SWRL Rule Development to Automate Spatial Transactions in Government

Premalatha Varadharajulu[1,2](✉), Lesley Arnold[1,2], David A. McMeekin[1,2], Geoff West[1,2], and Simon Moncrieff[1,2]

[1] Curtin University, Perth, WA 6845, Australia
p.varadharajulu@postgrad.curtin.edu.au
[2] Cooperative Research Centre for Spatial Information, Carlton, Australia

Abstract. The land development approval process between local councils and government planning authorities is time consuming and resource intensive because human decision-making is required to complete a transaction. This is particularly apparent when seeking approval for a new land subdivisions and administrative boundary changes that require changes to spatial datasets. This paper presents a methodology that automates the approval process by developing. Feedback on the transaction is communicated to the land developer in real-time, thus reducing process handling time for both developer and the government agency. This paper presents an approach for knowledge acquisition on rule development using Semantic Web and Artificial Intelligence to automate the spatial transaction process. The Web Ontology Language (OWL) is used to represent relationships between different entities in the spatial database schema. Rules that replicate human knowledge are extracted from government policy documents and subject-matter experts, and are defined in the form of Semantic Web Rule Language (SWRL) and based on geometry and attributes of database entities. The SWRL rules work with OWL-2 (spatial schema and vocabulary) ontologies to enable the automatic transactions to occur. These rules are implemented using an ontology and rule reasoner, which accesses the instances of data elements stored in the underlying spatial database. When the developer submits an application, the software checks the rules against the request for compliance with the relevant government policies and standards. This paper presents results for dealing with road proposals and road name approvals.

Keywords: Spatial transaction · Spatial data supply chain · Artificial intelligence · Semantic Web · Ontology · Rule-based reasoning · OWL-2

1 Introduction

Land developers and local government authorities are required to submit proposals for new subdivisions to land and planning departments for approval. These new subdivisions include new land parcel boundaries, roads and road names, and changes to local authority boundaries. The approval process often spans many

C. Grueau et al. (Eds.): GISTAM 2016, CCIS 741, pp. 122–142, 2017.
DOI: 10.1007/978-3-319-62618-5_8

work teams and new information, such as property addresses may need to be generated. This manual process can be time consuming and resource intensive.

New methods are required to reduce data handling and support the automation of transactions with government. Current workflows are characterised by several decision points and a trail of paper documents are often created to formalise the decision-making process and to provide a reference point for legal transactions further along the land administration process [1]. As a result, there is often a time delay of several weeks during which a new subdivision is considered by authorities from the various land development and planning perspectives.

This research seeks to automate the spatial transaction process using artificial intelligence with ontologies to create rules that replace the human decision-making process for land development approvals. A case study examining new road proposals, road names and land administration boundary changes is used to demonstrate the approach. This research is being conducted in conjunction with the Western Australian Land Information Authority (Landgate). Landgate is the approving authority for all new subdivisions in Western Australia, and is responsible for land administration boundary changes resulting from land development activity.

The Semantic Web was first introduced by Tim Berners-Lee who imagined it as “a web of data that can be processed directly and indirectly by machines” [2]. This research is inspired by the increased bandwidth of the Internet and advances in Semantic Web technologies, which now make it possible to automate many of the human elements of the decision-making process on the Web.

Rule-based systems have been used for decision support in the past but these are typically closed client bases systems. However the advantage of the Semantic Web is that the data, ontologies and rules are described using well defined standards (w3c.org) and can be made available over the Web as published resources, typically in one of a number of machine (and human) readable formats [3]. The vision of the Semantic Web is that, ontologies, especially those of a general nature, can be shared and re-used in many applications. In our case, it is envisaged that once a working solution for the approvals process has been validated for one jurisdiction (Western Australia), the ontologies and rules can be used in other jurisdictions (Victoria, New South Wales etc.) and domains.

The work is part of a research program into Spatial Data Infrastructures being conducted at the Cooperative Research Centre for Spatial Information (CRCSI), Australia. One of the objectives of the research program is to automate spatial data supply chains from end-to-end to enable access to the right data, at the right time, at the right price [4].

This research is focusing on the first stage in the spatial data supply chain process, which is the creation of spatial data generated through a land development business process. Instead of paper-based systems, the method enables the capture of spatial information in machine-readable form at its inception point. This is a significant step towards achieving downstream workflow automation. It also supports the recording of data provenance in machine-readable form at the commencement of a spatial transaction to support legal and data quality attribution.

The development consists of two stages. In the first stage, a GUI-based interactive system called Protégé is used to design ontologies and rules from spatial data schema and various documents including policies. The second stage uses a runtime environment (Jena and Java) to process the ontologies and rules along with existing and proposed road data to determine compliance with policies etc.

2 Background and Related Research

Methods for spatial data processing and integration have been researched and developed over the past few years, however little work has considered the automation of the decision-making process using the semantic web where spatial data is an input to the approval process.

One of the objectives of the Semantic Web is to evolve into a universal medium for information, data and knowledge exchange, rather than just being a source for information. To attain this, it uses the well known http protocol and technologies [5,6], such as URIs (Universal Resource Identifiers), RDF (Resource Description Framework) and ontologies with reasoning and rules.

One of the most important components is the RDF, which is a language for representing information about resources on the Web (http://www.w3.org/RDF/). RDF aims to organize information in a machine-readable format by representing information as triples: <subject, predicate, object>, a concept from the artificial intelligence community.

Traditionally, data is generally stored in relational databases. This has been a suitable model for the last few decades as it enables reasonable computers to store the data and allow searching. The advantage is that each piece of data is only stored in one place and each piece of data is atomic. The disadvantage is that the database tables have to be developed in advance usually from entity relational diagrams, the tables do not naturally relate to reality, and it is hard to link various databases together, especially if they are across different systems.

A more natural representation for the Internet (and Web) is the network or graph model. Data items are defined as nodes and the relationships defined as the arcs. A graph can represent anything and allow different pieces of disparate data to be related to each other. Extra links can be added on the fly without the need to redefine databases. For spatial data e.g. parcels in a cadastre where the norm is one person owns one parcel, it is easy to add links to show ownership of many parcels by one person, multiple people owning one parcel etc. Such changes can be made on the fly by the user as required, and there is no need for a data supplier to redesign databases to accommodate such changes.

RDF and triples are a way of defining a network as the triple <subject, predicate, object> defines two nodes (subject, object) and the link (predicate). Spatial data currently held in relational databases can be converted to triple stores and managed with software such as Fuseki. Current relational databases can be made into virtual triple stores as well. Triple stores can be queried using SPARQL (SQL for triple stores).

Importantly, each element of a triple can be a URI (or IRI for different languages), allowing further distribution of data and definitions. For example, if a predicate is called "near", the IRI can point to a location where the concept is defined. It may be the Euclidean distance between two points (spatial) or the distance between people in a family tree.

Of importance to the semantic web, RDF enables access to knowledge and rules, as well as the data allowing sophisticated user defined operations to occur, again without the data supplier having to configure systems specifically for a user. Ontologies and rules allow high level queries and processing to occur by many users on the fly, which is currently not possible.

