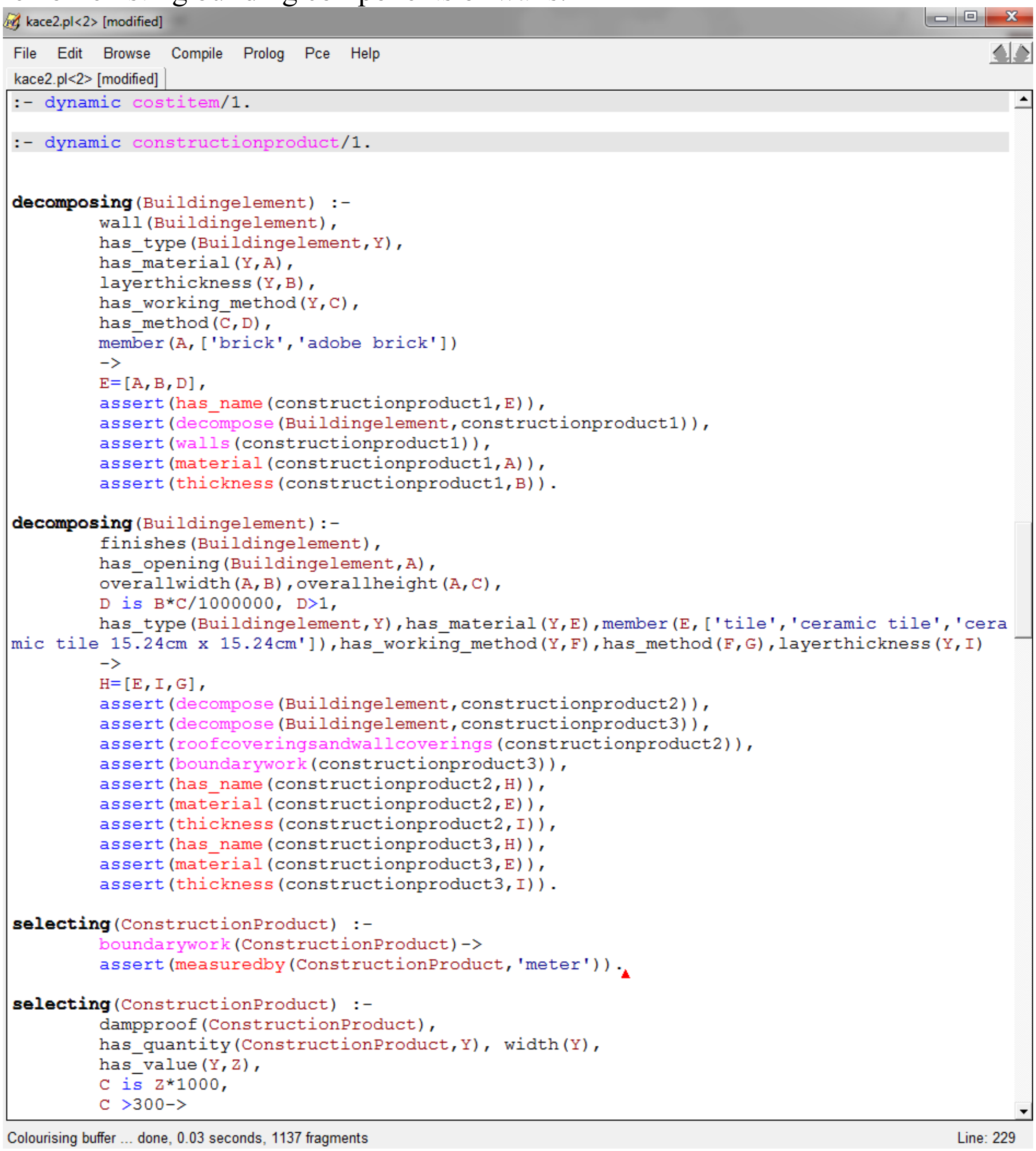


Figure 11 Rules developed

Table *8* displays the outcome of a contrast between the planned and existing methods of cost estimation. Specialists were invited to produce the cost item using the same information that is provided. Specialist 1 reported that there were not enough details given with regards to tiles and additional quotations were needed, while specialist 2 did not initially identify the requirement for boundary work. Through the feedback loop they concurred that the measurement requirements highlighted the need for this work and that the cost could be translated from the database supplied. Specialist 3 identified that, due to the height of the wall, on-site productivity may fall by 20%, which impacts the labour unit cost (rising by 20%). They also noted the possibility that the height may affect the tile use within the boundary works. Additionally, our application includes a particular piece of expert knowledge that the wall having a height of between 2.74m and 3.96m, rolling scaffolding can be used for construction which improve our productivity by 30%, which is an example provided in Staub's research [85].

Table 8 Result of comparison (Exterior Tile Finishes Boundary work)

Components	Traditional method				Proposed method
	1 (4 years)	2 (8 years)	3 (2 years)	4 (12 years)	
Cost Item Description*	A, D	Missed	A, B, C	A,B,C,D,E	A,B,C,D,E
Measured Quantity	9m	Missed	9m	9m	9m
Material Unit Cost	unknown	Missed	33.1$/m	33.1$/m	33.1$/m
Labour Unit Cost	unknown	Missed	34.18$/m	34.18$/m	34.18$/m
Adjust on Labour Unit Cost	0	Missed	Increase 20%	0	Decrease 30%
Scoped Labour Unit Cost	unknown	Missed	41.02$/m	34.18$/m	23.93$/m
Total	0	Missed	$667.08	$605.52	$513.23

*Note: A: Material, B: Working Methods, C: Size, D: Location, E: Additions

The third section provides a contrast between cost estimation processes. In comparison with the advanced BIM-based cost estimation system, a method originating from semantic analysis, providing additional itemisation of the costing process is examined, as shown in Table *9*.

Table 9 Comparison between existing cost estimation methods and proposed cost estimation method

Estimation methods	General process of cost estimation Part 1			
	Describing cost item	Infer working method	Find unit cost	Form unit price
