Towards Transparent Governance by Unifying Open Data

Paramjeet Kaur, and Parma Nand

Abstract—**Open data initiatives have been gaining increasing momentum in recent times both with national government as well as regional governing bodies. Increasing number of governments are realising the need for the role of transparency in governance. In congruence to this, a number of governments have been working towards opening up government operational data to the public, however the data is extremely large, disparate and segmented, hence is hardly useable by general public as well as data consumers. A number of countries have worked on the open data initiatives and attempted to link the data using diverse frameworks and publishing tools. This paper presents an analysis of the works that have developed ontologies and frameworks to link the data released by governments as part of open data initiatives. It also provides a critical evaluation of the techniques used and suggestions of novel techniques that can be used to improve the frameworks and ontologies. We also present the results of a case study that used an ontology-based linking of data released by governments as part of the open data initiative. The results show that and an ontology mapping of such raw data drastically enhances the usability and the quality of the raw data.**

Index Terms—**Open data Initivatives, Semantic Web, Ontology, Open Data, RDF triples, Semantic links.**

I. Introduction

World Wide Web has been able to connect the world by the use of hyper links between the web documents. These hyperlinks are used to navigate between html pages containing information as free text. In this way, all of the documents can be accessed via a single web link which integrates other links embedded in pages. An accelerated growth of informed society means people are growing awareness of their rights and they want to know, and get increasingly involved, with governance mechanisms.

This has given rise to the necessity of having the transparency by governing bodies resulting in releasing data to the public based on the rationale of making policy decisions transparent. Open data model proposed by [1] highlights the benefits and use of open data in research which can help researchers in making data driven decision. The fundamental principle of Open Government is that the general public has the privilege to access data, information, and activities generated by government agencies. The raw data, without any interpretation, plays a vital function in this situation [2]. By releasing data related to governance by the ministries, the government gains the trust of their citizens. This gives a clearer picture to the public on the expenditure and policy decisions by increasing trust between the public and the government.

Manuscript received April 06, 2021; revised Oct 12, 2021.

P. Kaur is a PhD candidate of Auckland University of Technology Auckland, 1010, New Zealand, e-mail:param.kaur@aut.ac.nz

P. Nand is a Senior Lecturer in Auckland University of Technology, Auckland, 1010, New Zealand, e-mail: parma.nand@aut.ac.nz

In 2009, the president of USA Barrack Obama announced the concept of open data [3]. According to him, it is the right of the people of the country to know, what is happening in the government and how the government is investing their money. A lot of the countries were not in favour of the idea at the time and Obama got a lot of objections for this announcement. However, following this, in 2012, the UK government launched its own open data initiatives. After that, a lot more of the developing and developed countries have become proactive, and they have embarked on their own open data initiatives.

These (ODI) open data initiatives release raw data via various data portals which are disparate and difficult to use, especially by the general public. The data files which have been linked on these data portal does not provide the semantic links among the data sources, hence is difficult to draw the nexuses between data entities. This has created a need for techniques and ontologies that can relate the data with useful semantic links and make it easily accessible over the Internet so that the general public can easily access and use it.

II. Open Government Initiatives

Open Government Initiatives are becoming increasingly critical among the developed as well as the developing countries. The aim of Open government initiatives is to provide transparency and wide data reuse. The availability of open data has numerous benefits for the public and other stakeholders. However, the open government initiative data is disparate and published in heterogeneous formats which makes it challenging to link and reuse the data.

In recent years, a number of (OGD) open government data movements sprung up around the world with two major aims, transparency and data reuse. Some examples are, Barrack Obama's open data initiatives in 2009 [3], open Government Partnership in 2011 [4] and Open Data Charter in 2013 [5]. These movements have given rise to open data portals such as data.gov, data.gov.uk, data.gov.gr, open.data.al and data.gov.nz to enable stakeholders and residents to obtain information on any particular ministry of the government. The (OGP) Open Government Portal was announced in 2011 at which point only 45 countries choose to become members. However, since then more countries have joined the OGP and the number has gone up to a total of 94. Moreover, statistics show that New Zealand is at the 8th place in the OGP index with 68% open data sets [4].

These facts demonstrate the fast-growing rate of data transparency and availability in New Zealand. Figure 1 shows the 2014 Global Open data Index of number of places, where Grey colour cells shows available open data sets whereas, the black colour cells show the data sets not yet open.

