19th International Computer Science and Engineering Conference (ICSEC)
Chiang Mai, Thailand, 23-26 November, 2015

# An Applied Ontology: a Semantic Query Builder for Health GIS System

Theeradol Boonprapasri
Geographic Information System Division
Bangkok Metropolitant Administration
173 Dinsoa Road,Pranakorn, Thailand
theradol@bangkok.go.th

Gridaphat Sriharee
Department of Computer and Information Science
King Mongkut's University of Technology North Bangkok
1518 Pracharat Road, Wongsawang, Bangsue, Thailand
gridaphat.s@sci.kmutnb.ac.th

***Abstract*— GIS information is important for e-Health surveillance because such information is required for monitoring, exchanging and shared across health GIS systems. However, interacting with the health GIS system is difficult for the users who may have less knowledge about GIS systems. This paper proposes a semantic query builder that provides a query service and can interact with health GIS systems. The semantic query builder receives the semantic query and translates such queries into a SQL query to allow querying information from databases stored in the health GIS system. It delivers GIS tuple data – the data from the map service is provided by the health GIS system. With the proposed semantic query builder, the users who may have less knowledge of SQL and GIS system are able to specify the query flexibly which meets their requirement.**

***Keywords—Health GIS, GIS system, query builder, ontology***

## I. INTRODUCTION

E-HEALTH GIS system provides information for e-Health monitoring services. The system maintains geographical information of GIS and Health information of a civilization. The e-Government service system may exchange such information with the e-Health GIS system for monitoring, and providing services to a civilization in crisis situations such as disease or earth disaster. Thus, the e-Health GIS information needs to be in real-time on demand and should be provided in some formats that support exchanging, parsing and sharing across the involving systems.

The health department of Bangkok provides 68 health centers to support the health care service to the population of Bangkok. Each health center provides health care services to the people in each particular area. The mission of the department also includes promoting health support and disease control. The public health staff in each health center will survey health information of all members in a family. Also, they will monitor and give care at home for some particular patients. Some evaluations of the well-being of each individual person, family and community are analyzed using the gathered health information. The health GIS system also supports planning for home visits and survey. The computer experts of the GIS system may also need to support the public health staff when querying information from the health GIS system. If the requirement changes frequently, this may cause inconvenience for both participants.

A growing of research has employed mobile technologies and geographic information systems (GIS) for enhancing health care and health information systems, but there is yet a lack of studies of how these two types of systems are integrated together [1]. Many research proposed the use of ontology as intermediate information model to retrieve the GIS information. The integration of the ontological information and the database information of GIS system is focused. However, the implementation of those systems is still difficult and inflexible for the users who may have less knowledge of both GIS system and database query. This paper has two contributions:

(i) to present a transformation model according to the structure and query language of the ontology-based information and database information of the health GIS system, and

(ii) to demonstrate an implementation of the proposed system - a semantic query builder which deploys both ontology-based information and spatial information for query specifying.

A semantic query builder supports the public health staff to specifying the query and to retrieve information from the health GIS system. The semantic query builder formulates the semantic query and such query is transformed into the SQL query for the health GIS system. The semantic query represents an attribute-based query and spatial query. The query results are geospatial data that represents the locations of the patients' home for health care services. Visual representation of such result can be implemented for mobile applications and desktop applications that both interact with the map service on the health GIS system.

Section II provides the background knowledge of health information and the authors' motivation to conduct this research. Section III presents the design of semantic query builder and its interaction with the health GIS system. Section IV is the description of the health information model and the additional ontology used in the semantic query builder. Section V is the description of the development of the

978-1-4673-7825-3/15/$31.00 ©2015 IEEE

proposed semantic query builder. Section VI is related works and section VII is conclusion and discussions.

## II. BACKGROUND AND MOTIVATION

Public health preparedness and surveillance are important for a big city. Bangkok has a recorded population in the registration system of about 5.7 million people and the city area covers 1,568.737 square kilometers. The health department of Bangkok is responsible for 68 health centers which provide health service to the population. To monitor disease situation and surveillance, the public health staffs survey health information of the people in each year (see Fig.1). The health information consists of three parts: family health information, home visit health information, and home care information. The family health information consists of both demographic data and economic data of the persons in family.