Conventional cost estimation	Manual	Manual	Manual	Manual
BIM based cost estimation	Manual	Manual	Manual	Manual
Other studies	1	2	2	None
Framework based application (BIM+Ontology/Logic)	Auto	Auto	Auto	Auto
Estimation methods	General process of cost estimation Part 2			
	Calculate quantity	Identify construction condition	Adjust unit price	Review cost plan
Conventional cost estimation	Manual	Manual	Manual	Manual
BIM based cost	Automatic*	Manual	Manual	Manual

estimation				
Other studies	1	3	3	None
Framework based application (BIM+Ontology/Logic)	Auto	Auto	Auto	N/A

** BIM based quantity calculation is model based, doesn't comply with standard*

Note: 1. [31]*; 2.* [27]*; 3.* [32]

6 Discussion

This study has encompassed elements of both cost estimation and information systems. It is understood that the incorporation of technological elements into the everyday undertakings of construction professionals has significant value, particularly with regard to time reduction within cost estimation activities. Key research activities have been considered in producing an information system engineering process, such as knowledge engineering processes, semiotic frameworks, semantic analysis and norm analysis. The information has been collated from literature reviews, evaluation of systems and an appraisal of existing standards, prior to their assessment using the methods described above. This has led to the understanding the there is significant value in adopting a methodical approach, which has been carried out in this study.

In addition, the use of rules and ontology to produce intelligent solutions for construction projects is extensively recognised. Of the studies carried out between 1990 and 2012, 30.74% focused on knowledge-based and specialist systems [86]. In particular, the most recent studies concentrate on rule-based semantic methods, for activities such as regulation checking [35] and an ontological method relating to cost estimation [27]. This study emphasises the advantages associated with a rule-based semantic method of cost estimation and it has undertaken a systematic approach to establishing a rule-based semantic process. It is anticipated that the advantages gained by utlising OS can provide opportunities to discuss an appropriate process for transforming text-based documents into quantifiable clauses.