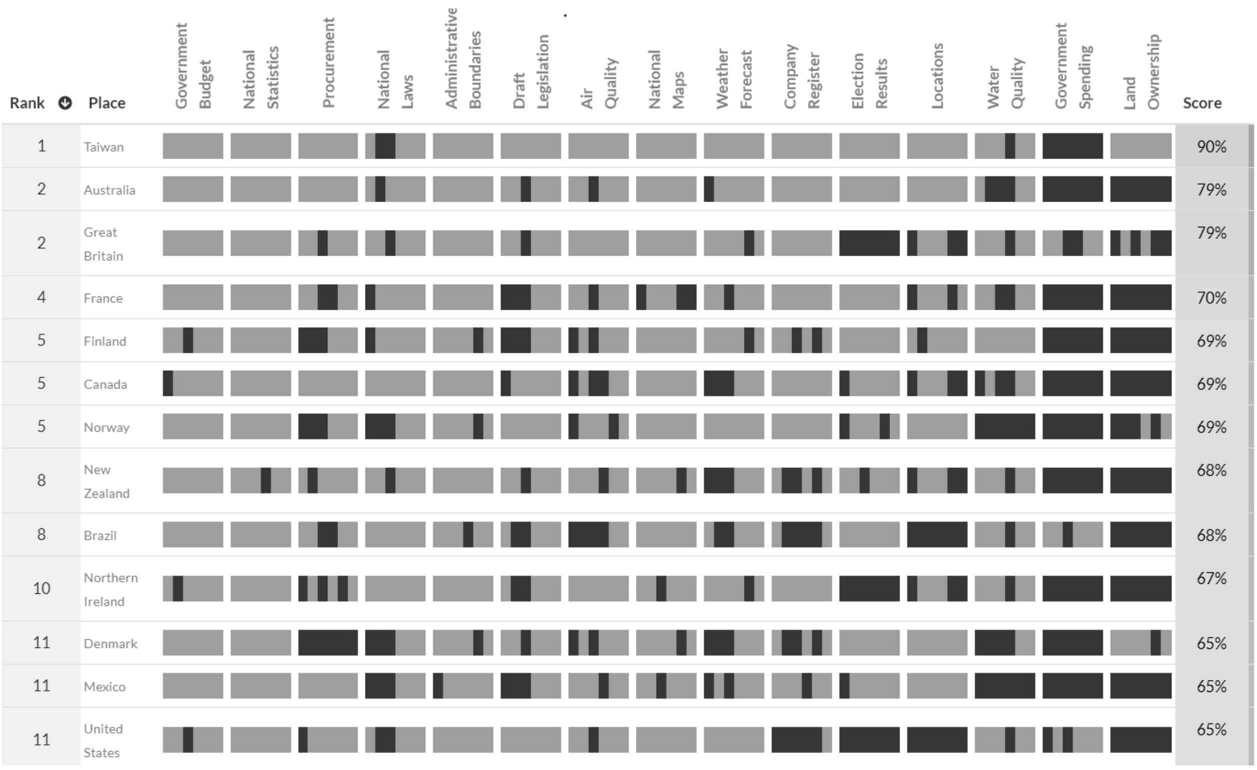


Fig. 1: Global Open Index (source: https://index.okfn.org/place/) ”public domain”.

The main aim of this index is to track whether the data published is accessible to all stakeholders and measuring the openness of the data on global level. The availability of open data has several valuable benefits for the general public as well as professionals who use data for making decisions and planning purposes. Encoding (OGD) Open Government Data as (LOD) linked open data, would enable a user to browse a data source and then to navigate via the links into other related data sources to get all relevant data at one place [6].

Furthermore, linked open data would also help other decision makers, such as business managers, in essential long-term planning activities because they would get the relevant information easily in one place. In summary, (LOGD) Linked Open Government Data will increase the government transparency and public awareness of government processes.

A. Linked Open Data

Linked Data is a set of guidelines to link related structured data on the web. The data is represented in triple form (subject, predicate, and object). The subject and predicate in a triple are always (URIs) Uniform Resource Identifiers and object is either literal such as string or number or another URI. Furthermore, linked data refers to create a link between the data that is from heterogeneous sources and in different formats such as (HTML) Hypertext markup language, (CSV) Comma separated value, (XML) Extensible markup langauge, (XSL) Extensible stylesheet language. It is sustained by more than one organization in diverse locations, or it specifies heterogeneous systems within a single firm.

In this way, linked data refers to the data distributed on the web so that it is machine readable, its meaning is defined, and it is linked to external data sets as well as within datasets [7]. Linked data is an extension of the web with worldwide data space connecting different companies, films, music, scientific experiments, television, radio programs, medical, drugs, clinical and many more. There are linked data engines that allow a user to crawl the web data, following the link between the data sources and provide expressive query capabilities. Linked data is based on two fundamental technologies:

- URI (Uniform Resource Identifiers): are the entities that exist in the real world. It also specifies the address for document and entities on the web. The entities of URIs use the http:// scheme, these URI entities can be looked by dereferencing the URIs over the (HTTP) hypertext transfer protocols.
- HTTP (Hypertext Transfer Protocol):provides the universal mechanism to retrieve the information or resource that is sequential as a stream of bytes (for instance: image of flower) or description of the entity that cannot be sent across the network (for instance: the flower itself)

1) Method of Linking: There are several possible formats that is used to publish OGD such as CSV, spreadsheets, HTML tables, and (PDF) Portable Document Format files. An important issue emerging from publishing data using multiple formats is that the users face problems in linking multiple datasets and initiating a data analysis process.

For example, assume that a user needs to find the number of schools located in a particular area and the road traffic in that area in the last 2 year so that he can find out which school is located nearby and the rate of traffic during normal and peak hours on those roads which go to that school. To achieve this the user will need to open two data sets to complete the analysis, in addition, this becomes more complex when performing analyses with more than two data sheets.

Therefore, there is a clear need for infrastructure that links the raw data while preserving the classification, where classification refers to a conceptualization of data according to domains [8]. One way of data linking, and analysis is Relational database and other is Graph database.

2) Relational Database: It is a traditional way of data storage in which tables with multiple rows and columns are created, known as a record. Each record consists of a set of fields to hold relevant information on a class of objects [9]. For instance, we could have three tables agriculture, land, and rainfall. These tables could be linked by using fields from one table to another. This would then enable one to extract the information from multiple tables. However, imposing complex queries on these tables is very tedious and requires lots of effort and if a table has changed, whole schema needs to be changed.

TABLE I: Agriculture

Year	Region	Types of Farms
2001	Auckland	Pig
2002	Wellington	Vegetable
2003	Christchurch	Beef

TABLE II: Land

Year	Region	Types of Land
2001	Auckland	Forest
2002	Wellington	Farming
2003	Christchurch	Volcanic

TABLE III: Rainfall

Year	Region	Rate of Rainfall
2001	Auckland	2345mm
2002	Wellington	1200mm
2003	Christchurch	2345mm

In the above tables I, II, and III we have information related to Agriculture, Land and Rainfall. These tables can be joined to access the information via a single query. However, if new rows are added to table I and III,the query wouldn't be able to retrieve accurate results. Every alteration brings change in schema. It requires manual efforts to alter the table and perform the joining again.

Similarly, if any deletion and alternation happens, the whole schema needs to change. It is easy if the tables are small but, very complex, challenging and time consuming if table size increases. Therefore, relational database has limitation of being static which can be overcomed by a dynamic database such as Graph databases.

3) Graph Database: A graph database is an operational database which is designed to operate (CRUD) create, read, update and delete processes on a graph data model. It uses graph structure for semantic queries with nodes, edges and properties mainly to store and represent the data [10]. It stores the relationship between records. So, a graph consists of nodes and relationships. Each node represents an entity such as, place, a person, thing, category or other data. Relationship represents how two or more nodes are connected or associated with each other.

Graph database uses well defined data models. The most common graph data models are property graph, hypergraphs and (RDF) resource description framework triples. Figure 2 is an example of property graph. It illustrates the relationship among agriculture, land and rainfall in a graph database. The nodes are labelled as agriculture, land and rainfall. It is connected with relationship describing how each node is connected. As we see below, vegetable depends on land and rainfall for farming, where farming depends on the rate of rainfall. Moreover, hypergraphs are isomorphic in nature hence can be represented as a property graph but not vice versa. Graphs are the means of drawing information via diagrams.

This method of depicting graph is precisely easy to read for humans if the graphs are small. But, with thousands or millions of nodes, it will be difficult. Moreover, to store and process these graphs in computer systems is also a complex task. Therefore, to extract the information from large graph, we need to divide it into smaller parts where each part can be stored independently. To transform such complex data structures into linear strings is known as serialization [10].