RDF was originally considered as metadata but now covers data as well. RDF triples can be used to represent tables, graphs, trees, ontologies and rules because it describes the relationship between subject and object resources where a 'object' in the <subject, predicate, object> triple can be another subject enabling subjects to be linked together. Each of the triple components can also be a URI so information can be linked across the Web. RDF formatted data is much easier to process, because its generic format contains information that is clearly understandable as a distributed model.

Reasoning and rules are an important part of this research and in the Semantic Web, the Ontology Web Language (OWL-2), based on RDF, is used for defining Web ontologies that include rules, axioms and constraints allowing inferencing (discovery of new knowledge) to be performed.

The Semantic Web has been used for queries by a user for natural events using observation sensor data [7,8]. In particular [7] describe a number of ontologies used to model various sensors and rules used to map queries such as flooding in an area to the need to sample a number of point water sensors. Methods have been proposed that have potential to automate land development approval processes. For example, the Sensing Geographic Occurrences Ontology (SEGO) model supports inferences of institutionalized events [9] based on time. However they do not resolve any conflicts arising if an event qualifies based on both policy and business rules. This research does not cover the sensor-specific technical details [9], but instead concentrates on the business knowledge rules.

A large number of open source and proprietary tools are available for semantic web research and development. This research uses the Protégé framework (http://protege.stanford.edu/) to develop ontologies and rules because its GUI environment allows fast design, interactive navigation of the relationships in OWL ontologies and visualization. It allows some rule-based analysis to be performed and can read and write RDF-based files in a number of different formats. Rules are defined in the form of ontological vocabularies using Semantic Web Rule Language (SWRL). Like many other rule languages, a SWRL rule has the form of a link between antecedent and consequent. The antecedent refers to the body of the rule, consisting or one or more conditions, and the consequent refers to its head, typically one condition. Whenever the conditions specified in the antecedent are satisfied, those specified in the consequent must also be satisfied [10]. Once ontologies and rules have been defined, they can be imported into the Apache Jena framework complete with the Pellet reasoner (http://clarkparsia.com/pellet/) to support OWL for runtime querying and analysis [11]. Combining both Jena and OWL API libraries, Pellet infers logical consequences from a set of asserted facts or axioms.

3 Case Study

Landgate administers all official naming actions for Western Australia under the authority of the Minister for Lands. The relevant local government authority generally submits all naming proposals for ratification by Landgate. All new proposals must satisfy government policies and standards. The current process has an online submission form, but for the most part the process is paper-based and requires significant human involvement. Current methods often require negotiation between the parties involved (i.e. local government and Landgate). While there are specific rules applying to new road name approvals, there are grey areas within policy that are often challenged and can only be resolved by an experienced negotiator. A request for a new road name may be transferred back-and-forth until an outcome is achieved that is satisfactory to both parties. Outcomes may be different depending on the expertise of the negotiator/approver.

Automation is needed to reduce the manual overhead by extracting expert knowledge for road name approvals to create a standard set of rules. The notion is to create a self-service online mechanism for developers to submit new road names for approval, underpinned by a complex rule-base and querying process. Complexity comes from the flow on effect of such changes. A new land development results in a change to the surrounding road network. This has a flow on impact to property street addressing and an administrative boundary change.

The case study uses the Landgate geographic road names database, called GEONOMA, to process the road name proposal. The current online submission process has the following issues that complicate the approval process:

- The online form is only used to test whether new road names are allowable based on a set of road names that have been reserved for use. If a proposed name is a reserved road name then the request will fail. There is no opportunity to contest the decision.
- A maximum of ten names per application is allowed; meaning separate applications are required for larger subdivisions. It is not possible to conduct cross-reference checks against other submissions and therefore the process is open to error.
- The current system does not consider the spatial extent of roads. Figure 1 shows a schematic submitted for road name approvals that does not represent the actual proposed location of roads. Roads do not actually meet up; they are stylized with solid and dashed lines with arrows etc. Manual editing and digitising is therefore necessary to extract the full topology of the proposed road network complete with coordinates of junctions.
- The current system does not permit checks on phonetics and this is an issue for similar sounding names (e.g., Bailey, Baylee, Bayley, Baylea). Similar or 'like' names (e.g. Whyte and White) are not allowable under policy guidelines as they can cause confusion for applications such as emergency services

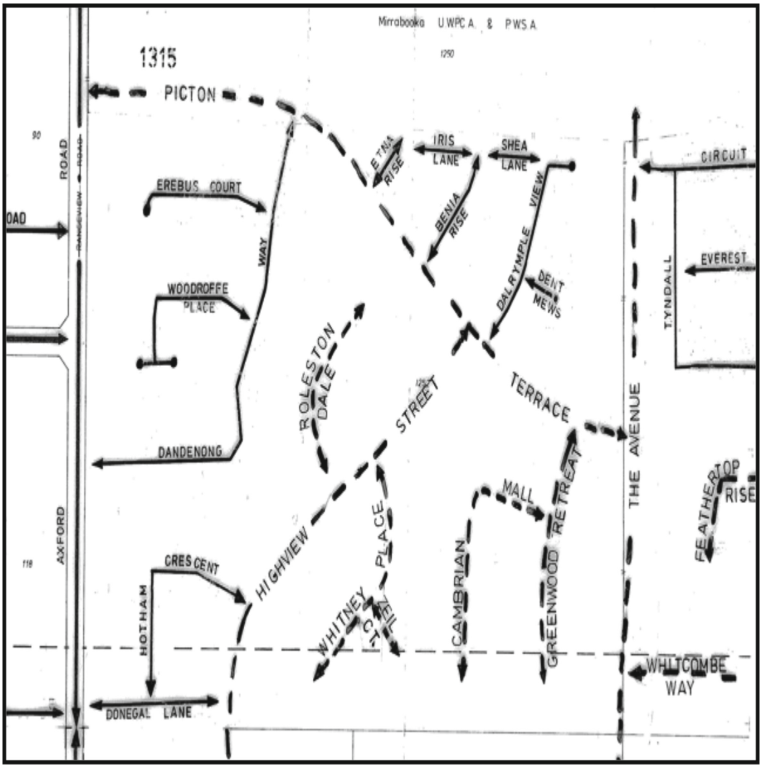

Fig. 1. Hardcopy road network plan with road name application.

dispatch. Similarly, the same road name or a similar sounding road name is not permitted within close proximity.

- Where an extension to an existing road occurs or where a road 'type' (e.g. cul-de-sac, highway) changes, the current system is unable to return an extension to a road name or change to road suffix, respectively.

4 Approach

Figure 2 shows the different phases in the land transaction process from knowledge acquisition to final feedback. Data is extracted from the various databases in formats such as html, JSON, csv and xml and converted to RDF. Ontologies in OWL are created from database schema and models in the interactive GUI based Protégé environment. Rules are generated in SWRL by an expert. Once the system has been developed, the data, ontologies and rules can be used in the runtime environment Jena with a rule engine by a developer to process road changes.