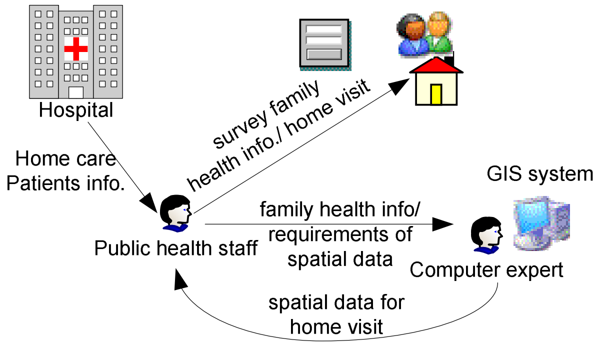



Fig. 1. The participants

The health department of Bangkok monitors non-communicable diseases (NCDs) and communicable diseases. The NCDs includes 6 types such as curable disease, long-term chronic disease, long-term with mild disabilities, long-term extreme disabilities, terminally ill with independence and special groups. They are obligated by the hospital to give home care service to the patients. For surveying and monitoring, planning of home visit locations is specified using the GIS system. The public health staff of each health center provides requirements to the computer experts of the GIS department of Bangkok. The computer experts create the database for the health GIS of the health department for Bangkok. They maintain the monitored and surveyed health information from 68 centers. They also provided the queries for spatial requirements for monitoring and home visit planning specified by the public health staff (see Fig. 2). When the requirements change, it is difficult to communicate to each other and it takes time to deliver spatial data because the computer expert has to prepare the new software to deliver such data. In addition, if the exchange of information between the public health staff and the computer expert is not clear, the report from the health GIS system may give inaccurate information.

From these problems of data management, this paper proposes a system called the semantic query builder that supports the public health staff to specify the query with the health GIS system directly without exchanging the requirement to the computer expert. Doing so, the public health staff will be free to define their own queries.

In this paper, the ArcGIS system is adopted for integration with semantic query builder. In ArcGIS system, the GIS information consists of the data in many layers and several maps are overlaid on each other. The users are able to query the spatial data using the semantic query builder. The semantic query builder returns spatial data obtained from the GIS system.

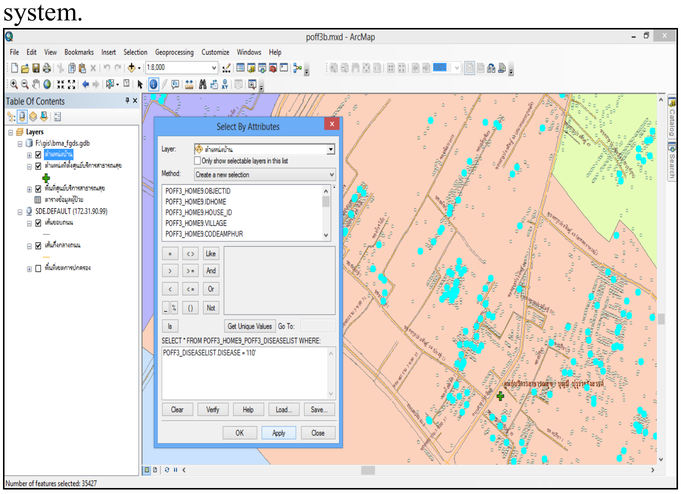



Fig. 2. ArcGIS application for querying spatial data

## III. THE SEMANTIC QUERY BUILDER

GIS-database query and spatial query, will be supported by the proposed query builder for specifying their query. When using the proposed semantic query builder, users can specify their query without awareness of structure changes in the GIS database, and they are not required to have much knowledge of the SQL language and GIS system.