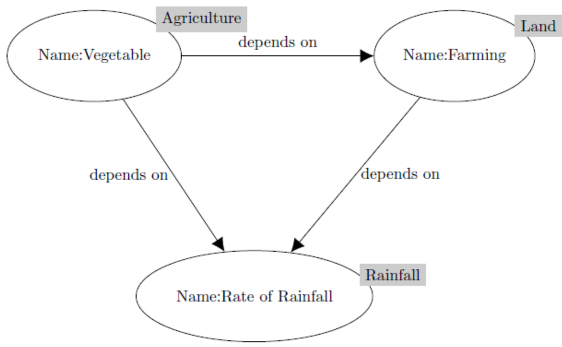


Fig. 2: An Example of the Property graph

4) RDF Triples: RDF, is a type of graph database which stores the semantic facts and information. RDF is a model of data publishing on web standardized by (W3C) World Wide Web Consortium which supports semantic queries. The main aim of RDF is to allow applications to transfer data on the web without compromising their original meaning. Furthermore, data in RDF is stored in triples which consists three elements- subject, predicate and object. RDF format is able to take any subject or concept and relate it to any other object using the predicate (verb) which shows the type of relationship between the subject and the object [11].

RDF document is a directed graph where both nodes and edges are tagged with identifiers, are also known as URIs. Generally, URI's are used for the subject and predicate, where object can either be another URI or a literal such as, string or number. Further, a literal can have a type which can be a URI. This specifies that triples can have up-to 5 bits of data. Figure 3 is an RDF graph with three triples, that specify agriculture depends on land. Further, vegetable and farming are literals of agriculture and land URI's respectively. Agriculture depends on land which specifies that vegetable is a type of agriculture that depends on farming land.

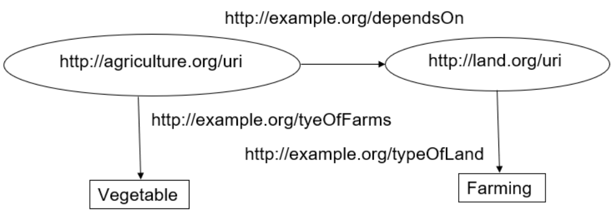


Fig. 3: An RDF graph describing data values with literals

In addition, RDF triples are written in turtle as follows:
<http://agriculture.org/uri>
<http://example.org/dependsOn>
<http://land.org/uri>.
<http://agriculture.org/uri>
<http://example.org/typeOfFarms>"Vegetable".
<http://land.org/uri>
<http://example.org/typeOfLand>"Farming".

Here URIs are written in angular brackets, quotation marks are used for literals and a full stop is used to terminate all statements.

5) Semantic Links: RDF triplestore is used to publish and manage the linked open datasets such as GeoNames and DBPedia. RDF encoding provides a fast query retrieval by using SPARQL[12], which is a semantic query language similar but not same as (SQL) Structured Query Language. It can be used to retrieve and manipulate the data stored in RDF which is the basic representation format for the Semantic Web.

An Ontology supports the organization of Linked Data based on the conceptualization. Semantic links bring the concept of linked open data, in which URIs have been created between data files so that different department's data can be linked together based on same fields available in the data files. This will solve the problem of missing links and data will be machine readable.

More than 70% of the data on the web today are unstructured text [13] [14]. This applies to government data as well where, large quantity of data is hindered in natural language text documents making it difficult for humans to understand quickly.

6) Ontology: An ontology can be coded using both OWL (Web Ontology Language) and RDFS (RDF Schema). However, OWL is more expressive than RDFS hence is more suitable for small scale ontologies [15] [16] [17]. Ontologies are used to process, capture, reuse and communication of knowledge. It can be defined as a "specification of conceptualization". The domain structure is captured by an ontology, conceptualization represents knowledge about domain.

An ontology is the study of entities in real life and the relationship which these entities carry with each other. It is used in different fields such as life science, artificial Intelligence, libraries and recently in Computer science [18].

The usage and purpose of ontologies varies from application to application [19]. The entities in an ontology encapsulates the concepts while the taxonomy represents the relationship between them. In recent times, domain ontologies have become the integral part of numerous knowledge-base and semantic applications [20]. However, the design process of such ontologies and publishing them with correct individuals is a tedious and time consuming task which requires open data as an input so that triples can be generated to make it linked open data. Linked open data and semantic web technologies provide the mechanism for sharing the knowledge which comes from diverse sources [7].

According to Tim Berners Li's principle of five star linked open data. The combination of RDF, URI (Uniform Resource Identifier) along with SPARQL query make 5 star linked open data. Figure 4 depicts, Berner's 5 star Linked Open Data rule. In nutshell, graph databases have significant better performance in structured type queries and full text character searches as compare to relational databases [21].

III. Related Work

A. Linked Open datasets using String matching

The first systematic study of linked open dataset was reported by Hassanzadeh et al[22] which connects several existing movie web resources. String matching measures are used to discover similar links among movie datasets. For evaluation, all string matching results are compared using the owl relationship type, "sameAs" links. Therefore, the exact matching of movies is not so possible and can give small amount of similarity, hence false results.

Additionally, by providing related links about the entities along with the Meta data can enhance the results. The accuracy of the results are obtained manually by finding all the matching. Thousands of links are inspected manually, which is a time consuming process. However, if an automatic mining techniques for the data sources is provided, it will save the time. Moreover, more internal links can be provided among the related entities such as movies with same title. The links can be found using the similar string matching techniques. Nevertheless, the linked movie database can be extended further to make it easy to use so that the users can provide feedback on the quality of the links.

B. LOD Algorithms for Ontology Alignment

A number of researchers have reported Linked open data algorithm for alignment of ontologies of data sources such as, zoology, geospatial and genetics [23] and data sources GeoNames, LinkedGeoData, DBPedia, GeoSpecies, and GeneID. In order to determine ontology alignment an approach has been discussed in which, extensions of the classes are compared and restriction classes are defined over the ontology. When the source of the ontology is elementary, restriction classes help to get the redefined set of classes. In an ontology, restriction classes are used to identify existing and derived set of classes. Moreover, five pairs of data sources are used to evaluate the alignment of the algorithm by using experimental evaluation.