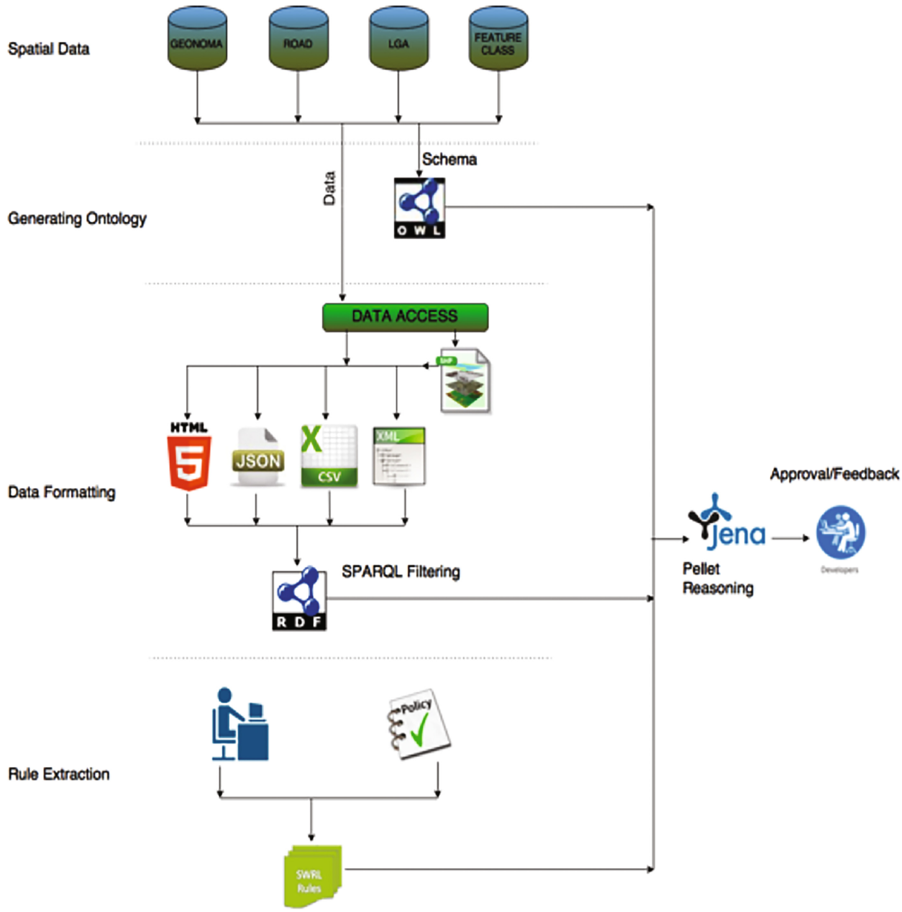

Fig. 2. Data integration/reasoning architecture.

4.1 Knowledge Acquisition

Knowledge acquisition was used to extract, structure and organise knowledge from policy documents, data dictionaries and by interviewing subject matter experts. This knowledge was then used to create the road naming rules. A combination of knowledge acquisition methods are used including organising explicit knowledge and eliciting tacit knowledge.

1. Organising explicit knowledge
   General procedures for spatial transactions are mentioned in policy documents, standards and dictionaries. These documents were reviewed to build the general rules on process. Establishing rules from explicit knowledge uses the following strategies:
   (a) Rules sourced from policy standards:

- A road name cannot be used if it already exists within a 10 Km radius of the new road in city areas or 50 Km in rural areas.
- A road name may not be used more than 15 times in the State of Western Australia.

(b) Rules sourced by accessing data dictionaries:
- Discriminatory or derogatory names are not allowed.
- A name in an original Australian Indigenous language will be considered for a new road name with reference to its origin.

2. Eliciting tacit knowledge
Currently polices and standards do not completely capture the human knowledge required for geographic naming processes. This makes it difficult to translate procedural knowledge into a computer-understandable form. In order to overcome this problem, knowledge elicitation techniques have been used to elicit procedural knowledge by conducting interviews, focus groups and observations etc.

(a) Rules sourced by interviewing subject matter experts:
- A name must not relate to a commercial business trading name or non-profit organisation
- A name must not sound like an existing name
- A name with the suffix type ‘place’ or ‘close’ cannot be assigned to a road greater than a specified length (200 m)
- A historical name, such as ANZAC, cannot be used
- A name with road type ‘rise’ can only be used for roads that have elevation or are at an incline
- Abbreviated names derived from the suburb name are not acceptable for new road names

With the current traditional naming process, satisfying the rules identified above is time consuming because of the back-and forth process between developer and approver. As an example, from a process perspective, when a land developer or local authority requests a new road name within a development site, a spatial validation process is run to test whether the proposed name:

- is already in use in the local authority and if so, whether it is within 10 Km of the new site; and
- has already been used 6 times within metropolitan area and 15 times across the State.

In addition to policy rules, subject matter experts use broader contextual knowledge when determining if a new road name is valid. For example, during the approval process experts check the scope for the proposed subdivision within the wider development site to avoid subsequent changes resulting from incorrect initial decisions.

Figure 3 presents a further example of where expert knowledge in the road naming process, from initial application to final approval, is required. During the negotiation phase with the land developer, documents are transferred back and forth between both parties; each making changes to a paper plan by way

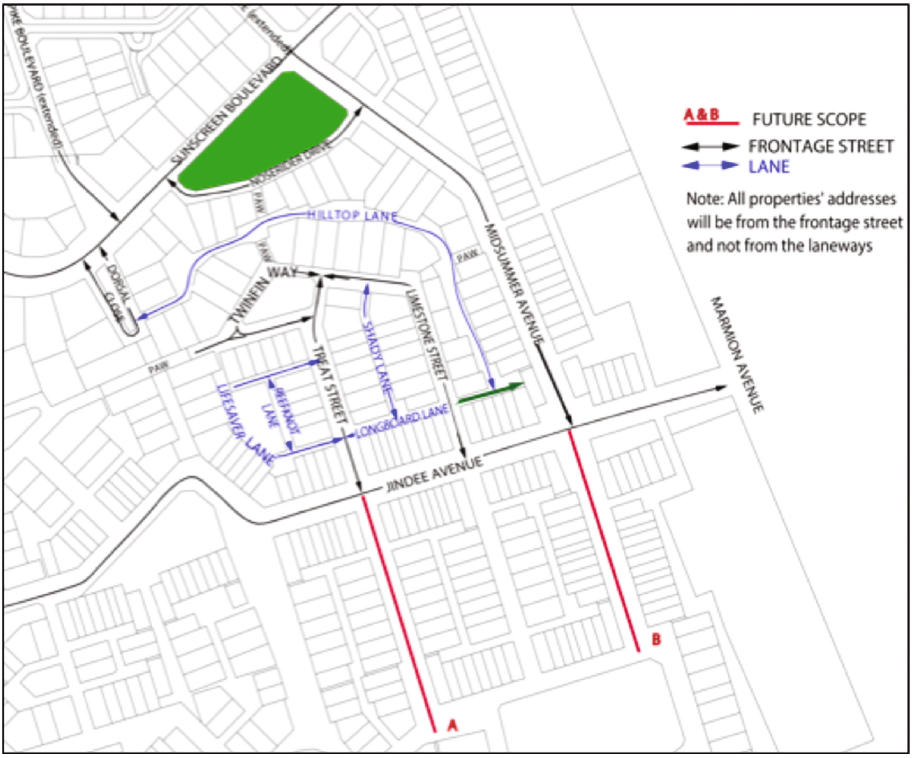

Fig. 3. Road naming process in Jindalee - City of Wanneroo Western Australia. (Color figure online)

of communication. The following notes, written by Landgate to the developer, illustrate typical negotiations (See Fig. 3):

- Jindee Avenue: The road type is suitable, however the name Jindee is not. Apart from sounding similar to the suburb name, this is also an abbreviated name derived from the suburb name and is not acceptable. A replacement name is required.
- Limestone Street and Twinfin Way: The street is continuous so one street name can be used for this street.
- Noserider Drive: The name is suitable, however the road type Drive is not (as this road is adjacent below in this case) to a future open space then relevant types are Way, Vista, View, or if it shaped like a crescent, then Crescent can be used).
- Longboard Lane: The name complies with policy, however it is too long a word for that road. Also a portion of the extent is a part of Hilltop Lane (mentioned in green). A short name with its origin is required. Alternatively, the developer can hold the name Longboard for future use when a long road name is needed in the vicinity.