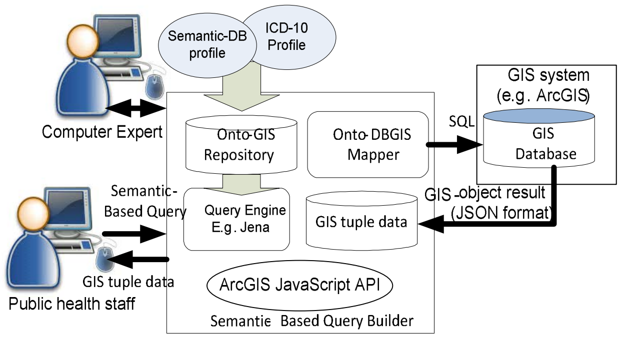



Fig. 3. The semantic query builder

Fig. 3 depicts the semantic query builder, its components and integration with the health GIS system (e.g. ArcGIS). The computer expert who has the knowledge about ontology creates the semantic-DB profile using a tool (e.g. Protégé [2]). The semantic-DB profile is the profile that is transformed from the database schema of the health information of health

GIS system. The semantic-DB profile is described by OWL [3] and it is stored in the Onto-GIS repository. The query engine (e.g. Jena) of the semantic query builder reads the Semantic-DB profile and generates the user interface for specifying the semantic query. Additional ontology is available to support specifying query. Here, ICD-10 profile is provided and it is the description of the NCDs disease.

The semantic query is transformed into the SQL query which will be the input of the health GIS system. Doing so, the Onto-DBGIS Mapper is implemented. The Onto-DBGIS generates the SQL script from the input specified in the semantic query builder. It uses the mapping rule specifying the mapping between ontological information and the database information (see Table I). The GIS system responses to the semantic query builder with GIS object results (JSON format) which will be transformed and stored in the GIS tuple data database as the tuple data (records). The semantic query builder uses the ArcGIS JavaScript API to connect the GIS server for map rendering. The tuple data can be exported in terms of the CSV (Microsoft Excel) data. The semantic query builder connects to the map service of ArcGIS and performs map rendering using the tuple data.

## IV. ONTOLOGY-BASED HEALTH INFORMATION

Ontology provides formal specification that describes the entities and their relations in particular domain. The ontology model is mostly used as shared information for multiple systems. There are numerous efforts that proposed the ontologies and that are available on the web with the aim to share, exchange, and link the data from various systems together. In the health domain, there are a number of ontologies provided such as ICD-10 [4] and ICD-11 [5]. The purpose of using such information is sharing to aid common understanding of disease in a particular domain. In this paper, the ICD-10 is selected because the health department of Bangkok uses such information in present for referencing the health care service obligated by 68 health centers.

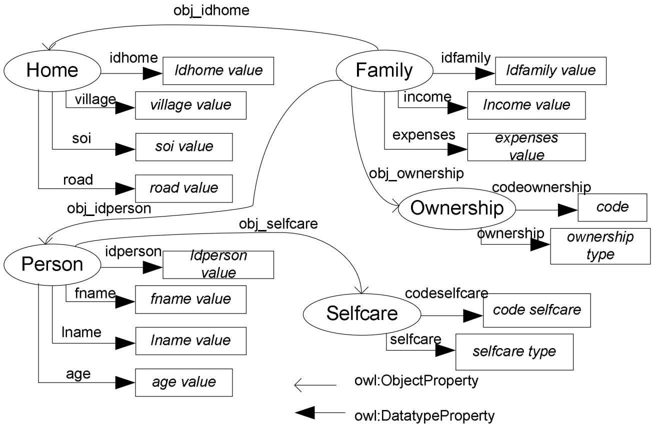



Fig. 4. An ontology-based health information model

Fig. 4 depicts an example of the ontology-based health information model. The model is the description in the semantic-DB profile. The model is created from the database schema of the health GIS system. In this example, some concepts are defined such as *home, family, selfcare*, *ownership* and *person*. The home concept is the information of home location. The family consists of the economic data and ownership information. The person concept is the personal health data containing questions of how the person takes care him/her self. The ownership represents the type of the house of the family. The property of each concept may be represented by the data value or may link to another concept.

This paper adopts the technique for transformation from database schema into the ontology from [6]. The mapping details are shown in Table I. Fig. 5 depicts an example of the OWL profile (semantic-DB profile) of the health information model. The profile is read by the semantic query builder.