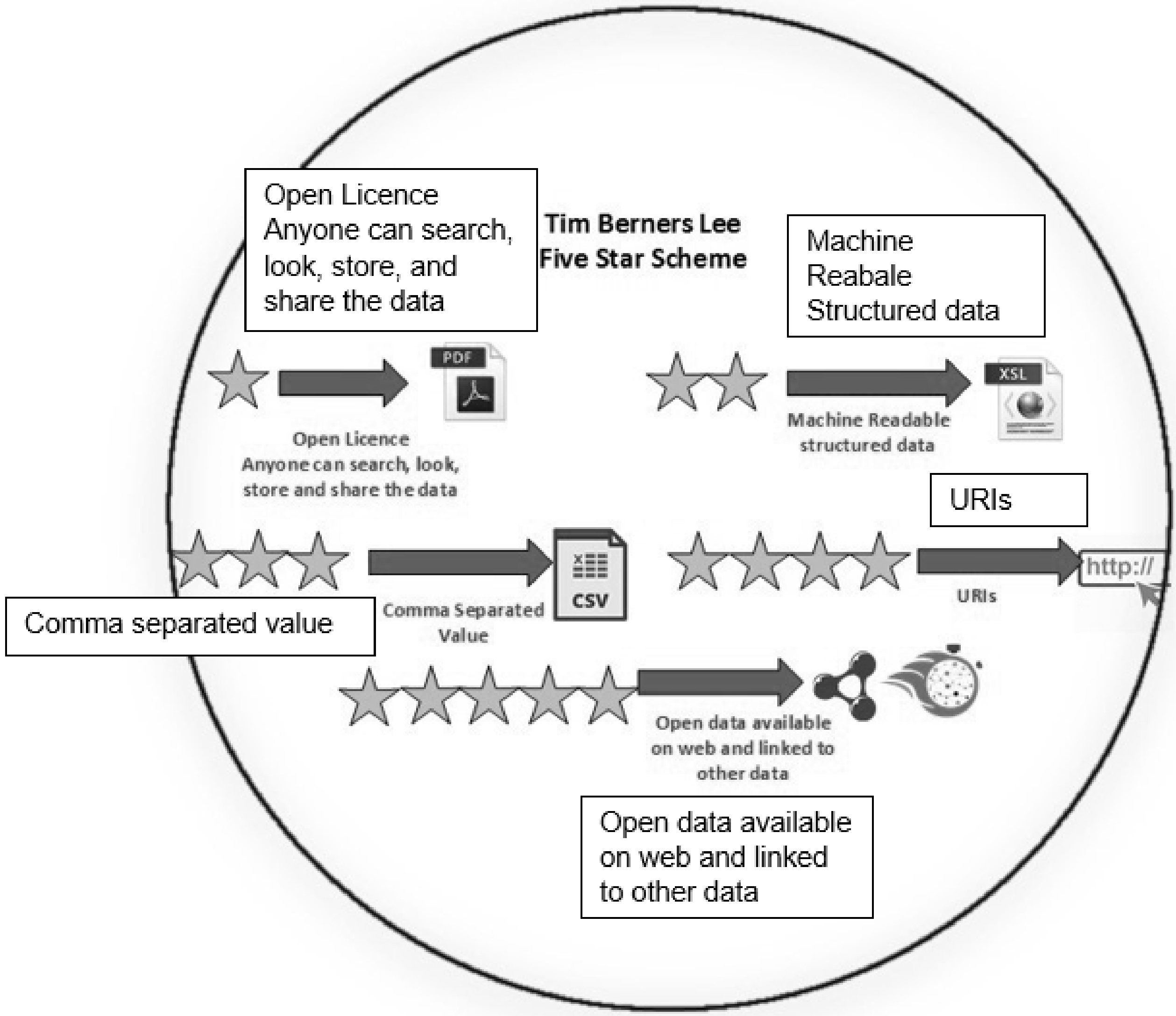


Fig. 4: 5 Star Linked Open Data (figure redrawn from (https://5stardata.info/en/) (CC BY 1.0 Universal).

In this evaluation, linked instances are used but properties not generating useful restricted classes are removed. This provides the reduced data sets, which highlight the applied usage of the equivalence links and properties pertinent to the domain. However, by creating alignment theories the scalability of this approach can be improved. Therefore, experimental investigation of the space will enhance the performance of the algorithm. This algorithm can be applied to align data resources in biomedical contexts.

C. LOD Algorithms to detect hidden links in datasets

Le et al.[24] has designed an algorithm which is used to detect the hidden links in US Census, DBPedia, GeoNames and World factbook. In this algorithm, linked data has been converted into graphs where nodes represent the URIs and edges represent the links between URIs. The designed algorithm has been compared with naive Bayes algorithm which uses only URIs names to make predictions.

The proposed algorithm resolves the problem of ambiguity in terms of different geographic locations with same name. In order to get accurate results, multiple patterns of graphs were generated. Missing entities were creating a problem so, a separate pattern was created and the algorithm was applied to predict new entities that can be linked to the missing entities.

However, lot of uncertainty still exists about the usage of naive Bayes algorithm for comparison with the proposed one. Also, the challenge is to detect hidden links between well matched URIs and noise.

D. LOD Prototypes implementation using OGD data sources

Other studies have proposed architectures for indirect determination of links between entities to build Linked data. The main aim of this was to link different OGD sources by creating owl: SameAs links among the URIs. Further, a prototype scenario is designed in which three actors are considered- a school, local directorate of secondary education of Athens and the ministry of Education. Five implementation steps are performed to access the consumption of data related to a specific school.

SPARQL endpoints are used to access the specific datasets [25]. However, the whole process is manual which is time consuming. A Silk framework which is an open source framework for integrating heterogeneous data sources, has been discussed, but there is no details on how it has been enforced. Further investigation is required to understand the relation among the data models and political priorities.

Liu et al.[26] projected a prototype implementation based case study for linking the Australian government data for sustainability science research. Data sets of energy consumption, population, economic, rainfall and temperature are used to examine the practical value of linking the data of Australian government.

The main focus is, reusing published data sets. Numerous challenges were encountered related to data description, discoverability and analysis. However, the process of discovering relevant datasets is manual. The automation of this process will enhance the analysis functions of linked data. Moreover, SPARQL endpoints are discussed partially but, not implemented for data access.

The discussed studies so far have only focused on implementing algorithms and prototypes using linked open data. The work reported by Haslhofer et.al [27] goes a step further and implements the first version of linked open data using open datasets of libraries, archives and museums sectors. The Technical architecture of the linked open data pilot comprises (ESE) European Semantic Elements metadata. Semantic annotations are created using Apache Solr tool and all Meta data files are accessible via Europeana data portal.

At second layer of the architecture, ESE has been converted to (EDM) Europeana data Model where, dumps are created from semantic enrichment and ESE to EDM conversion. Further, the dumps are stored as RDF which can be accessed by the data portal of Europeana. Whenever, the linked data clients access data in RDF specific internet type from data.europeana.eu. The request will be accepted otherwise, HTTP requests are redirected to European portal. Using this architecture, one can access the Meta data about Europeana resources. However, this alignment requires further evaluation so that, it can be extended to other data sources such as DBPedia and other relevant initiatives.