Table 6 demonstrates that when the empty space is larger than $1m^2$, a new activity is required to develop a new cost item for boundary work. At the same time, a property 'inter-section area' is absent from the available data. The situation that is likely to be confronted during the addition of metadata into a code of practice document is an element that has not been included in the data structure. This particular situation will be met on a regular basis within actual construction projects. In these circumstances the quantity surveyor must ensure the reliability of the bill of quantity or risk complaint regarding contract variations. Under these circumstances, if a cost estimation is undertaken, the literature identifies three outcomes. Firstly, an interface could be created that allows the specialist to enter data directly into the system. In this instance this may relate to the boundary work's location, type and material required to produce a new cost item or alternatively the value of the inter-section of the specific elements. Secondly, a request could be made to the developer of the data format to include details of the absent element. Thirdly, using a semantic method, the SWRL could be employed to surmise the details. These possibilities are not appropriate for the research currently being carried out however for the following reasons: the first entails extensive amounts of manual data input due to variances in each wall covering and the values for each inter-section needing to be computed separately for each element. The second would involve time-consuming involvement from the data format developers. To put this in context, an IFC developer takes at least six months to produce a new item [87]. The third option presents an innovative method of cost estimation however it is not entirely practical [88]. The proposed method, however, makes use of Horn clauses by stating pragmatics in order to employ the ontology. For this reason, the new approach will be included in the norm specification and the 'inter-section area' will be computed by selecting a specific activity to be completed, as shown in Table 10 Adding of norm specification.

Table 10 Adding of norm specification

N1.2:	whenever [quantifying **building element**] if [**the voids**] [**greater than**] [**certain value**] then the [quantity surveyor] is [obliged] to [create a boundary work cost item]

This paper depicts a framework of developing a knowledge based system based on IFC schema for cost estimation. The translated rules are based on the new rules of measurement which is suitable for most construction works for housing [89]. A developed knowledge base can be reused for real projects with constraints. Firstly, the measurement standards should be NRM, and the system is applicable to UK practice; secondly, the construction project should be a residential housing project; thirdly, the IFC file should be a version of 2*4; fourthly, the level of details should reach LOD500, which is just before tendering; moreover the costing database should be R.S. Mean and, lastly, only the wall case in a residential villa project has been tested. Furthermore, there are more than 350 rules in the NRM and translating all rules and testing them will be a manual intensive process. It requires large amounts of labour efforts but it is believed that it can benefits the company as the rules are reusable for individual similar projects. Fully developed system and related evaluation is not in the scope of this research paper.

This study makes an initial attempt to establish an overarching framework, resulting in a highly conceptualised mapping. The elements detailed allow for the activities of the standards of practice for the framework to be easily adapted, without impacting either the file formats containing area information or how the overall execution of the concept is achieved. The system in use requires analysis of the individual concepts, which in turn ensured that data formats and procedures relating to producing a methodical implementation were entirely successful. This led to the production of many information sources for the overall framework. Additionally, it should be highlighted that this allows further data formats to be created, notwithstanding the original goal of focusing solely on the area of construction pricing.

The novelty of this framework is that it explicates the functionality of the three focal aspects, i.e. syntax, semantics and pragmatics. Each aspects serves certain functions within the framework and by integrating them it can bring tangible benefits to the domain experts. Furthermore, from the practical perspective of semantic technologies [90] and data provenance [30], the proposed framework brings together the entity, agent, and process into an industrial usage through the deployment of various logics.

7 Conclusion

This paper has detailed the activities undertaken in the application of the framework for construction cost estimation, which has been evaluated using NRM. Moreover, a representative example has been demonstrated to define the use of the model by industry specialists in the production of a trustworthy cost estimation framework that meets their requirements. The framework produced as an outcome of this study has been shown to be effective, dependable and suitable for the completion of quantity surveying on building data, as authenticated by NRM.