TABLE I. MAPPING BETWEEN DATABASE AND OWL COMPONENT

| Database component | OWL component |
|---|---|
| Table | Class |
| Attribute with value | Datatype property with specified domain and range class. |
| Primary key attribute | Functional datatype property with specified domain class. |
| Foreign key attribute | Object property with specified domain and range class. |

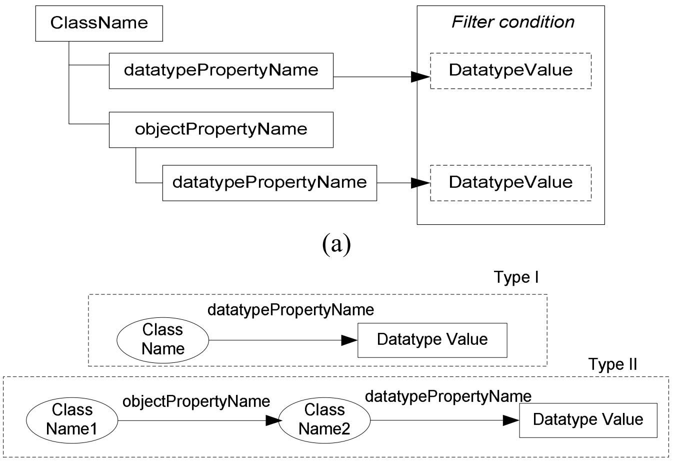
```
<owl:Class rdf:ID="Family"/>
<owl:Class rdf:ID="Home"/>
<owl:Class rdf:ID="Person"/>
<owl:Class rdf:ID="Ownership"/>
<owl:Class rdf:ID="Selfcare"/>
<owl:ObjectProperty rdf:ID="obj_idhome">
  <rdfs:domain rdf:resource="#Family"/>
  <rdfs:range rdf:resource="#Home"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID=" obj_ownership">
  <rdfs:domain rdf:resource="#Family"/>
  <rdfs:range rdf:resource="#Ownership"/>
</owl:ObjectProperty>
<owl:DatatypeProperty rdf:ID="expenses">
  <rdfs:domain rdf:resource="#Family"/>
  <rdfs:range rdf:resource="&xsd;int"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="income">
  <rdfs:domain rdf:resource="#Family"/>
  <rdfs:range rdf:resource="&xsd;int"/>
</owl:DatatypeProperty>
```

Fig. 5. Semantic-DB profile of health information model

## V. THE DESIGN OF SEMANTIC QUERY BUILDER

### *A. The design of user interface*

The semantic query builder reads the ontology profiles: Semantic-DB profile and ICD-10 profile, and it displays such information on the user interface. Designing the user interface is important because the semantic query builder must read the OWL-based profile and display information on the user interface. The input data specified on the user interface will be transformed into the SQL query that will be parsed into the health GIS system. The authors have considered the mapping between the input data and SQL query pattern. Moreover, considered how the query builder can read the ontology profile to generate information on the user interface.




(c.1) See graph type I.
*SELECT ?ClassName*
*WHERE ?ClassName <rdf:Type> <owl:Class>.*
(c.2) See graph type I.
*SELECT ?datatypePropertyName*
*WHERE*
*<ClassName> ?datatypePropertyName ?dontcare.*
*?datatypePropertyName <rdf:Type><rdf:DatatypeProperty>.*
*SELECT ?datatypePropertyValue*
*WHERE*
*<ClassName> <datatypePropertyName>*
*?datatypePropertyValue.*
(c.3) See graph type II.
*SELECT ?objectPropertyName*
*WHERE*
*<ClassName1> ?objectPropertyName ?dontcare.*
*?objectPropertyName <rdf:Type> <rdf:ObjectProperty>.*
(c.4) See graph type II.
*SELECT ?datatypePropertyValue*
*WHERE*
*<ClassName1> <objectPropertyName> <ClassName2>.*
*<ClassName2> <datatypePropertyName>*
*?datatypePropertyValue.*