In an analysis of open data for e-government Theocharis et al [28] discussed the Greek open data initiatives opportunities and benefits. The authors have debated the challenges of opening the data for the public including the availability, accessibility, reuse, simplicity, global participation and redistribution. Moreover, the basic steps for converting data to open data have also highlighted by proposing an architecture for linked open government data.

E. Internationalization of Linked Data using Framework implementation

The study of linked open data is further carried out by Dimitris et al [29]. They examined the internationalization of linked data where language specific DBPedia framework has implementions to publish linked data in non-Latin languages.

However, it is challenging to create manual SPARQL query using URIs in non-Latin language. The aim is to make it easier for the incredible amount of information in DBpedia to be used in new and interesting way and to inspire new mechanisms for linking and navigating. The proposed work provides new ways to improve the Wikipedia because DBpedia serves as a vital programmatic interface for Wikipedia. Nevertheless, it is challenging to create manual SPARQL query using URIs in non-Latin language. It gives rise to the problem of consistency in terms of triple serialization formats.

Another example of work using linked open data was conducted by Charalampos et al [30]. They demonstrated a methodology to analyze the open data movement of Public data landscape from three different perspectives: semantic, functional and technological. This gives statistical results of the availability of open data in each perspective. However, the government bodies consider their websites to be advanced than technical architecture that open up their data to others to use. If the government focus on providing reliable, simple and easy to use technical architecture to expose the data, it will encourage the private sectors to share useful information or knowledge to stakeholders.

Other studies by Juan et.al [31] have proposed a theoretical framework to analyse open data based on supply and demand mechanism. Data sets from Investment portfolio of Mexican federal budget has used to evolve an example of data visualization. Tableau software has used to visualize the resulted graphs. These visualization results will enhance the decision making progress. While, accessing the graph, user can apply numerous filters related to the details of project such as nature, resources allocated and branch. The graphs are user friendly and helping the non-technical users to access the information. However, this case is related to Mexican only and how the proposed framework can benefit other stakeholders and countries need to be explained.

Additionally, Judie et al [32] discussed the opportunities, challenges, issues, hindrances of the open government initiatives. Moreover, guidelines have been discussed to publish the open data. This systematic study also, identify the impacts on the stakeholders who are involved in the usage of open data initiatives.

In the frequent study Fragkou et al [33] implemented the methodology described in [34]. The main goal of this implementation is to create link between open data initiatives by using an interface which is based on the E-GIF ontology and Jena framework. Jena is an open source semantic web framework for Java. It provides API for data extraction from RDF. Further, SPARQL endpoints are implemented to query the datasets. However, the main focus is on the working and characteristics of Jena framework. It seems the proposed methodology is based on the effectiveness of the tool. Further, the pre designed E-GIF ontology has used for creating RDF models. There is no clue how this E-GIF ontology has created.

Another survey such as that conducted by Li et al [35] have shown that value realization of OGD in information policy, technology and economy. The impact factors of value realization of Open government data include information policy, technology, investment, infrastructure and information system management. However, OGD area is still lacking in theoretical groundwork and experimental study. Moreover, the study of linked open data was further carried out by Azevedo [36] in Geographical information system. This is the project of the Brazilian federal government, two case studies are used to get a cost effective decision making process to minimize the damages from the flood.

In the first case study, data from diverse sources are collected and converted into CSV form using MS Excel Software. Further, RDF datasets are created from CSV files using D2RQ platform. In addition, SPARQL queries are implemented and visualized by using (GIS) Geographic Information System web application. In the second case study, the designed prototype is used to identify the competency by collecting qualitative data from people having informal group discussions. For this purpose, DB4Trading web application was built so that, the users can validate data for semantic repository using their own criteria. Although, this framework seems user friendly but, how the damages of flood will be minimized using this framework is not so clear. Moreover, the datasets can be extended to improve the relevancy. The Data visualization application DB4Trading needs to be discussed in detail.

In another major study [34] authors demonstrated the applications of linked open data technologies on the data available in ERMIS Government portal for Public Administration. Jena framework is used for the proposed methodology. The HTML web pages has given as input in the Jena framework from which instances of RDF triples have created and passed to the open link virtuoso server. Further, SPARQL endpoints are provided to access the data.

In addition, the Greek and English version of the page has passed to E-GIF Ontology and then to Jena framework. Furthermore, D2R server with (TDS) triple database store has installed and used to pose the SPARQL queries on the triples. However, the interlinking and identification of data is performed manually. A framework which automate the link discovery process can be used.

A survey of OGD in Russian Federation by Koznov et al [37] analysed the OGD trends by using the (OECD) Organization for Economic Co-operation and Development analytical framework. The study highlights the progress and implementation of OGD portals in Russia. Numerous e-services have implemented using OGD. However, it requires more efforts to redefine the process further for systematically utilization of OGD.

Recent evidence suggests that linked open data is very useful in linking the data of cities to make them smart cities [38]. Consoli et al. proposed a data model which integrates data from heterogeneous sources such as public transportation, road maintenance, municipal waste collection, Geo and urban fault reporting. This prototype linked data portal is available on-line, which helps the citizens to access information under free license and programmers can access the ontology and data via SPARQL queries. However, the public transport data can be further aligned with advanced open data standards which is able to capture more details of the traffic and passenger needs. In addition, this prototype can be aligned to the RDF data cube vocabulary, because, it focuses on publishing multidimensional data using (SDMX) statistical data and meta data exchange which is an ISO standard helpful in exchanging the meta data and statistical data among organizations.

A small scale study by Agarwal et al [39] was conducted to construct graphs by analysing the availability of open government data. The findings are highlighting the sectors such as animal husbandry and agriculture which can be used to create a semantic mesh. Moreover, basic framework for open data can be built using the findings of the survey.

F. Data Mapping and Visualization tools and Applications to enhance the consumption of LOD

The analysis shows that Mutuku et al.[40] conducted a study on an experiment which brings the subject matter experts in education, transport, water, and local country sectors with open data converters and software developers. The purpose of this experiment is to design an approach to identify OGD applications. The main aim of this study is to find the best practice to increase the consumption of open data via tools and mobile applications. The proposed idea of having an application or tool to access open data that can benefit the citizens in spite of their literacy level is unique, however, no technical and implementation details have been shared to achieve this toolkit.