The selection of this situation as an example of the activities undertaken by costing experts within construction has been appropriately confirmed. It is also the case that they are suitable for use with BIM products, due to their ability to: (1) display standards of practice without any requirement for additional software products; (2) analysing and updating the standards of practice as they alter by each area; and (3) identifying specialist analytical instructions for individual use. This process of improving the standards of practice and analytical instructions can be carried out in a simple manner through the development of semantic assessment and semantic recording tools created at the outset of the process. In addition, the inclusion of semiotic procedures to assess the BIM-based cost estimation details confirms the concept that the IFC structure lacks a degree of analytical ability, which is a key requirement of advanced cost estimation processes. It has also permitted the implementation of supplementary items of data to improve the IFC provisions, through the added support of expense prediction actions. These elements can be transferred back to buildingSMART where required in a much clearer set of cost estimations.

The current sample model employs a number of software products, including IFC Parser, Protégé and Prolog, which verifies the original theories connected to the system. Moreover, the system also has the ability to be improved by: (1) incorporating plug-ins with the initial BIM system (such as Benley Microstation and ArchiCAD BIMx) as a means of automating the costing documents; (2) employing an interface of specific procedures and guidelines; or (3) producing a spontaneous ontology based on semantic assessments. It is anticipated that natural language processing and efforts to build in spontaneous ontologies will bring about semi-automation of various elements within the planned system, in particular the partial removal of the original semantic and pragmatic details from text based documents. Finally, it is expected that, in light of the growing requirement for electronic services within the industry, standards of practice will be detailed in a strongly semantic style, and therefore clearly defining the connections between items.

It should be noted that adding further metadata and obtaining the clauses is carried out manually at present. There is the possibility of improving the system to allow this to take place automatically. On the other hand, as reliance on automated systems increases, it is anticipated that codes of practice will be delivered in a semantically focused manner, allowing the connections between items to be clearly identified. This conceptual change in the presentation of codes of practice will ensure that they are produced in such a way as to allow them to be automatically processed, with the ability for human understanding therefore being an output of the automated system rather than an input. Beach et al. (2015) support this approach in their current goals of achieving automated regulation checking processes.

It is also believed that the framework can be generalized to a wider domains which are in the nature of expert reasoning based upon rules/clauses. By decoupling the syntax, semantics and pragmatics, the domain can be fully understood and expert knowledge can be captured for computing. Meanwhile, it has a practical impact on the industrial implementation of knowledge-based systems in that the publishing of ontology is separated from the expert reasoning processes. This means that company/expert competitiveness can be guaranteed.

8 Reference

[1] M.-Y. Cheng, H.-C. Tsai, W.-S. Hsieh, Web-based conceptual cost estimates for construction projects using Evolutionary Fuzzy Neural Inference Model, Autom. Constr. 18 (2009) 164–172. doi:10.1016/j.autcon.2008.07.001.

[2] M.N. Beltramo, Beyond Parametrics: The role of subjectivity in cost models, Elsevier Eng. Costs Prodcution Econ. An Int. J. Ind. 14 (1988) 131–136.

[3] Stapel, S., 2002, March. The Eurostat construction price surveys: history, current methodology and new ways for the future. In International Conference on ICP, World Bank, Washington, DC, March (pp. 11-13).

[4] H. Karanci, A Comparative Study of Regression Analysis, Neural Networks and Case-based Reasoning for Early Range Cost Estimation of Mass Housing Projects, 2010.

[5] U.S. Department of Energy, Cost Estimating Guide, Washington, D.C., 2011.

[6] International Society of Parametric Analysts, Parametric Estimating Handbook, 2008.

[7] M. Shivani, R. Ghanshyam, A. Khandekar, M.E. Student, Review on Cost Estimation Methods for Software Development, (2013) 132–134.

[8] S.-H. An, G.-H. Kim, K.-I. Kang, A case-based reasoning cost estimating model using experience by analytic hierarchy process, Build. Environ. 42 (2007) 2573–2579. doi:10.1016/j.buildenv.2006.06.007.

[9] R. Roy, P. Souchoroukov, E. Shehab, Detailed cost estimating in the automotive industry: Data and information requirements, Int. J. Prod. Econ. 133 (2011) 694–707. doi:10.1016/j.ijpe.2011.05.018.

[10] L.C. Briand, K. El Emam, F. Bomarius, COBRA : A Hybrid Method for Software Cost Estimation , Benchmarking , and Risk Assessment COBRA : A Hybrid Method for Software Cost Estimation , Benchmarking , and Risk Assessment, 24 (n.d.).