Fig. 6. Designing user interface for ontological information retrieval

Fig. 6 depicts an example of the user interface structure (a) for browsing the ontological information that is derived from the graph model (b). The graph type I presents the model of information that queries the datatype property of a class and type II presents the model to query the value of object property of a class (ClassName1) that refers to another class. The semantic query builder reads the Semantic-DB profile (health information model) and queries using SPARQL patterns (i.e. (c.1) – (c.4)) for querying class, datatype property of a class and its value, object property of a class and the value of the object property respectively. Note that the name specified with symbol < > represents the name of ontological information, and those specified with symbol ? represent the variable in SPARQL. The datatype property value can be specified with filter condition in the semantic query builder. Fig. 7 depicts the user interface of the semantic query builder. Part 1 on the interface is the browsing tree to display ontological information (class, datatype and object property). Part 2 presents the selected datatype property and part 3 specifies the filter condition. The selected filter conditions appear in part 4. Spatial condition (e.g. coordination points) can be specified in part 5. Note that the system supports Thai users and hence, the interface is represented in Thai language.

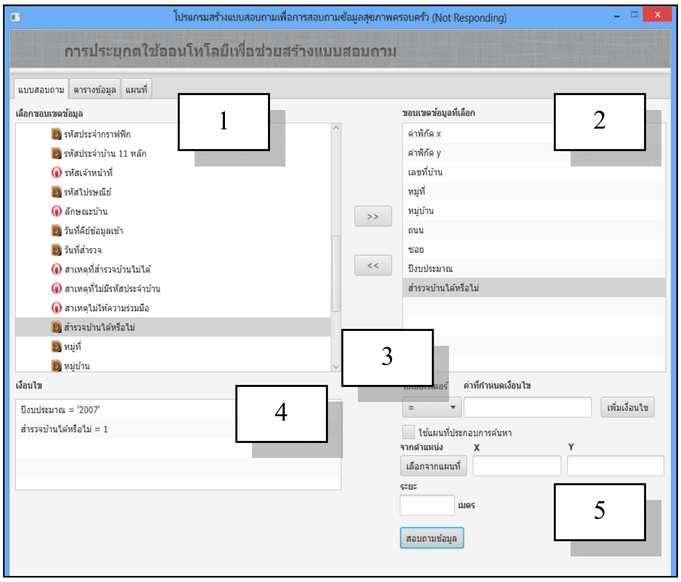



Fig. 7. UI of semantic query builder

### *B. Transformation the input data to SQL query*

The input data of the semantic query builder will be transformed to a SQL format. The authors selected SPARQL query model for transformation. Here, SPARQL query model is defined as shown in Fig. 8. The subgraph section (e.g. subgraph 1) presents the triple statement (subject, predicate, object) of the graph model that queries the datatype property value (e.g. Fig. 6 (c.2)). The SPARQL filter condition specifies the condition of the specified datatype property of the subgraph. The connected subgraph specifies the query on the object property that links two subgraphs (e.g. see Fig. 9).

These models are derived to the SQL query model as shown in Fig. 10. The semantic builder reads the input data and queries to consult type of the specified ontological information to the query engine (e.g. Jena) and generates the SQL query.

```
SELECT ?datatypeValue1 ?datatypeValue2
WHERE
{ /* Subgraph 1*/
 TripleQueryDatatypeValueOfAClass
Filter {?datatypeValue1   operator literal_value }
}
{ /* Subgraph 2*/
TripleQueryDatatypeValueOfAClass
FILTER {?datatypeValue2 operator literal_value }
}
… /* Subgraph N*/…
{ /* Connected subgraph */
  <ClassName1> <ObjectPropertyName><ClassName2>.
}
```

Fig. 8. SPARQL query model of the semantic query

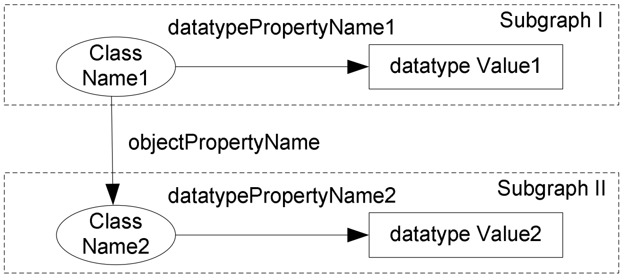



Fig. 9. SPARQL query model for connected subgraphs

```
SELECT  datatypePropertyName1,
           datatypePropertyName2
FROM ClassName1, ClassName2
WHERE  datatypePropertyName1= datatypeValue1 AND
           datatypePropertyName2= datatypeValue2 AND
           ClassName1. objectPropertyName =
           ClassName2.datatypePropertyName2
```

Fig. 10.Tranformed SQL query model

A compatible query question to the model in Fig. 9 is "finding the home with the village named 'Bangkapi' that the family has the monthly income more than '5,000' baht". Here, the subgraph in Fig. 9 can be the graph model type I and II (see Fig. 6).

