A Linked Data Approach to Planning Collaboration amongst Local Governments in Thailand

Lerluck Boonlamp
College of Computing, Faculty of Technology and Environment
Prince of Songkla University, Campus Phuket,
Phuket, Thailand
Lerluck.b@phuket.psu.ac.th

***Abstract*—This research focuses on developing a new framework for the Sub-district Administrative Organization (SAO) of local government in Thailand. The goal is to publish linked data amongst local governments which allows discovering, accessing and encouraging interoperability. In order to achieve efficient interoperability of information systems, linked data play an important role in relating data. Linked data use semantic web technologies to publish structured data on the web and set links between data. Structure data are encoded as RDF. Furthermore, the data model supports integration of data sources by a linking semantic vocabulary, which is an effective way of exchanging data on the web. We propose a new framework for Thai local government with linked data technology to relate data that was not previously linked to enable exposing, sharing, and connecting pieces of data, information, and knowledge on the semantic web using URLs, RDF, data model and ontology.**

Keywords—Linked Data; Local Government; Collaboration planning; RDF Data Model; D2R Technology; SPARQL Language;

I. INTRODUCTION

A significant amount of data on the web is stored in relational databases (RDB). This is particularly true of the data of local governments in Thailand. Linked data technology can provide a method to relate data in RDBs amongst local governments and expose otherwise hidden linkages. Linked data from other sources may be connected as well. These are ways of exposing and sharing as resources on the Web and interlinking them with semantically related resource [1]. Providing links between data on the web of local governments would be very useful for SAOs (Sub-district Administrative Organization) to make a decision for their strategic and project plans. Moreover, with open publication of more information to the public, it attempts to increase trust from citizens and in support of government accountability and transparency initiatives [2].

In this paper, we proposed a new framework for local government. Our proposed system focuses on linked data technology to convert relational databases to RDF, mapping the RDF data model with semantic vocabulary and ontology. Moreover, we convert expressions in Thai language into SPARQL, an ontology query language.

Thailand policy reforms practice decentralization to break down the grip of central government. This significant reform drive [3] is motivated due to the performance of local government authorities, including the capabilities of these institutions to extract local revenues, deliver development services and respond to local preference. Local governments are territorially based organizations with administrative and fiscal autonomy. In addition, the citizens in the community are part of a decision making in the local government management and activities to expand their community development. Furthermore, local governments consists of several communities within a given region which share environment context such as natural resources, topography, geographical proximity and due to the existence of coordinating structures and financial management.

Indicating a gap, administration and development of local governments are based on strategic planning which derived from province policy. In general, strategic plans are designed as individual planning of each organization without collaboration. Lacking collaboration plans among local governments especially in the same province lead to difficulty in indicating if a province's strategic goals are achievable. In addition, it is hard to monitor and evaluate their developmental plans and supportive strategies. A further problem is duplication in planning and strategizing amongst local governments should be shared and cooperated.

With growth of the internet and web, technologies such as semantic web and linked data are increasingly used. Linked data publish structured data on the web and set links between data among different data sources. Following this, data are encoded using Resource Description Framework (RDF) as a triple in form of <*subject, predicate, object or value*> to set relations between things and enable new types of applications and web browsers that allow users to browse one data source and then navigate along links related data sources, examples

are presented as Fig. 1 and 2. The two basics of linked data employ RDF data model to publish data on the web and RDF links to interlink data from different data sources[1].

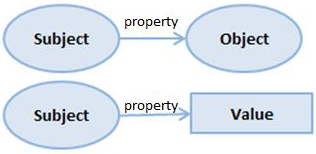

Fig. 1. A simple RDF graph

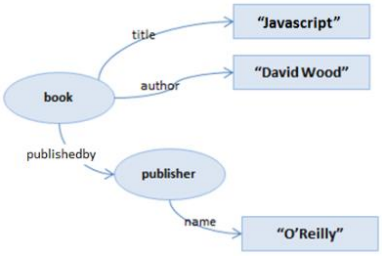

Fig. 2. The RDF graph describing the relationship between book and publisher

Linked data and related works in other domains such as Chinese medical and multimedia domains are described as following:

Interestingly, Chinese medicine (CM) [4] has used the linked data approach to publish CM knowledge in English language in order to integrate with Western medicine (WM). The results, the linked data approach made CM knowledge accessible and facilitated the association of CM and WM.