[11] N. Sinclair, P. Artin, S. Mulford, Construction cost data workbook, in: World Bank (Ed.), Conf. Int. Comp. Progr., Washington, D.C., 2002. http://4estimators.ucoz.com/_ld/0/5_dms.final.doc (accessed March 4, 2014).

[12] S. Isakowitz, NASA Cost Estimation Handbook, NASA HQ, Washingt. DC, Spring. (2002). http://www.nasa.gov/pdf/263676main_2008-NASA-Cost-Handbook-FINAL_v6.pdf (accessed March 4, 2014).

[13] Mair, C., Martincova, M., & Shepperd, M. (2009). A literature review of expert problem solving using analogy.

[14] RSMeans, RSMeans Building Construction Cost Data 2013, 2013. https://www.rsmeans.com/.

[15] BCIS, Standard Form : Property Occupancy Cost Analysis Standard Form of Property Occupancy Cost Analysis, RICS. (n.d.).

[16] Kiziltas, S., & Akinci, B. (2009). Contextual information requirements of cost estimators from past construction projects. Journal of Construction Engineering and Management, 135(9), 841-852.

[17] R.D. Stewart, R.M. Wyskida, J.D. Johannes, Cost Estimators Reference Manual, 1995.

[18] Department of Defence, Parametric Estimating Handbook, 2nd Ed., DoD, 1999.

[19] R.A. Mileham, C.G. Currie, A.W. Miles, D.T. Bradfor, A Parametric Approach to Cost Estimating at the Conceptual Stage of Design, J. Eng. Des. 4 (1993) 117–125.

[20] R.T. Hughes, Expert Judgement as an Estimating Method, Inf. Softw. Technol. 38 (1996) 67–75.

[21] S. Staub-French, M. Fischer, J. Kunz, B. Paulson, A generic feature-driven activity-based cost estimation process, Adv. Eng. Informatics. 17 (2003) 23–39. doi:10.1016/S1474-0346(03)00017-X.

[22] Staub-French, S., Fischer, M., Kunz, J., Paulson, B., & Ishii, K. (2002). An ontology for relating features of building product models with construction activities to support cost estimating. Center for Integrated Facility Engineering Working Paper, (70).

[23] S. Staub-french, A.M. Asce, M. Fischer, J. Kunz, B. Paulson, M. Asce, An Ontology for Relating Features with Activities to Calculate Costs, (2003) 243–254.

[24] F. H. Abanda, J.H.M. Tah, C. Pettang, M. B. Manjia (2011) An ontology-driven building construction labour cost estimation in cameroon, ITcon Vol. 16, pg. 617-634, http://www.itcon.org/2011/35

[25] G. Fidan, I. Dikmen, A. Tanyer, B. M., Ontology for relating risk and vulnerability to cost overrun in international projects., J. Comput. Civ. Eng. 25 (2011).

[26] Z. Ma, Z. Wei, Framework for Automatic Construction Cost Estimation Based on BIM and

Ontology Technology, in: Proceedings of the CIB W78 2012: 29th International Conference – Beirut, Lebanon, 17-19 October