A further example of work that uses RDF triples of open government data carried out by Hoxha [41]. Authors proposed to use XLWrap wrapper tool to create RDF triple of the open government data collected from diverse sources.The main objective is to provide transparency via data. Data has collected from various sources and semantic integration techniques are applied to ease the integration and publication of linked data. However, there is no evidence of the fact that from which particular sector's data sets have used to populate on the tool. The objective was to achieve data transparency and data visualization which has attained. Therefore, by adding the details of the RDF triple creation and population on web, the process can become more effective.

By drawing on the concept of DadosGov catalog, authors [42] have been able to show that Triplify tool can be used to convert the XML, (JSON) JavaScript Object Notation data sets into RDF triples. Triplify is used for conversion because this tool only takes relational database as input and produce output as RDF triples. The mapping file defines how database schema concepts must be presented in terms of RDF classes and properties. The creation of data mashup is time consuming and complex task. In order to make a comparison among the datasets of Brazil and USA another set of RDF triples have created by considering the vocabularies defined in data.gov and database schema of DataGov. Although, the finding shows that the tool doesn't provide necessary support during the conceptual modeling stage.

The study of linked open data is further carried out by Khalifa et al [43], who discussed linking and publishing of Linked data cloud and Arabic content on Web. Users can access complex semantic data by querying in Arabic language. Currently, there is no DBpedia chapter for Arabic DBpedia which can act as a source of knowledge. The proposed mapping of Arabic DBPedia with Arabic Wikipedia pages will help the users to access complex semantic data by querying in Arabic language. Further, natural language tools have used to extract information from Arabic DBPedia. However, there are some inconsistencies of design which can be improved by enhancing the knowledge of Arabic DBPedia. Moreover, to extract the English version of the datasets it can be linked with English DBPedia.

G. Linked open data Ontology implementation of public data

An, RDF triple by definition consists of subject, predicate and object which can be used to create the classes and properties of an ontology. The first systematic study of linked open data ontology of public spending using triples was reported by Vafopoulos [44]. It was designed by using data.gov.uk data portal. The input to the proposed architecture is given by Diavgeia which is a XML based (API) Application Programming Interface and first Greek Government Open Data Portal. Output can be seen via SPARQL endpoints. However, the ontology has very basic class and relationship definitions. The data properties, objects, concept-restrictions and rules have not been created for the proposed ontology.

In another study Theochairs, discussed an ontology [45] using protege 4.2 to link the public administration data. The built in reasoner of protege is used to find out semantic errors in the designed ontology. This study is an attempt to present a part of the ontology concerning the characteristics of administrative acts. Therefore, proposed ontology can be extended further by adding more concepts and their properties. It will contribute in forming a knowledge-base for the management and development of open data. Furthermore, human evaluation is conducted in the form of posing questions to the ontology using SPARQL endpoints. However, human evaluation can be challenging because, human are disposed to make mistakes.

Another example is an attempt to enrich the Greek e-GIF ontology [46] where protege is used to design the entities. A number of entities are added in the existing ontology and a comparison has done with (PSCs) Point of Single Contact of other European countries such as Cyprus, Malta, Spain and Slovak Republic. However, more attempts are required for further enrichment of the entities. A refined version of ontology along with good comparison results will be used for semantic enrichment of higher elements of the web pages with URI properties. This is vital for the conversion of open government data into Linked data.

Some further examples of linked open data are exploratory study of Brazilian initiatives based on the principles of linked open data is conducted by Ricardo et al [47]. Brazilian portals are three stars which means data sets are in XML, CSV and HTML. In addition, James et al [48] have done discussion on the data.gov portal and use of linked open data. They have focused on various sectors where linked data has utilized.

In an analysis of linked open data, Zhao [49] retrieved graph based ontology from the diverse data sources available publicly. The ontology alignment methods have applied to identify classes and properties of ontology from the data sets. Related classes and properties from different data sets are combined to find the missing "SameAs" links. This semi automatically created cohesive approach solves the heterogeneity problem of ontology. Moreover, it finds out the missing and wrong properties in the data sets. However, only four datasets have been selected which is DBPedia, GeoNames, NYTimes and LinkedMDB. In order to extend the alignment process, more data sets are required and Map Reduce method can be used to deal with big data sets.

In an attempt to semantify open data, Khalifa [50] proposed a lightweight approach for re-using existing ontologies from Hoxha [41]. The main objective was to contribute in the knowledge of semantic web and enabling data exchange and linking with other semantic sources over web. However, the conversion process was not fully automated. Furthermore, there is no open data portal available in Saudi which gives rise to the problem of data extraction. If common vocabularies to access the data are develop, it will help to link open data initiatives world widely.

Moroever, the OGDL4M ontology will be used for qualifying datasets in order to improve the accuracy of their legal annotation. The Ontology also aims to connect each applicable legal rule to official legal texts in order to direct legal experts and reusers to primary sources.

H. Challenges of Open data

In a study Mou et al [51] demonstrated the opportunities, challenges and negative impacts of open government data. The six dimensions open data policy of Taiwan has discussed. According to six dimensions policy, first, need of strong legislation policy for open data in Taiwan. Second, current data sets should be extended further. Third, cloud computing should be used as a fundamental facility for distribution of open government data. Four, a single portal to access the government open data. Fifth, public awareness is must to utilize the open data and lastly, open data should be free to the public. There should not be any license fee for downloading the data.

I. Open data Management Tools and Activities

In another study Bojars [52] discussed (PDH) Peak data hackathons, which is an activity initiated by group of volunteers. The main purpose of PDH is to use available open data and transformation of data into user friendly form. After completion, the data has been submitted to data.opendata.lv. However, the work-done by the volunteers is not static because, they have no right to hold and maintain the data. The datasets used does not come under open license and not up to date. Therefore, an automation process is required to update the catalogue based on Meta data related with open data.

In a case study focused on evaluation, effectiveness and capabilities of open data of Canadian Municipal level [53] Roy has analysed open government data and open data governance in Canada with particular focus on Municipal governments. This will bring systematic transparency and overcome the challenges of conceptualization, role of media, political and data culture and holistic governance framework. However, open and innovative governance is required to bring progression in open data movement.

One of the most significant current discussion is that, ODI can be managed using online APIs, tools and websites. A survey conducted by Correa et al [54] over 20 municipalities to analyse the access to information law in Brazil. The findings suggest that 10% of the documents are in open format. The usage and importance of (CKAN) Comprehensive Knowledge Archive Network [55], which is used to manage and publish Open government data is also highlighted. A number of research institutions, local and national governments are using CKAN. The published data can be previewed in the form of graphs, tables and maps. It is an open source data portal platform software, so it is used by a number of countries. Internal modes of CKAN has been used to store the Meta data about diverse sources and it has presented on the web so that users can search the data. However, challenges of change in culture still have to be dispersed across local governments. Table IV provides the complete information related to the work done in linked open data.