In multimedia domain [1], vast amount of multimedia are provided by different sources. Consequently, they use metadata as linked data on the web for creating and maintaining proprietary knowledge bases and increasing quality of organization, search, and retrieval of multimedia contents. In addition, increase their visibility and expand the coverage of their content to a Web scale.

In this paper, we present the linked data approach in Section 2. In Section 3 we introduce issues in collaboration amongst Thai local governments and propose a system to apply the linked data approach. Next, experimental results are presented in Section 4 and we conclude and future work in section 5.

II. Linked Data Approach

Linked data refers to "a set of best practices for publishing and connecting structured data on the Web" [6]. Linked data aims to share a common methodology for publishing data that avoids the heterogeneity of data sources. This mechanism is based on the four simple principles [7]:

1. Use URIs as names for things.
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
4. Include links to other URIs so that they can discover more things, example as presented in Fig. 3.

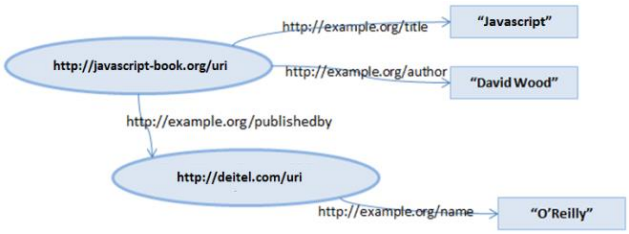

Fig. 3. Four principles of the linked data represents by RDF graph

These principles provide a framework for publishing data on the Web. The principles entail to share a common data format based on URIs and RDF, as well as use SPARQL as a common language to manipulate the data. In addition, the linked data [8] require to identify an entity via a single HTTP schema based URI. The identified entity is represented by URI. According to the third principle, this data be represented using RDF. RDF is a generic, graph based data model that represents information based on triples.

The forth linked data principle is to set RDF links into other data sources on the Web and enable applications to discover additional data sources.

There are three important types of RDF links[2]:

1) *Relationship Links* point at related things in other data sources. For example, relationship links enable people to point to background information about the place they lives, or to bibliographic data.
2) *Identify Links* point at URI aliases. It enables clients to retrieve further descriptions about an entity from other data sources.
3) *Vocabulary Links* point from data to the definitions of the vocabulary terms that are used to represent the data, as well as from these definitions to the definitions of related terms in other vocabularies. It enables linked data applications to understand and integrate data across vocabularies.

RDF allows us to formulate statement which consists of a subject, a predicate and an object. The subject and predicate in a statement must always be resources; the object can either be a resource or a literal node.

III. Application of Linked Data to Thai Local Government

We start by presenting the collaboration amongst local governments in Thailand in the first subsection. In the next subsection we formulate our linked data approach to collaboration.

A. Collaboration in Thai Local Government

Normally, strategies and projects are designed individually without cooperation among organizations within the region. Moreover, the application that they are using do not provides a channel for interchange, distribution and searching data

[1] http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/

[2] http://linkeddatabook.com/editions/1.0/

between organizations. These lead to ineffective management, duplication and might not achieve the province goal. The linked data technology allows the linking and browsing between data in different sources.

One medium technology used by Thai Local Government (TLG) is e-Plan an online Web application where existing data are stored in a relational database at the Thai Local Administration Department. However, the information which is recorded on computerized base cannot be distributed and reused. Ideally, the data should be distributed, published and used for further development. From Tim Berners-Lee [5], the basic assumption behind linked data technology is the value and usefulness of data increases the more it is interlinked with other data.

B. Linked Data Approach to Collaboration

This paper proposes a linked data approach to overcome the limitation of discovery in the TLG system. This section first motivates its development and describes the new framework of a TLG system. In this place, the local government ontology is proposed.

The proposed system, which we call new framework of Thai local government, is illustrated in Fig. 4. The TLG data are stored in RDB. Several technologies can publish RDB to RDF such as Tripify, Vituroso and D2R but we selected D2R because D2R is faster and can stores vast amounts of data [9]. Therefore, we publish the RDB to RDF by using D2R technology. D2R Server[3] is a tool for publishing the content of relational database on the Semantic Web. Database content is then mapped to RDF by a mapping which specifies how resources are identified and which properties are used to describe resource. Based on this mapping, a D2R Server allows the RDF data to be browsed and searched with the SPARQL language.