- Lifesaver Lane: the name is suitable, however it appears that there will be a third entry off Twinfin Court. Clarification of this will be necessary and an additional name for a portion (i.e. the northern east/west portion) will be needed.
- Midsummer Avenue and Treat Street: extensions are suitable because there are possibilities for the future development. The roads on the south side of Jindee Avenue (A & B) are currently unnamed as they are part of a later development stage.

4.2 Ontology Development

Ontology is one of the technologies listed within the Semantic Web Technology Stack [12]. Although it is used within the information sciences the term ontology has its origin in philosophy and is the study of being or existence [13] and it has been considered to be a branch of metaphysics looking at the nature of being. It is from these origins that the disciplines of Computer Science and Information Science borrow ontology and now it is used as a way to represent knowledge [13].

The term ontology is used with various different meanings and at different points in time these different definitions can be contradictory01 [14]. Bergmen [15] listed more than 40 different terms that are used which could all be called types of ontologies or at least ontological frameworks. With this number of terms often used in reference to ontologies it is quite understandable that there may be misunderstandings as well as misinformation about ontologies. Table 1 shows some of the various names that could loosely mean ontology. It is crucial that when using the term ontology it is clearly laid out how it is being used. Within

Table 1. Terms used to describe ontology (http://www.mkbergman.com/374/an-intrepid-guide-to-ontologies/).

Tag cloud	Social bookmarking	Topic Maps
Controlled vocabulary	Tags	Concept Maps
Thesauri	Tagging	Synsets
Collaborative tagging	Taxonomy	Glossary
Folk taxonomy	Folksonomy	WordNet
Directory	Classification	Data Reference Model
Subject Map	Categorization	Facets
Semantic Web	RDF	Structure
Cladistics	Metadata	Dublin Core
Markup languages	OPML	Typology
Ontology	XOXO	OWL
Microformats	Subject Trees	Information Architecture
Data dictionary	Phylogeny	

the work here within this paper the term ontology is used to describe the spatial aspects of land data and extract the rules to handle the decision making process.

Once the rules behind both policy standards and business processes are understood, the next step is to generate the ontology model from multiple sources of information. This ontology is developed as a global schema that means that while it works with the Landgate GEONOMA database, it can also be used in conjunction with other databases that link the spatial extent of a road to the road naming process. Figure 4 presents an overview of the generated `Geo_feature ontology` containing classes, data and object properties, and instances. Links show relationships such as domain, range and subClassOf. The ontological components are summarised below.

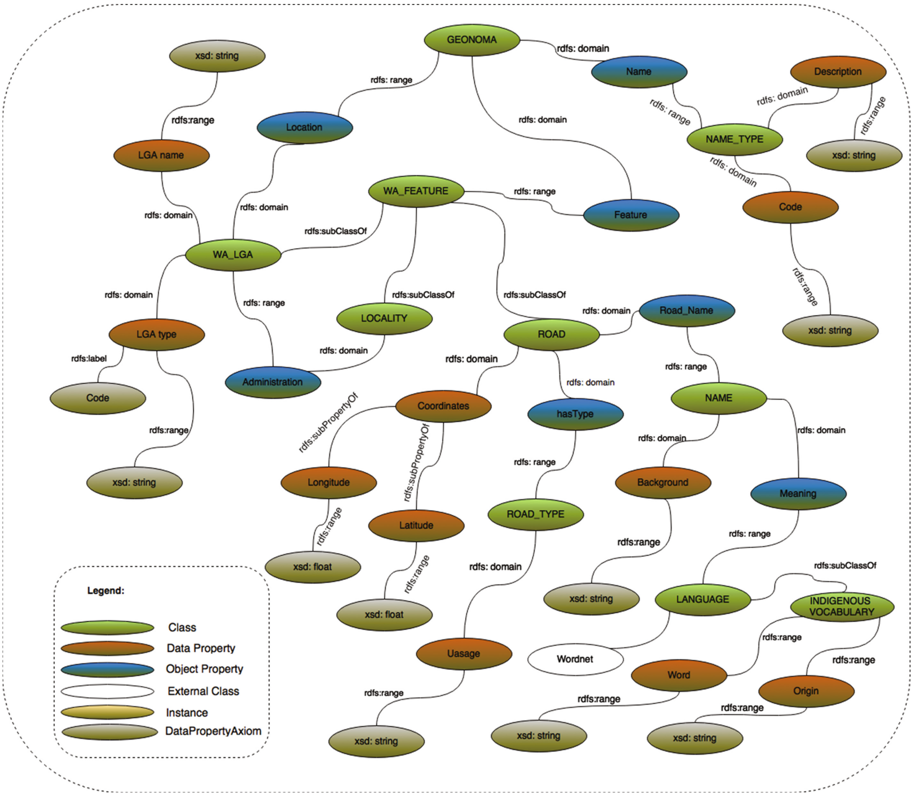

Fig. 4. An overview of Geo_feature ontology.

Geo_feature Ontology. The GEONOMA dataset is exported to XML and then imported into Protégé to help with the ontology generation process. Protégé was chosen, as it is an open source tool with wide community support that supports ontology development and reasoning, and importantly OWL DL, W3C description logic standard. The `Geo_feature ontology` consists of OWL classes, data and object properties, and individuals and is expressed in the form of OWL-2. Each OWL class is associated with a set of individuals. Object properties link individuals of one class to other class individuals. Data properties link one individual to its data values. Value constraints and cardinality constraints are used to restrict the attributes of the individual. For example each ROAD instance much have only one ROAD_TYPE through an object property link. Figure 5 shows the relationships between class instances. An example for a ROAD_TYPE instance is shown at bottom right. It has property restrictions handled by cardinality constraints. Each instance must have information about its type, description and whether it is a cul-de-sac or an open ended road type. Typically, further work is required to create the full semantics in the ontology. `Geo_feature ontology` comprises of more than one ontology such WordNet ontology and homophone ontology. All semantic relationships (links) between data components are needed because mapping from datasets directly is not adequate to explain the full model [16]. For example, every instance of ROAD, LGA and LOCALITY has a link with an instance of GEONOMA. Similarly every ROAD has a link with LGA and LOCALITY. These are inferred in Protégé by invoking the OWL-DL rule reasoner.