### *C. Semantic Query*

The semantic query is specified into two types: attribute-based query and spatial query. The former queries information from the database tables by attribute-based matching. The latter includes geospatial data (coordination points) in the query.

Most GIS systems support three broad categories of spatial questions: direction questions, distance questions, and topology questions. In this paper, the useful questions in two categories: distance question and topology question are used. For example:

- What is the distance from the health center to the patients home?
- Which patients' homes are close to the health center within a range of 1 kilometer?
- Which patients' homes have a distance to the health care center over 1 kilometer?
- Which patients' homes are located within the specified area (e.g. area name, coordinate points) ?
- Within the specified area, which patients' homes are located within the range of 1 kilometer from the health center?

These questions are represented by Fig. 11 (a) to (f). The semantic query builder provides the functions that communicate with the ArcGIS functions i.e. QueryTask.

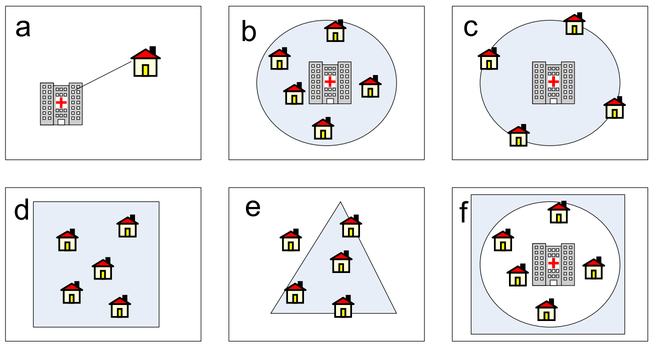



Fig. 11. Spatial model for spatial query

Fig. 12 presents the result of a map displayed on the semantic query builder.

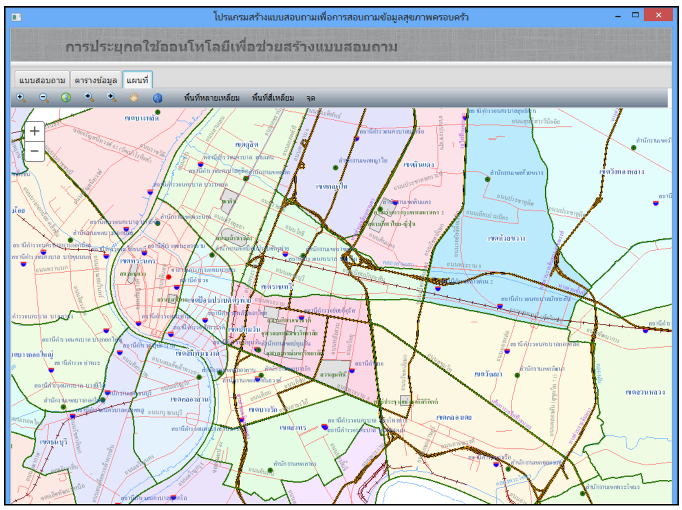


Fig. 12. Spatial model for spatial query

## VI. RELATED WORK

The OGC GeoSPARQL standard [7] is released since 2011 to support representing and querying geospatial data on the Semantic Web. GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data. In addition, GeoSPARQL is designed to accommodate systems based on qualitative spatial reasoning and systems based on quantitative spatial computations. However, it requires particular framework of the GIS system that supports the GeoSPARQL. Using the GeoSPARQL, the users are required to have the good knowledge of SPARQL which may be difficult for the users. Regarding ontology applying, the research work [8] proposed ontology as a share knowledge for querying the GIS database. They defined a collection of vocabularies describing the geographic entities. The geographic entities associate to geometry feature. They provided the mediator to link such vocabularies to the GIS databases. The provided ontology is stored in the database. The research work of [9] proposed the RDF-based ontology to express spatial and non-spatial objects according to the transportation domain. Some mapping rules that map the RDF components and database schema are presented. They consider SPARQL for the formulation of the RDF query to SQL query. The research work [10] proposes OWL 2 QL query to interacts with the GIS system.