[27] S.S.-K. Lee, K.-R.K. Kim, J.J.-H. Yu, BIM and ontology-based approach for building cost estimation, Autom. Constr. 41 (2014) 96–105. doi:10.1016/j.autcon.2013.10.020.
[28] F.H. Abanda, J.H.M. Tah, R. Keivani, Trends in built environment semantic Web applications: Where are we today?, Expert Syst. Appl. 40 (2013) 5563–5577. doi:10.1016/j.eswa.2013.04.027.
[29] H. Grzybek, S. Xu, S. Gulliver, V. Fillingham, Considering the Feasibility of Semantic Model Design in the Built-Environment, Buildings. 4 (2014) 849–879. doi:10.3390/buildings4040849.
[30] B. Plale, S. Miles, C. Goble, P. Missier, R. Barga, Y. Simmhan, et al., The open provenance model (v1. 01), Univ Southampt. TR. (2008).
[31] Z. Ma, Z. Wei, X. Zhang, Semi-automatic and specification-compliant cost estimation for tendering of building projects based on IFC data of design model, Autom. Constr. 30 (2013) 126–135. doi:10.1016/j.autcon.2012.11.020.
[32] S. Staub-French, M. Fischer, I.K. Kunz J, B. Paulson, A feature ontology to support construction cost estimating, Artif. Intell. Eng. Des. Anal. Manuf. 17 (2003) 133–154.
[33] M.P. Nepal, S. Staub-French, R. Pottinger, J. Zhang, Ontology-Based Feature Modeling for Construction Information Extraction from a Building Information Model, J. Comput. Civ. Eng. (2012) 120814090309004. doi:10.1061/(ASCE)CP.1943-5487.0000230.
[34] Z. Ma, Z. Wei, P.D. Candidate, Z. Liu, Ontology-Based Computerized Representation of Specifications for Construction Cost Estimation, (2013) 707–716.
[35] T.H. Beach, Y.R. Rezgui, H. Li, T. Kasim, A Rule-based Semantic Approach for Automated Regulatory Compliance in the Construction Sector, Expert Syst. Appl. 42 (2015) 5219–5231. doi:10.1016/j.eswa.2015.02.029.
[36] R. Stamper, K. Liu, M. Hafkamp, Y. Ades, Understanding the Roles of Signs and Norms in Organisations-A semiotic approach to information systems design, J. Behav. Inf. Technol. 19 (2000) 15–27.
[37] Information in business and administrative systems,John Wiley & Sons, Inc. New York, NY, USA ©1973, ISBN:0471820458
[38] K. Liu, Semiotics in Information Systems Engineering, Cambridge University Press, Cambridge, 2000. doi:10.1017/CBO9780511543364.
[39] F.S. Staub, M. Fischer, J. Kunz, K. Ishii, B. Paulson, S. Staub-French, et al., A feature ontology to support construction cost estimating, Artif. Intell. Eng. Des. Anal. Manuf. 17 (2003) 133–154. doi:10.1017/S0890060403172034.
[40] ISO 10303-21, Industrial automation systems and integration -- Product data representation and exchange -- Part 21: Implementation methods: Clear text encoding of the exchange structure, (2002).
[41] ISO 10303-28, Industrial automation systems and integration—Product data representation and exchange—Part 28: Implementation methods: XML representations of EXPRESS schema and data, (2007).
[42] buildingSMART International Limited, IFC Standard, Build. Int. Ltd. (2013). http://www.buildingsmart-tech.org/specifications/ifc-overview (accessed July 25, 2016).
[43] The Royal Institution of Chartered Surveyors, RICS New Rules of MEASUREMENT Bill of Quantities for Works Procurement, Coventry, 2011.
[44] Stamper, R., & Liu, K. (1994, January). Organisational dynamics, social norms and information systems. In System Sciences, 1994. Proceedings of the Twenty-Seventh Hawaii International Conference on (Vol. 4, pp. 645-654). IEEE.
[45] H. Grzybek, S. Gulliver, Z. Huang, Inclusion of Temporal Databases with Industry Foundation Classes-A Basis for Adaptable Intelligent Buildings., in: ICISO, 2010. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Inclusion+of+Temporal+Databases+with+Industry+Foundation+Classes+-+a+basis+for+adaptable+intelligent+buildings#0 (accessed June 24, 2014).
[46] Nicolle, C., & Cruz, C. (2010, April). Semantic Building Information Model and multimedia for facility management. In International Conference on Web Information Systems and Technologies (pp. 14-29). Springer Berlin Heidelberg.
[47] J. Steel, R. Drogemuller, B. Toth, Model interoperability in building information modelling, Softw. Syst. Model. 11 (2010) 99–109. doi:10.1007/s10270-010-0178-4.
[48] Z. Shen, Semantic 3D CAD and Its Applications in Construction Industry — An Outlook of