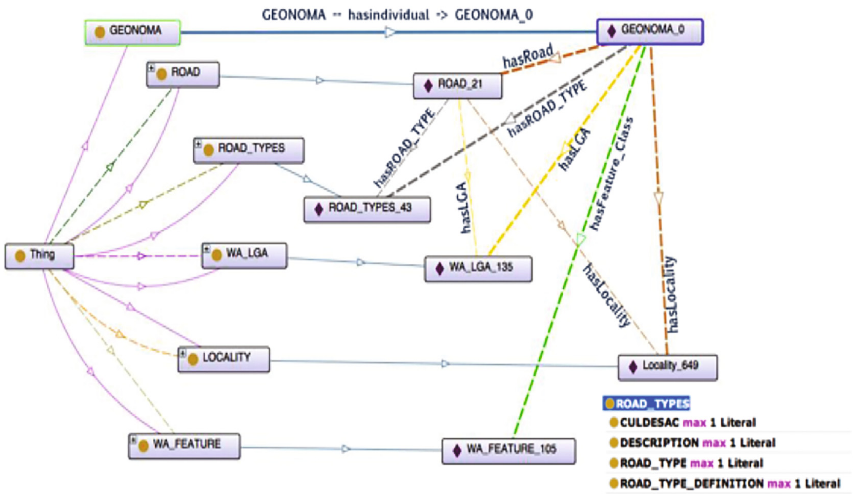

Fig. 5. OntoGraf representation for classes and instances.

Ontological Classifications and Spatial Relations. The resulting `Geo_feature ontology` represents the spatial relationship between several datasets including the road network, local government authority boundaries, locality and language. These datasets combined are used in the road name approval process and checked for constraints. The spatial relationship distinction is mainly based on source datasets. However, from a realistic viewpoint, these source datasets can only supply certain details relating to a feature name. To make it more meaningful there is a need to add additional vocabularies such as the Australian indigenous language dictionary and the WordNet ontology. The Australian indigenous language dictionary gives insight into the Australian indigenous naming specifics and WordNet ontology resembles a thesaurus of English words. By adding these we can check the meaning of a name and whether or not it complies with the chosen road-naming theme. To process a road request the road structure needs to be examined. By adding road coordinates it is possible to check where the proposed road will be actually developed.

4.3 Rule Development

Figures 4 and 5 show several relations between spatial datasets, such as the link between road and locality. Many of these relationships are inferred by the rule-based mechanism automatically from constraints, axioms and links defined in the ontology, thereby reducing the need for manual specification for all instances. The Pellet reasoner is used to infer decisions from these SWRL rules in Protégé. These inferred decisions are then communicated to the developer as a feedback. More complex, nested conditions can be handled by Boolean operators in SWRL rules are executed with the rule engine [17].

```xml
<owl:NamedIndividual rdf:about="http://www.semanticweb.org/ontologies/LandInfo#GEONOMA_0">
    <rdf:type rdf:resource="&GEONOMA;GEONOMA"/>
    <GEONOMA:FEATURE_NUMBER rdf:datatype="&xsd;integer">100062011</GEONOMA:FEATURE_NUMBER>
    <GEONOMA:NAME_TYPE rdf:datatype="&xsd;string">Approved Name</GEONOMA:NAME_TYPE>
    <GEONOMA:FEATURE_STATUS rdf:datatype="&xsd;string">CURRENT</GEONOMA:FEATURE_STATUS>
    <GEONOMA:DISPLAY_NAME rdf:datatype="&xsd;string">Chancery CT</GEONOMA:DISPLAY_NAME>
    <GEONOMA:FULL_NAME rdf:datatype="&xsd;string">Chancery CT</GEONOMA:FULL_NAME>
    <GEONOMA:SECURITY_CLASS rdf:datatype="&xsd;string">Complete Access</GEONOMA:SECURITY_CLASS>
    <GEONOMA:LOCALITY_NAME rdf:datatype="&xsd;string">Forrestfield</GEONOMA:LOCALITY_NAME>
    <GEONOMA:LGA_NAME rdf:datatype="&xsd;string">Kalamunda, Shire of</GEONOMA:LGA_NAME>
    <FEATURE_CLASS rdf:datatype="&xsd;string">Public Road</FEATURE_CLASS>
    <CATEGORY rdf:datatype="&xsd;string">Road</CATEGORY>
    <hasLGA rdf:resource="&LGA;WA_LGA_135"/>
    <hasFeature_Calss rdf:resource="&Ontology1445038673;WA_FEATURE_105"/>
    <hasRoad rdf:resource="&Ontology1445238488;ROAD_21"/>
    <hasLocality rdf:resource="http://www.semanticweb.org/ontologies/LandInfo#Locality_649"/>
    <hasROAD_TYPE rdf:resource="http://www.semanticweb.org/ontologies/LandInfo#ROAD_TYPES_43"/>
</owl:NamedIndividual>
```

Fig. 6. Source data in RDF format.

4.4 Data Formatting/Conversion

Once the ontology and rules have been developed the next stage is to access the source datasets to reason with the ontologies. To make this happen it is necessary to convert the source dataset into RDF triple format. In this way all data are accessible in one common format and ready for initial reasoning [18]. There are many data conversion and integration tools (Karma, MASTRO, OpenRefine and TripleGeo) that can be used for this conversion. MASTRO has been shown to be a successful Ontology-Based Data Access (OBDA) system through a series of demonstrations [19–23]. It can be accessed by means of a Protégé plugin. The facilities offered by Protégé can be used for ontology editing, and functionalities provided by the MASTRO plugin can be used to access external data sources. Openrefine (http://openrefine.org/) is used to convert data to RDF format. Spatial information from a shape file can converted into RDF triples [24] (https://github.com/GeoKnow/TripleGeo). Figure 6 shows an RDF instance. Having the data instances in RDF format, Apache Jena, with the help of MAVEN repositories is used to link all the ontologies, instances and rules at runtime.

5 Process/Operation

5.1 System Implementation

Figure 7 shows the runtime system architecture, which has been implemented using Jena in Java. The ontology repository consists of multiple ontologies