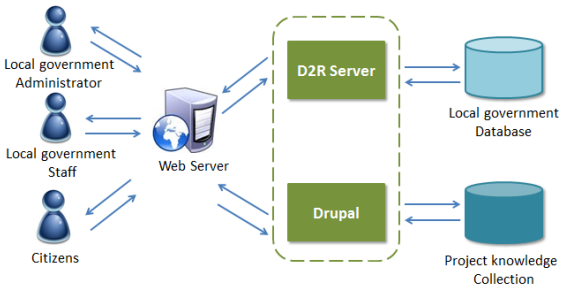

Fig. 4. A new framework of Thai Local Government System

Another part of the framework, the SAOs will use Drupal application as open participation channel. Drupal is a content management system, applied for network communication and project knowledge collection. This framework can categorize three groups of users who had a different authorization on the system.

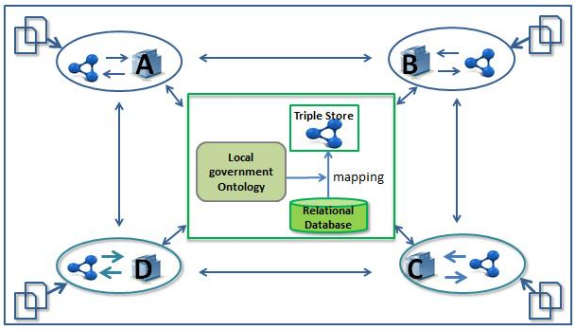

Fig. 5. System Architecture

As Fig. 5 is represents system architecture of the TLG. This system architecture uses linked data technologies to links data in different sources and systems. Two main data storages embedded in a framework are central and individual spaces can be described as follows:

- Part I: the central part, all of local governments' data are stored in the RDB. Therefore, we will convert RDB to RDF triples store. Mapping data in RDB and publish it to the RDF as triples according to the concept of ontology. The local government ontology vocabulary will be extended from standard ontologies. A further the local government ontology states relationship between classes and subclasses as well as properties as expresses in a data model. The data model is entities relationships diagram to describe structure of data of domain interested. The following step is to apply existing standard ontology to data model as presented in Fig. 6.
- Part II: at organization, the system will be developed on Drupal and data are stored as RDF. The metadata will be defined to describe entities following the same standard property as the local government ontology.

From the system architecture introduced above, two resources linked from different data sets on the system, allowing data in one data source to be linked to other data sources.

Fig. 6. Overview of Thai Local Government Ontology and Data Model

Thai local government ontology contains three kinds of classes. Those three concepts are project, organization and strategy. As a result, it is able to describe that project has

[3] http://lod2.eu/Project/D2R-Server.html

associated with which strategic and is developed by which organization. The RDF data model represents information as node-and-arc-labeled directed graphs. Therefore, defining standard ontology to data model enables discovering other data that relate projects and strategy as demonstrated in the section 4.

IV. Experiments

To evaluate the part of the proposed system which publishes RDF to RDF, D2R server technology is used. Database schema is the most essential part which needs to be clarified before we start generating a mapping file. After that, ontology and data model are designed in order to state semantic vocabulary.

D2R Server uses the D2RQ mapping language [10] to capture mapping between application specific databases schemas. A D2RQ mapping specifies how resources are identified and how property values are generated from database content. An example mapping language is presented in Fig. 7.

```
# Table t_plan_project
map:t_plan_project a d2rq:ClassMap;
      d2rq:dataStorage map:database;
      d2rq:uriPattern "project/@@t_plan_project.YY@@/@@t_plan_
      d2rq:class local:Project;
      .

map:t_plan_project_PROJECT_NAME a d2rq:PropertyBridge;
      d2rq:belongsToClassMap map:t_plan_project;
      d2rq:property local:projectName;
      d2rq:property rdfs:label;
      d2rq:column "t_plan_project.PROJECT_NAME";
      d2rq:datatype xsd:string;
      .
```

Fig. 7. Example of D2RQ mapping language

D2R Server enables RDF and HTML browsers to navigate the content [11]. From the example presented in Fig. 8, a road construction project is represented as a subject where other properties and object values links display in the page interrelating additional nodes such as organization, vision, province node and other projects in the organization.