Construction Data Visualization, (2007).
[49] Venugopal, M. (2011). Formal specification of industry foundation class concepts using engineering ontologies. Georgia Institute of Technology.
[50] M. Venugopal, C. Eastman, J. Teizer, Formal Specification of the IFC Concept Structure for Precast Model Exchanges, Comput. Civ. Eng. 2012 (2012) 213–220. doi: 10.1061/9780784412343.0027.
[51] T. Hartmann, A Semiotic Analysis of Building Information Model Systems, Comput. Civ. Eng. (2012) 381–388. doi:10.1061/9780784412343.0048.
[52] S. Xu, K. Liu, L.C. Tang, Incorporation of expert reasoning into the BIM-based cost estimating process, in: A.B. Raidén, E. Aboagye-Nimo (Eds.), Procs 31st Annu. ARCOM Conf., Lincoln, Uk, 2015: pp. 1–7.
[53] Tan, S., Liu, K. and Xie, Z., 2004. A Semiotic approach to organisational modelling using norm analysis. In 6th Int. Conf. Enterp. Inf (pp. 1-15).
[54] ISO 16739, Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries, (2013).
[55] BuildingSMART, IFC4 - the new buildingSMART Standard, 2013. http://www.buildingsmart-tech.org/ifc/IFC4/final/html/
[56] The Royal Institution of Chartered Surveyors, RICS New Rules of Measurement - Order of Cost Estimating and Elemental Cost Planning, Coventry, 2007.
[57] Ma, H., Ha, K. M. E., Chung, C. K. J., & Amor, R. (2006, June). Testing semantic interoperability. In Proc. of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering (ICCCBE) (pp. 1216-1225).
[58] S. Xu, K. Liu, W. Li, Knowledge-based Design Cost Estimation through Extending Industry Foundation Classes, in: 16th Int. Conf. Enterp. Inf. Syst., Institute for Systems and Technologies of Information, Control and Communication, Lisbon, Portual, 2014: pp. 161–168.
[59] buildingSMART International Limited, Freeware - Ifc, (2015). http://www.ifcwiki.org/index.php/Free_Software (accessed March 27, 2015).
[60] R. Stamper, The semiotic framework for information systems research, in: H. Nissen, H. Klein, R. Hirschhaim (Eds.), Inf. Syst. Res. Contemp. Approaches Emergent Tradit., North Holland, Amsterdam, 1991: pp. 515–528.
[61] G.R. Smith, A.S. Hanna, Factors influencing formwork productivity, Can. J. Civ. Eng. 20 (1993) 144–153.
[62] Portas, J., & AbouRizk, S. (1997). Neural network model for estimating construction productivity. Journal of construction engineering and management, 123(4), 399-410.
[63] B.H.R. Thomas, I. Zavrs, P. Ractice, Construction Baseline Productivity: Theory and Practice, (1999) 295–303.
[64] RSMeans., Building construction cost data, 2005. https://www.rsmeans.com/.
[65] E. Koehn, G. Brown, Climatic Effects on Construction, J. Constr. Eng. Manag. 111 (1985) 129–137. doi:10.1061/(ASCE)0733-9364(1985)111:2(129).
[66] A. Touran, M. Asce, Probabilistic Model for Cost Contingency, (2003) 280–284.
[67] A. Russell, Computerized daily site reporting, J. Constr. Eng. Manag. 119 (1993) 385–402. http://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9364(1993)119:2(385) (accessed February 12, 2014).
[68] M. Liberda, J.Y. Ruwanpura, G. Jergeas, Construction productivity improvement: A study of human, managerial and external factors, in: Proc., ASCE Constr. Res. Congr., ASCE, Reston, Va., Honolulu, 2003.
[69] E. Choy, J.Y. Ruwanpura, Situation based modeling for construction productivity, in: Proc., Constr. Res. Congr., ASCE, Reston, San Diego, CA, USA, 2005: pp. 662–666.
[70] A. Boussabaine, T. Elhag, Knowledge discovery in residential construction project cost data, 15th Annu. ARCOM Conf. 2 (1999) 15–17. http://www.arcom.ac.uk/-docs/proceedings/ar1999-489-498_Boussabaine_and_Elhag.pdf (accessed July 25, 2016).
[71] R. Akbaş, Geometry-based modeling and simulation of construction processes, Standford University, 2003. http://cife.stanford.edu/sites/default/files/TR151.pdf (accessed February 12, 2014).
[72] S. Staub-french, B. Columbia, Feature-based Product Modeling for Building Construction, (1991) 1–10.
[73] K. Liu, L. Sun, J. Barjis, J.L.. Dietz, Modelling dynamic behaviour of business organisations—extension of DEMO from a semiotic perspective, Knowledge-Based Syst. 16 (2003) 101–111. doi:10.1016/S0950-7051(02)00077-1.