Further, a study of open government data [69] discloses that legal requirements which are unspecified make open data complex to use. The perplexity of legislation requirements indicates the necessity of having an ontology that facilitates the analysis of the licenses, conditions and legal notice of OGD. These analysis techniques can be automatic or semi-automatic.

The work [70] proposes a prototype search engine for automatic and manual linking of the concepts in the transport domain ontology. The result has shown quality search and more efficient open data search. However, the proposed prototype can be improved by adding more datasets as well as by adding accurate metadata description.

More recent work by [71] used the water supply management datasets of Valencia(Spain) to generate an ontology based framework for publishing the linked open data. The proposed ontology is capable of identifying the correlations among the water supply, leakage and population. However, the interlinking process can be improved by adding more data sets of different domains.

IV. Benefits of Linked Open data

A. Transparency

Linked open data brings transparency among the various sectors of the government. In addition, common citizens of the country would be able to see the performance of the various government departments. Every sector's data will be accessible at one place which will help the users to navigate amid the links to extract the useful information [56]. The citizens be able to participate in the political and social activities because, once the information is readable and understandable it will encourage the citizens to show more interest in political and other social affairs.

The linked open data will create trust among the general public towards the government [32]. Everyone will get equal access to the information and it can help the government to improve the services and process for the citizens. It means the stakeholders can see, use and distribute the data freely. In order to achieve a transparent democracy, the country should open up the data so that the stakeholders can see the activities of the government [57].

B. Public Participation in Government

Open data provides an opportunity for the public to participate widely in the government activities. The general public can impose questions on certain areas. It will help the government make fair policies and to take better decisions [58]. For instance, if the government wants to introduce new changes in current policies and intends to seek the viewpoints of the public by disclosing the operational cost and accompanying benefits, it will help the people to make a clear and firm decision in terms of supporting or opposing the decision [59].

C. Social Value

Government is responsible for handling large amount of data relating to education, health, environment, agriculture, employment and budgets. By opening these data to the public, government can encourage people to use it and invent new ideas, thus helping others create social value [60].

D. Reliability

Open data initiative will lay down the foundation of a strong and open government which work and care for people. The data available by the government will help people to believe in the government. They can take clear decisions based on the data sources provided by the government [61].

E. Economic Growth

One of the biggest benefit of linked open data will be economic growth and stimulation of competitiveness. It will stimulate the innovation and contribute towards the improvement of various processes services and products. Moreover, the information will be available for various investors and companies which will help them to invest their money wisely and ultimately, it will add value to the economy [62].

F. Data Sustainability and Re-usability

This is the most generic criteria for quality of open data. By re-usability means how easily the published data can be reused. It means the data sets collected once can be used multiple time without recollecting the data. This will reduce the unnecessary duplication and associated cost. In addition, open data will bring sustainability of data [57].

G. Decision Making

Open data will bring the tremendous benefit of fair decision making by enabling the comparison of various associated data sources. The new data sources can be created by combining the available or existing data sources [13].

H. Integration and Availability of Information

Open data will bring the concept of information availability. All vendors and companies can access the useful information from the data portals. The operational and technical benefit of open data is to integrate, merge and mesh private and public data [63].

V. Challenges of Linked Open Data

A. Data Protection and Privacy

There is conflict between the aim of transparency, accountability, open data and data protection. Even though the data is nondescript before merging and publishing, still possibility result in the discovery of some personal information. For instance, if the data of garbage collection is published along with personnel timetable, in that case it will easy for the data consumer to identify the route of a particular employee. Therefore, this issue requires more research in order to come up with some predefined guidelines that can provide solution to this privacy concern. Moreover, if some restrictions will be imposed on the data access, can solve this problem [64]. However, applying this approach will restrict the openness of the open data on several levels.

B. Complexity

Open government initiatives are growing so rapidly but, still there is lack of ability to discover the appropriate data. Data consumers can only access the processed data. Moreover, the meaning and explanation of the data is highly complex. Users have difficulty in browsing and searching due to lack of index or other useful means to ensure easy search for finding the right data. The data format and datasets are too complex to read and use. There is no support to extract the information from such big initiatives. There is focus on one datasets therefore, the actual benefit will come from linking various data sets together [16].

C. Lack of knowledge

The users of the open data have no motivation to access the information. There is lack of knowledge among the citizens to make use of the data. They are lacking in capability to use the information. Further, they are not keen in the data due to several reasons such as license fee, big data initiatives, redundant registrations before download, no statistical knowledge, unexpected intensified cost, unsupported behaviour of the public organizations and threats of lawsuits [59].

D. Legislation

Although most of the open government initiatives fit in to existing legal frameworks, yet there is lack of an open government policy. The issue lies in security and privacy violation [65]. Most of the countries doesn't have any license on the use and consumption of open data while, others are preferring to have written permission required prior to gain access to the data. In this way, an authenticated person will get the access. The written agreement between the data consumers and legal entities will bring security. Moreover, if any one breach the agreement, will be liable for punishment [66].

E. Quality of the Information

The quality of the information available in the data initiatives are not appropriate. Most of the data is invalid and obsolete. The incomplete information is a challenge to tackle, when part of the information is visible and rest of the information is missing or not loading [67]. However, the information may appear irrelevant when viewed in isolation, but when linked and analysed it can result in new insights. Moreover, similar information has stored at several places which bring confusion among the data consumers.

F. Technical

The open data initiatives are facing plenty of technical challenges. It is problematic to access the data if it is not readable by the machine and the format is not well defined. Currently, there is no architecture and standards for processing of the open data [68].

VI. Motivation

The tabular data from the real world is revealed over the web to provide an open platform for linking to other data resources. As a result, these information is stored in obsolete data format. Due to their complexity, it is difficult to import data into a web context in order to conduct reasoning and analysis over it. In order to carry out reasoning and analysis, the machine must be able to comprehend the semantic details of the data. Semantic links bring the concept of linked open data, in which URIs have been created between data files so that different department's data can be linked together based on same fields available in the data files. This will solve the problem of missing links and data will be machine readable. More than 70% of the data on the web today are unstructured text [13],[14]. This applies to government data as well where, large quantity of data is hindered in natural language text documents making it difficult for humans to understand quickly.