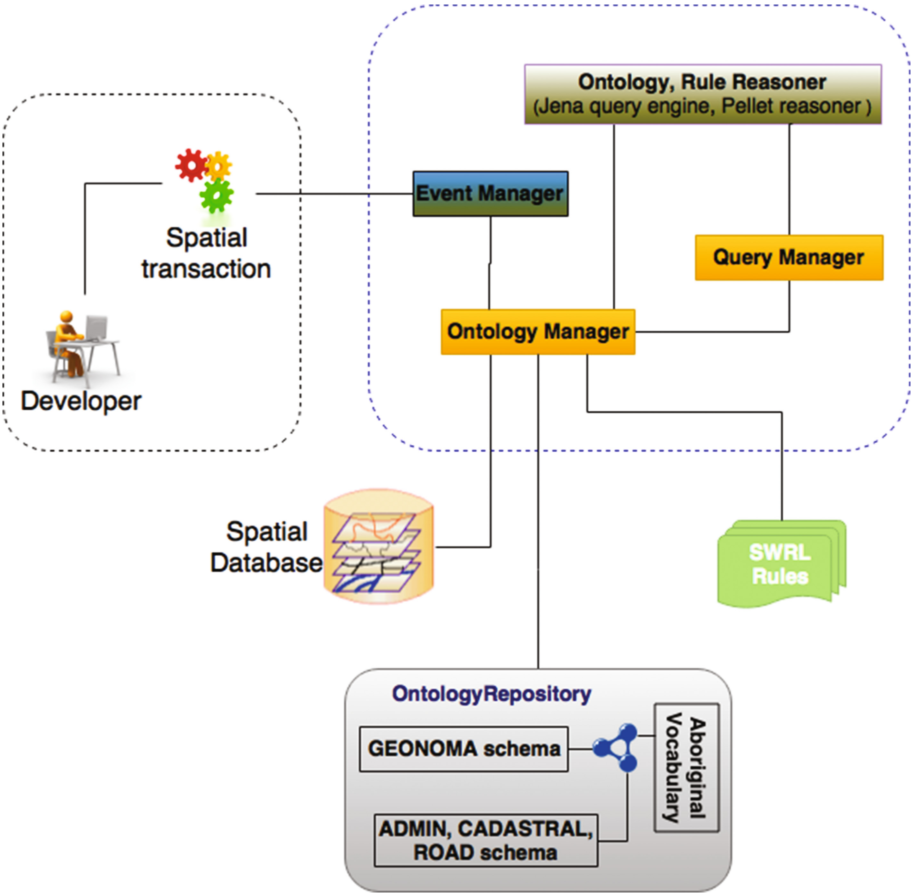

Fig. 7. System architecture.

derived from the data schema, data individuals, and rules, as well as non-specific ontologies such as Aboriginal vocabularies. The event manager collects the land transaction information and supports the ontology manager to infer the information relevant to that application. For example, if the application relates to a new subdivision, then it will gather the details spatially related to that land area, or if the proposed road name relates to a road name change, then it will gather information related to naming from the policy. The Ontology Manager collates the land information from the spatial database into the knowledge base.

5.2 Reasoning

The initial stage of reasoning is carried out in Jena with the Pellet OWL reasoner that checks the logical consistency of the model, processes the individuals (current, approved and proposed roads), infers new information including links and relationships, and updates the model with the inferred information. Through consistency checking, the system confirms whether or not any contradictory facts appear within the ontology. For example, the domain and range constraints on the feature relation: GEONOMA Features: `Feature_Class`. Constraints on the relation mean that GEONOMA has features, which come under only one of the `Feature_Class` categories. The reasoner will throw relevant errors if any ontological inconsistency appears given the proposed roads, for example if an instance of GEONOMA is linked to an instance of a ROAD and missing any property restriction relations.

Similarly, assigning an individual to two disjointed categories such as LGA and Locality will make the ontology inconsistent. Consider the case where every GEONOMA instance is represented with the ROAD feature type; it must have at least two coordinates and link to other road instances. This is declared as a necessary and mandatory condition for instances of the ROAD category in the OWL class description. When an individual in OWL satisfies such a condition then the reasoner automatically deduces that the individual is an instance of the specified category.

As well as the reasoning described above, to gather more information additional reasoning is required. Rules are expressed in terms of ontological vocabularies using SWRL. Table 2 shows some examples of implemented rules. As mentioned earlier, in each rule, the antecedent refers the body of the rule and the consequent refers to the head. The head and body consist of a conjunction of one or more atoms. Atoms are stated in the form of C(?R) P(?R,?X), where C and P represent an OWL description and property, respectively. Variables representing the individuals are in the form, for example ?R, where the variable R is prefixed with a question mark. Table 2 shows some examples of rules related to the application.

Table 2. SWRL rules with the action of each of the rules.

No	Purpose	SWRL rules
1	Relate a road link with existing road either directly or thru another proposed road	NEWROAD(?N), ROAD(?R), HASROADLINK(?N, ? R), STATUS(?N, "New"), STATUS(?R, "Existing"), notEqual(?N, ?R) − > isAllowed(?N, true)
		NEWROAD (?N1), NEWROAD (?N2), ROAD (?Old), HASROADLINK (?N1, ? N2), HASROADLINK (?N2, ?R), notEqual(?N1, ?N2), notEqual(?R, ?N2), STATUS (?R, "Existing"), STATUS (?N1, "New"), STATUS (?N2, "Aproved"), − > isAllowed(?R1, true)
2	Find the road duplication within the radios	NEWROAD(?N), ROAD(?R), ROAD_NAME(?N, ?RN2), ROAD_NAME(?R, ?RN1), stringEqualIgnoreCase(?RN2, ?RN1) − > hasRoad(?N, ?R)
3	Check the road name with definite article	DEFINITEARTICLES(?D), NEWROAD(?R1), FULL_NAME(?D, ?DN), ROAD_NAME(?R1, ?RN), stringConcat(?MSG, "Road name cannot contain definite article", ?DN), stringEqualIgnoreCase(?RN, ?DN) − > isAllowed(?R1, ?MSG)
		DEFINITEARTICLES(?D), NEWROAD(?R1), FULL_NAME(?D, ?DN), ROAD_NAME(?R1, ?RN), startsWith(?RN, ?SW), stringConcat(?MSG, "Road name cannot contain definite article", ?DN), stringConcat(?SW, ?DN, "") − > isAllowed(?R1, ?MSG)
4	Check for similar sounding names	NEWROAD(?N), ROAD(?R), LGAS(?N, ?L2), LGAS(?R, ?L1), METAPHONE_ALTERNATE(?N, ?MN2), METAPHONE_ALTERNATE(?R, ?MN1), ROAD_NAME(?N, ?RN1), ROAD_NAME(?R, ?RN2), containsIgnoreCase(?L1, ?L2), notEqual(?RN1, ?RN2), stringEqualIgnoreCase(?MN2, ?MN1) − > maySoundLike(?N, ?R)
		NEWROAD(?N), ROAD(?R), LGAS(?N, ?L2), LGAS(?R, ?L1), METAPHONE_ALTERNATE(?N, ?MN2), METAPHONE_PRIMARY(?R, ?MN1), ROAD_NAME(?N, ?RN1), ROAD_NAME(?R, ?RN2), containsIgnoreCase(?L1, ?L2), notEqual(?RN1, ?RN2), stringEqualIgnoreCase(?MN2, ?MN1) − > maySoundLike(?N, ?R)
		NEWROAD(?N), ROAD(?R), LGAS(?N, ?L2), LGAS(?R, ?L1), METAPHONE_ALTERNATE(?R, ?MN1), METAPHONE_PRIMARY(?N, ?MN2), ROAD_NAME(?N, ?RN1), ROAD_NAME(?R, ?RN2), containsIgnoreCase(?L1, ?L2), notEqual(?RN1, ?RN2), stringEqualIgnoreCase(?MN2, ?MN1) − > maySoundLike(?N, ?R)
		NEWROAD(?N), ROAD(?R), LGAS(?N, ?L2), LGAS(?R, ?L1), METAPHONE_PRIMARY(?N, ?MN2), METAPHONE_PRIMARY(?R, ?MN1), ROAD_NAME(?N, ?RN1), ROAD_NAME(?R, ?RN2), containsIgnoreCase(?L1, ?L2), notEqual(?RN1, ?RN2), stringEqualIgnoreCase(?MN2, ?MN1) − > soundsLike(?N, ?R)
5	Check the road name against road type	NEWROAD(?R1), ROAD_NAME(?R1, ?RN), ROAD_SUFFIX(?R1, ?RT), stringEqualIgnoreCase(?RN, ?RT) − > isAllowed(?R1, "Road name cannot be the same as road suffix")
6	Check the road length to against road types	NEWROAD (?R1), ROAD_SUFFIX(?R1, ?RT), hasLength(?R1, ?200), SameAs (?T1, ?*Close*) − > isAllowed(?R1, true)
7	Check the road name with restricted words	ILLEGALWORDS(?I), NEWROAD(?R1), FULL_NAME(?I, ?IN), ROAD_NAME(?R1, ?RN), startsWith(?RN, ?SW), stringConcat(?MSG, "Road name cannot contain word", ?IN), stringConcat(?SW, ?IN, "") − > isAllowed(?R1, ?MSG)
		ILLEGALWORDS(?I), NEWROAD(?R1), FULL_NAME(?I, ?IN), ROAD_NAME(?R1, ?RN), stringConcat(?MSG, "Road name cannot contain word", ?IN), stringEqualIgnoreCase(?RN, ?IN) − > isAllow