This paper is close to [8], [9], and [10]. These work focused on the use of ontology for retrieving information from the multiple databases. However, some detail of the implementation such as the creation of the user interface and formulation of the query are not presented clearly in those works. This paper and the research of [9] both focus on the mapping of the SPARQL and SQL. However, their work considers retrieving of the features of the map while we focus on spatial query and attribute-based query. The developed system interacts to the ArcGIS system in which the ontology-based query is transformed into the database query parsed into the ArcGIS system. We consider the plug-in and the cooperation between client system and the map service of ArcGIS for displaying the result returned. The client system provides the user interface in which the users are able to interact with the health GIS system easily.

## VII. CONCLUSION

The prototype of the proposed semantic query builder is the proof of concept which is developed by JavaFX framework. It shows that this approach supports the users who may have less knowledge about GIS systems. Interacting with the health GIS system becomes easy and convenient to specifying queries. The prototype presents that the proposed components of the semantic query builder are sufficient for use. The authors interviewed end users who are public health staffs using questions in regards to usability of the provided tool and the result were very good (4.80 of 5). The interviewed staffs were satisfied with the system overall, particularly as they could specify their query whenever they need without cooperation with a computer expert. They can also plan their home visits quickly. In our approach, the ontology is created according to the database structure. The database table should be at least second normal form to reduce inconsistent data. Although, we proposed our system for the health domain however, the proof of concept shows that it is possible to adopt the proposed system for applying in any domains.

With the developed semantic query builder, the graphical user interface for specifying query is flexible. When the database structure is changed, the ontology profile that is transformed from the former can be used. However, if the ontology profile is very large, creating the input data for user interface may take time because the semantic query builder interacts with the query engine (i.e. Jena) to query the profile during specifying the query. To extend the capability of the system in regard to mobility, the GIS tuple data should be available over mobile devices or tablets in which the public health staff may access the data during home care visit. The mobile application may also include a geo-location service to provide the public health staff with the ability to know the distance from their current position to the target homes during their visits.

## REFERENCES

[1] J. A. Nhavoto and A. Grönlund, "Mobile Technologies and Geographic Information Systems to Improve Health Care Systems: A Literature Review", JMIR Mhealth Uhealth, 2014 April-June 2(2)

[2] Protégé. Available from http://protege.stanford.edu/

[3] OWL Web Ontology Language Overview, W3C Recommendation 10 February 2004, Available from http://www.w3.org/TR/owl-features/

[4] ICD-10. Available from https://dkm.fbk.eu/technologies/icd-10-ontology

[5] T. Tudorache, C. I. Nyulas, N. F. Noy, M. A. Musen, "Using Semantic Web in ICD-11:Three Years Down the Road", The Semantic Web – ISWC 2013, Lecture Notes in Computer Science Volume 8219, 2013, pp 195-211.

[6] A. Yajai, G. Sriharee, "EERtoOWL2: A Tool for Transforming RDB Data to OWL2 for Data Validation", Tools with Artificial Intelligence (ICTAI), 2012 IEEE 24th International Conference on Tools with Artificial Intelligence (ICTAI), 2012, November 7-9, 2012, pp. 970-975.

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

[8] F. Fonseca, M. Egenhofer, P. Agouris, and C. Camara, "Using Ontologies for Integrated Geographic Information Systems", Transactions in GIS, Volume 6 (3), 2002.

[9] T. Zhao, C. Zhang, M. Wei, Z. Peng, "Ontology-based Geospatial Data Query and Integration", In proceedings of the 5th International Conference Geographic Information Science, Lecture Notes in Computer Science Volume 5266, 2008, pp 370-392.

[10] Marek Sm´ıd and Zdenˇek Kouba, "OnGIS: Ontology Driven Geospatial Search and Integration", In Terra Cognita Workshop on Foundations, Technologies and Applications of the Geospatial Web. Tilburg: CEUR Workshop Proceedings, 2012, p. 27-38. ISSN 1613-0073.