[74] J.C. Meyer, Applications of Deontic Logic in Computer Science : A Concise Overview, (n.d.) 1–28.

[75] Xu, S., Liu, K., & Tang, L. C. (2015, March). Applying Organizational Semiotics for Developing Knowledge-Based Cost Estimation of Construction Project. In International Conference on Informatics and Semiotics in Organisations (pp. 80-91). Springer International Publishing.

[76] K. Samuel, L. Obrst, S. Stoutenberg, K. Fox, P. Franklin, A. Johnson, et al., Translating OWL and Semantic Web Rules into Prolog : Moving Toward Description Logic Programs, (2003) 1–21.

[77] J. Wielemaker, S. Ss, I. Ii, SWI-Prolog 6.6.6-Reference Manual, 2014. http://www.swi-prolog.org/download/stable/doc/SWI-Prolog-6.6.6.pdf(accessed June 1, 2014).

[78] F. Tan, T. Makwasha, “ Best practice ” cost estimation in land transport infrastructure projects, (2010) 1–15.

[79] Evans & Peck, Best Practice Cost Estimation for Publicly Funded Road and Rail Construction, 2008.

[80] V.O. Schinnerer, Construction Cost Estimating, (2007) 1–2.

[81] Shearer, R., Motik, B. and Horrocks, I., 2008, October. HermiT: A Highly-Efficient OWL Reasoner. In OWLED (Vol. 432, p. 91).

[82] S. Bechhofer, OWL : FaCT++, Dep. Comput. Sci. Kilburn Build. Univ. Manchester. (2004). http://owl.man.ac.uk/factplusplus/ (accessed January 7, 2015).

[83] E. Sirin, B. Parsia, B.C. Grau, A. Kalyanpur, Y. Katz, Pellet: A practical OWL-DL reasoner, Web Semant. Sci. Serv. Agents World Wide Web. 5 (2007) 51–53. doi:10.1016/j.websem.2007.03.004.

[84] P. Pauwels, J. Oraskari, IFC-to-RDF-converter, (2015). https://github.com/mmlab/IFC-to-RDF-converter.

[85] F.S. Staub, M.P. Nepal, Reasoning about component similarity in building product models from the construction perspective, Autom. Constr. 17 (2007) 11–21. doi:http://dx.doi.org/10.1016/j.autcon.2007.02.013.

[86] Z. Irani, M.M. Kamal, Intelligent systems research in the construction industry, Expert Syst. Appl. 41 (2014) 934–950. doi:10.1016/j.eswa.2013.06.061.

[87] Z. Ma, Z. Wei, W. Song, Z. Lou, Application and extension of the IFC standard in construction cost estimating for tendering in China, Autom. Constr. 20 (2011) 196–204. doi:10.1016/j.autcon.2010.09.017.

[88] B. Parsia, E. Sirin, B. Cuenca, E. Ruckhaus, D. Hewlett, Cautiously Approaching SWRL, (2005). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.222.7885.

[89] T.R.I. of C. Surveyors, The Royal Institution of Chartered Surveyors, RICS New Rules of Measurement - Order of Cost Estimating and Elemental Cost Planning, Coventry, 2007.

[90] T. Segaran, C. Evans, J. Taylor, Programming the semantic web, 2009. http://books.google.com/books?hl=en&lr=&id=XShw5YSHoRUC&oi=fnd&pg=PR5&dq=Programming+the+Semantic+Web&ots=DxVnXYcJL2&sig=qwUXB191zfC4c8bdgTSbspDe1F8 (accessed April 11, 2014).