- Rule 1 automatically infers information with the help of a road link between proposed and existing roads from the source dataset with reference to road coordinates and feature id. This rule is necessary as every road needs to link with at least one other road to allow access.
- Rule 2 checks the similar road names within the neighbouring LGA to avoid duplication of road names.
- Rules 3 prevents the definite article being used in the road name.
- Rule 4 checks for similar sounding names within the LGA and neighbouring LGAs to avoid confusion for first responders and visitors to the locality.
- Rule 5 checks the road name against its road type to avoid road naming as road suffix.
- Rule 6 checks road length against road type. Checking the road length for shortest road types (‘Place’, ‘Close’ and ‘Lane’) is necessary to avoid confusion with the preference for road usage.
- Rule 7 prevents the restricted words such as ‘CITY’, ‘SHIRE’ and ‘TOWN’ being used in the road name.

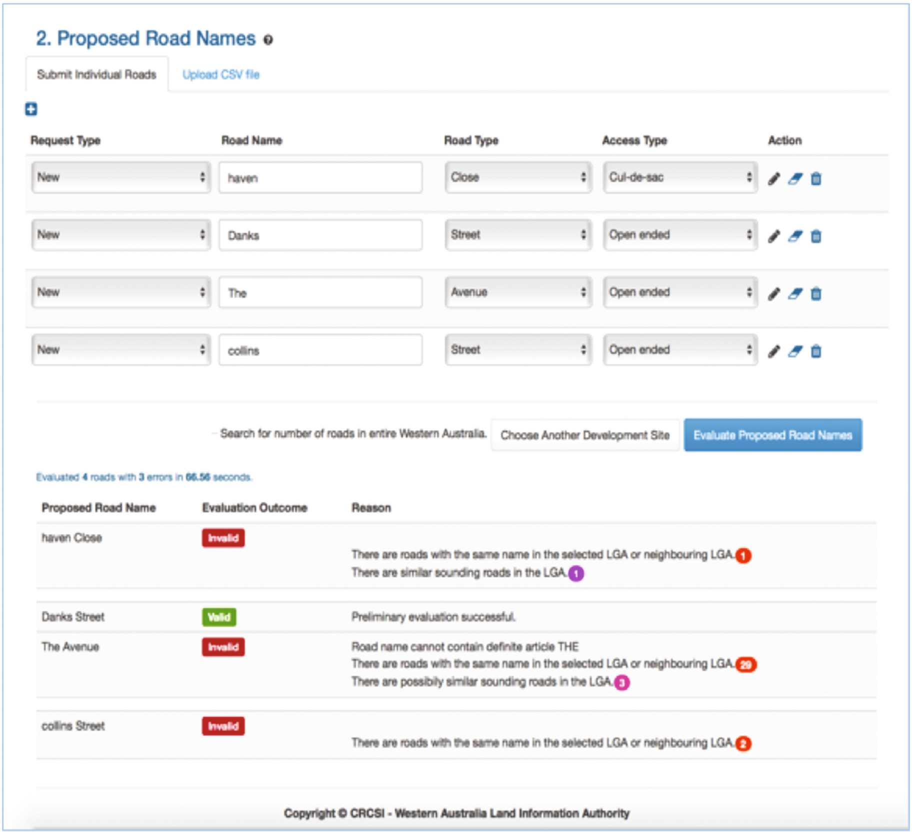

Fig. 8. Automatic spatial transaction application portal.

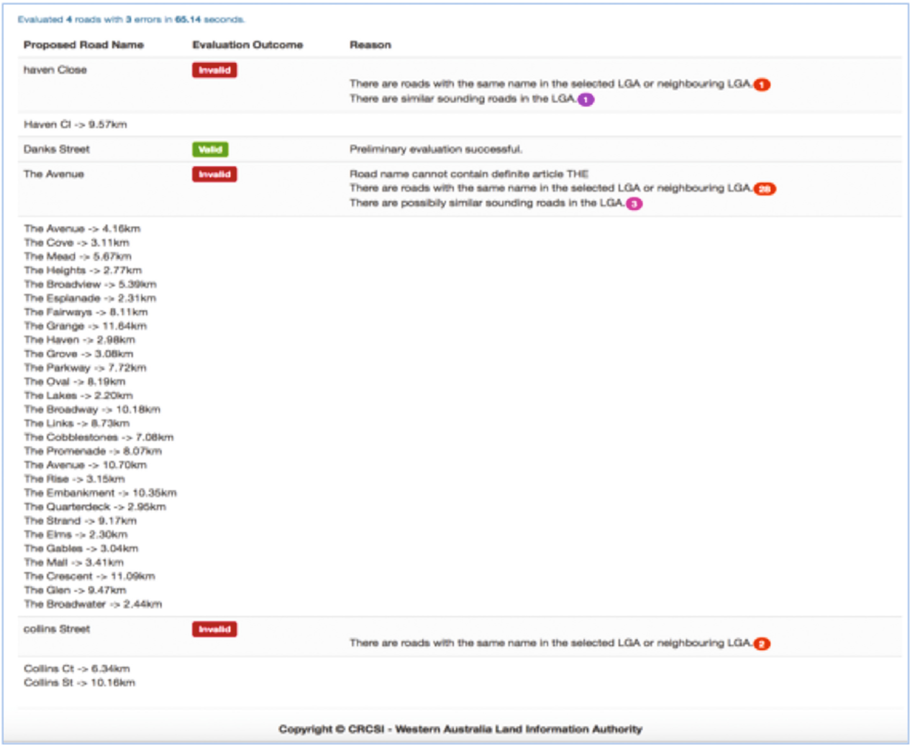

Fig. 9. Automatic spatial transaction feedback.

6 User Interface

The automated spatial transaction application has been developed in this research using Jena in Java. Firstly, the user interface was designed to obtain input from the end-user and, secondly, the rules for geographic naming were built using SWRL and then linked with the Jena rule engine. The Jena engine is used to link all the ontologies, instances and rules at runtime with the help of MAVEN repositories.