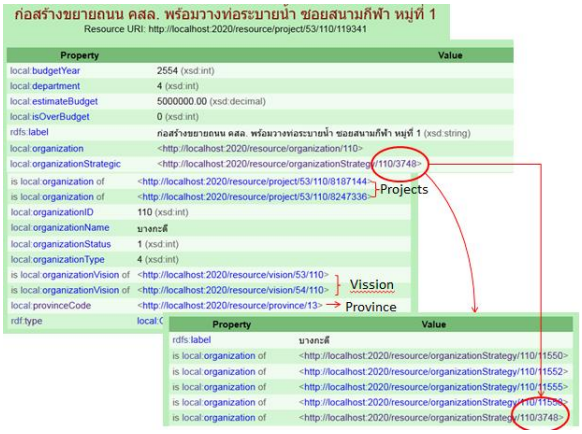

Fig. 8. Browse project from D2R

Search is by using SPARQL endpoint to query RDF data. As in Fig. 8, inquiry all projects in year 2553 (2010) from Organization number 110. The query outcomes display organization, year and project name. If we follow the organization links it will connect to organization information pages. At the project name column, it should be a node and not a literal value which needs programming.

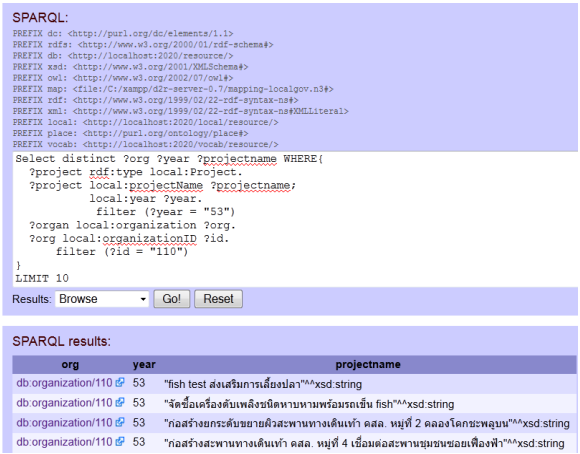

Fig. 9. SPARQL query and results

V. Conclusion and Future Work

In this research, a new framework of Thai local government system is presented to facilitate collaboration amongst, Thai local governments. The proposed method uses D2RQ mapping language into ontology and the linked data model. The derived strategies and projects pattern are relevant for SPARQL queries. Following the result links will yield useful information relating topics, which were not available before. These can help organizations in collaborating in strategizing and project planning.

For future works, existing standard ontology will be applied with the data model in the mapping file. In addition, organization communication channel will be developed by using Drupal application. Therefore, with the proposed ontology, people in community can discuss and exchange their experiences on project planning of organizations. Finally, metadata in Drupal system is defined with the intention of links information in different data sources and systems between Drupal and D2R.

References

[1] B. Schandl, B. Hasthofer, T. Burger, A. Langegger, and W. Halb, "Linked Data and multimedia: the state of affair". Journal of Multimedia Tools and Applications, volume 59, number 2 (2012), 523-556.

[2] B. Hyland, and D. Wood, "The Joy of Data – A Cookbook for Publishing Linked Government Data on the Web.

[3] M. Bratton, "Citizen Perceptions of Local Government Responsiveness in Sub-Saharan Africa". *In World Development, Elseview Ltd.*, 2011.

[4] J. Zhao, "Publing Chinese medicine knowledge as Linked Data on the Web". Chinese Medicine 2010, 5:27. http://www.cmjournal.org/content/5/1/27

[5] T. Berners-Lee, Fielding (Day Software), and Masinter (Adobe System). "Uniform Resource Identifier (URI): Generic Systax". *Network Working Group*, January 2005.

[6] C. Bizer, T. Heatlh, and T. Berners-Lee, “The Story So Far” *in International Journal on Semantic Web and Information System*s 5(3), 1-22, 2009.

[7] C. Bizer, T. Heath, and T. Berners-Lee, “Linked Data – The Story So Far”. International Journal on Semantic Web and Information Systems 5(3), 1-22, 2009.

[8] O. Hartig, and A. Langeger, “A Database Perspective on Consuming Linked Data on the Web”. Datenbank Spektrum (2010) 10: 57-66.

[9] Hert, M., Reif, G. and Gall, H.C. “A Comparison of RDB-to-RDF Mapping Language”. Proceedings of the 7th International Conference on Semantic Systems, page 25-32, ACM, 2011.

[10] C. Bizer, and A. Seaborne, “D2RQ: Treating non-rdf database as virtual rdf graphs” *in 3rd International Semantic Web Conference (ISWC2004)*, 2004.

[11] C. Bizer, R. Cyganiak, and T. Heath, “How to Publish Linked Data on the Web”, 2007. http://www4.wiwiss.fu-berlin.de/bizer/pub/linkeddatatutorial/