The user interface allows the developer to select the development site from the map layout. From the selected site the system buffers either 10 Km or 50 km radius depending on the location of the site. Figure 8 shows the user interface for road naming transactions. Once the developer selects the development site the application then allows the developer to enter new road details. In many cases the development site will require the approval of several new roads. For this reason, the application provides an upload facility for developers to lodge a CSV file format to save time. The system is designed so that the road names contained in the CSV file are assessed simply by pressing the evaluate button. If any of the given road information does not comply with the rules, then the application provides feedback to the user accordingly. An example is shown in

Fig. 9. Once all submitted roads comply with the rules, then the system requests the developer's details and all supporting documents as evidence for further land development proceedings.

7 Conclusion

Traditional methods in spatial transaction mainly involve manual assessment of applications that cause delay, as a consequence of a back and forth process is being required. Human involvements are very time consuming, expensive and may trigger errors. This emphasized the importance of automation that reduces the manual overhead by extracting expert knowledge for such critical spatial transactions.

This paper proposes a Semantic Web solution for automating the decision making process for spatially related transactions. Examples of such transactions are approvals for new roads names and property address change. The method develops a `Geo_feature ontology`, which comprises knowledge of roads and constraints, axioms and rules extracted from sources such as experts, policy, geometry and past decision documents. The method shows how ontologies and rules are manipulated with reasoning techniques to infer new information.

Semantic Web techniques are used as the solution because it allows the ontologies and rules to be published in RDF and made available for other application domains. For example, similar processing is envisaged for points of interest (bridges, parks), and the reconciliation of addresses. These ontologies can be used in other jurisdictions for similar transactions or other application domains.

This method has proven successful for the process that involves simple spatial queries, such as a request for new road name approval and updating existing road features. The User interface facilitates the developer and government agencies in naming proper road names by providing feedback with map layout that helps the developer to understand road name non-compliance faults in visual form. More rules and relationships with existing ontology elements are being developed as further examinations are carried out into the datasets and business rules. Future work is also examining reasoning over other information that can be used to aid the approval process. For example an approver may use aerial photography to check for the presence of vegetation, as the removal of trees may need approval, and digital elevation maps used to determine if the proposed roads are viable.

Acknowledgement. The work has been supported by the Cooperative Research Centre for Spatial Information, whose activities are funded by the Australian Commonwealth's Cooperative Research Centres Programme. The authors extend their thanks to Landgate for providing the example datasets for the case study and subject matter experts for rule formulation.

References

1. Varadharajulu, P., Saqiq, M.A., Yu, F., McMeekin, D.A., West, G., Arnold, L., Moncrieff, S.: Spatial data supply chains. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **40**(4), 41 (2015)

2. Lee, B.T., Fischetti, M.: Weaving the web: the original design and ultimate destiny of the world wide web by its inventor (1999)
3. Gupta, S., Knoblock, C.A.: A framework for integrating and reasoning about geospatial data. In: Extended Abstracts of the Sixth International Conference on Geographic Information Science (GIScience) (2010)
4. McMeekin, D.A., West, G.: Spatial data infrastructures and the semantic web of spatial things in Australia: research opportunities in SDI and the semantic web. In: 2012 5th International Conference on Human System Interactions (HSI), pp. 197–201. IEEE (2012)
5. Shadbolt, N., Berners-Lee, T., Hall, W.: The semantic web revisited. IEEE Intell. Syst. **21**(3), 96–101 (2006)
6. Millard, E.: The semantic web could enable even greater access to information. Promise of a better Internet. Teradata Mag. online (2010)
7. Devaraju, A., Kuhn, W., Renschler, C.S.: A formal model to infer geographic events from sensor observations. Int. J. Geogr. Inf. Sci. **29**(1), 1–27 (2015)
8. Yu, L., Liu, Y.: Using linked data in a heterogeneous sensor web: challenges, experiments and lessons learned. Int. J. Digit. Earth **8**(1), 17–37 (2015)
9. Reitsma, F.E.: A new geographic process data model. Ph.D. thesis (2005)
10. O'Connor, M., Knublauch, H., Tu, S., Grosof, B., Dean, M., Grosso, W., Musen, M.: Supporting rule system interoperability on the semantic web with SWRL. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, pp. 974–986. Springer, Heidelberg (2005). doi:10.1007/11574620_69
11. Segaran, T., Evans, C., Taylor, J., Toby, S., Colin, E., Jamie, T.: Programming the Semantic Web. O'Reilly Media Inc., Sebastopol (2009)
12. Hazaël-Massieux, D.: The semantic web and its applications. In: World Wide Web Consortium, SIMO - The Semantic Web and its applications, November 2003
13. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? Int. J. Hum. Comput. Stud. **43**(5–6), 907–928 (1995)
14. Noy, N.F., McGuinness, D.L., et al.: Ontology development 101: a guide to creating your first ontology (2001)
15. Bergman, M.: An intrepid guide to ontologies. ai3::: Adaptive information (2007)
16. Ghawi, R., Cullot, N.: Building ontologies from multiple information sources. In: 15th Conference on Information and Software Technologies (IT2009), Kaunas (2009)
17. Powell, J.: A Librarian's Guide to Graphs, Data and the Semantic Web. Elsevier, Waltham (2015)
18. Broekstra, J., Kampman, A., Harmelen, F.: Sesame: a generic architecture for storing and querying RDF and RDF schema. In: Horrocks, I., Hendler, J. (eds.) ISWC 2002. LNCS, vol. 2342, pp. 54–68. Springer, Heidelberg (2002). doi:10.1007/3-540-48005-6_7
19. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Poggi, A., Rodriguez-Muro, M., Rosati, R., Ruzzi, M., Savo, D.F.: The mastro system for ontology-based data access. Semant. Web **2**(1), 43–53 (2011)
20. Poggi, A., Rodriguez, M., Ruzzi, M.: Ontology-based database access with DIG-Mastro and the OBDA Plugin for Protégé. In: Proceedings of the 4th International Workshop on OWL: Experiences and Directions (OWLED 2008 DC), vol. 496 (2008)
21. Savo, D.F., Lembo, D., Lenzerini, M., Poggi, A., Rodrıguez-Muro, M., Romagnoli, V., Ruzzi, M., Stella, G.: Mastro at work: experiences on ontology-based data access. In: Proceedings of DL, vol. 573, pp. 20–31 (2010)

22. Rodriguez-Muro, M., Lubyte, L., Calvanese, D.: Realizing ontology based data access: a plug-in for Protégé. In: IEEE 24th International Conference on Data Engineering Workshop, 2008 (ICDEW 2008), pp. 286–289. IEEE (2008)
23. Zhang, Y., Chiang, Y.-Y., Szekely, P., Knoblock, C.A.: A semantic approach to retrieving, linking, and integrating heterogeneous geospatial data. In: Joint Proceedings of the Workshop on AI Problems and Approaches for Intelligent Environments and Workshop on Semantic Cities, pp. 31–37. ACM (2013)
24. Patroumpas, K., Alexakis, M., Giannopoulos, G., Athanasiou, S.: TripleGeo: an ETL tool for transforming geospatial data into RDF triples. In: EDBT/ICDT Workshops, pp. 275–278 (2014)