For example, if a new entrepreneur wants to buy a business and he is not sure exactly how to select an appropriate business which will bring more profit to him. LOD could be helpful in making the decision in selecting the right business. For achieving this, the data from business and financial sectors would be linked using LOD principles which will graph the progress of the business and articulate the year in which the business had reached a significant success rate. Thus, end users would be able to make worthy decisions. Moreover, LOD will provide transparency to general public, so that individuals can get to know what is happening in the country.

For instance, by making annual financial spending of the data available as LOD the public will have access to the data and they can easily figure out where the money is being invested. As a result, the government bodies need to justify their operations by providing data sets which will give a clear picture of all spending and decrease the chances of bankrupt and embezzlement. Hence, there is a need to have an approach that can assist the interlinking of diverese datasets where semantic encoding can be achieved.

VII. The Proposed Design Architecture

The Figure 5 depicts the proposed architectural design. The whole process is divided into three phases:

- CSV to OWL Conversion
- Generating the Semantic link
- Implementation of the SPARQL interface

The CSV to OWL conversion process starts when the user enters either URL of the CSV file or upload's the CSV file directly. User has to mention the name of the CSV file, later this name will be used as the name of the ontology. The subsequent concepts associated with each CSV file are created. To read the CSV files effectively, Apache commons CSV library is used, as it defines the format for the CSV file. Here default mode is used to parse the CSV files so that all headers and records of the CSV files can be parsed successfully. Once the CSV file is parsed, interpreted CSV tables are generated, the CSV to OWL converter process is initiated, which helped in transforming the CSV datasets into OWL triples. These OWL triples are stored in the local memory of the system which can be visualised using protege tool. To generate the semantic link the user must choose two or more ontology files. Because ontology files can contain many ontology classes, it is necessary to select the appropriate class from the ontology file. Additionally, the user must select common data properties from both ontology classes in order to perform semantic link operations. Once the user has generated all of the aforementioned ontologies. The user can choose between two distinct ontologies and create semantic links between them. Following the selection of two ontology files, the files must be combined to produce a union. In this case, the union contains all ontology classes, individuals, and data properties.The primary goal of generating the union file is to retain all information contained in the ontology files in order to avoid data loss. The class of the second ontology file is converted to tabular form as step of the procedure by employing the HSQL Database. The purpose of this conversion is to simplify the search operation of the semantic links by utilising the tabular form, which is simple to traverse when performing a linear search. The following step is to iteratively/sequentially retrieve the individuals from the first ontology class. Additionally, the data properties of the first ontology class individuals are compared with the tabular data by performing a linear search operation so that semantically linked individuals of the second ontology class can be identified. Individuals may have duplicate values, making it difficult to semantically link these datasets based on common properties. As a result, an appropriate approach for linking these ontology classes is based on a condition that determines the literal values of data properties of ontology classes that are similar. We overcame this issue by utilising the cosine similarity measure, which includes methods for comparing two strings and returning the similarity score. The linear search method locates individuals inside tabular data in order to establish semantic links between selected common properties. It will continue to browse through each piece of data progressively until a matching is discovered or the full table is searched. The produced semantic links are then integrated with the union file's second ontology class individuals. As a result, a semantically linked ontology for two distinct domains or data sectors will be created. Moreover, the semantic linking process is one directional only. In future versions, bi-directional linking will be considered.

Once we have constructed our semantically linked ontologies. The third and last phase is to use the SPARQL interface to query the produced ontology. The SPARQL interface is structured in such a way that the user can choose the ontology file to query. To query in SPARQL, the created OWL ontologies must be converted to an RDF format, such as Turtle. Composing SPARQL queries that contain complex OWL expressions varies in complexity from difficult to inconvenient due to the fact that SPARQL query syntax is built on Turtle [72], which is not designed for OWL. We used Apache Jena for the conversion process, a Java package that converts OWL files to RDF Turtle format and offers APIs for querying SPARQL from within a Java application. A diagram of SPARQL interface is illustrated in Figure 5.

VIII. Case Study

The approach described in Section VII has been evaluated using the data sets of agriculture, land and rainfall from New Zealand Government. The aim of our work is to introduce a prototype which will assist in automating the process of ontology creation by using government datasets. The datasets are extracted from URL https://data.mfe.govt.nz. The file parser parsed the entered URL, and the CSV datasets are converted to OWL files using an OWL converter. All OWL files are saved locally in the system memory. A visual representation of agriculture ontology can be seen in the Figure 6. Here Owl:Thing is the main/default class and Agriculture is the subclass ot it which is holding five data properties such as area_ha, farm_type, region, FID, year and 620 individuals. When the mouse pointer is hovered over an individual, the data property assertions for that individual are highlighted.

Likewise, an ontology is developed and stored for land and rainfall datasets. Furthermore, these generated ontologies will be semantically linked so that knowledgeable data may be retrieved by implementing SPARQL queries. To perform this task, two ontology files from the local hard drive of the system are selected, and then the shared properties are manually identified. The common data properties in both ontologies are year and region. As a result, these data properties are chosen for the generation of semantic links.

Furthermore, we do not want the ontology files to lose any information, thus a union file is formed by joining both of the ontology files. This union file was set aside and eventually combined with the resulting ontology. Here we have selected agriculture ontology as first ontology file and land as second ontology file. If the ontology has more than one class, we must ensure that we select an appropriate class from the ontology files. We selected the agriculture and land classes from the Agriculture and Land OWL file respectively. Individuals from the agricultural class are retrieved one at a time, whereas the land class has been converted into a tabular form to simplify the semantic links search operation because the tabular form is simpler to traverse while executing a Linear search operation.

Finally, the semantic links are created using a linear search on individuals in the agricultural class and tabular data from the land class. To avoid duplication, cosine similarity is employed. Prior to getting the full version of the linked ontology. The previously formed union file is merged with the resulting ontology file, allowing us to have both semantic and non-semantic data together.

Figure 7 shows a photographic image of the semantically linked ontology of agriculture and land datasets. The dashed arrow lines represent the semantic link among both the individuals of the land and agriculture classes. There are 892 individuals in total with 168 semantic links. We are unable to display all of the semantically linked individuals due to screen size constraints.

To evaluate our methodology, we must first examine the accuracy of the created ontology. We did this by feeding the framework certain sample SPARQL queries. The results to certain queries are recorded, and the consistency of those results is carefully assessed. SPARQL queries are performed to datasets to collect relevant information. Traditional data extraction approaches are time-consuming as they require a comprehensive reading of the datasets; however, as all data is stored as triples, an RDF query language makes the process simple. It is straightforward to find accurate data by using SPARQL queries.

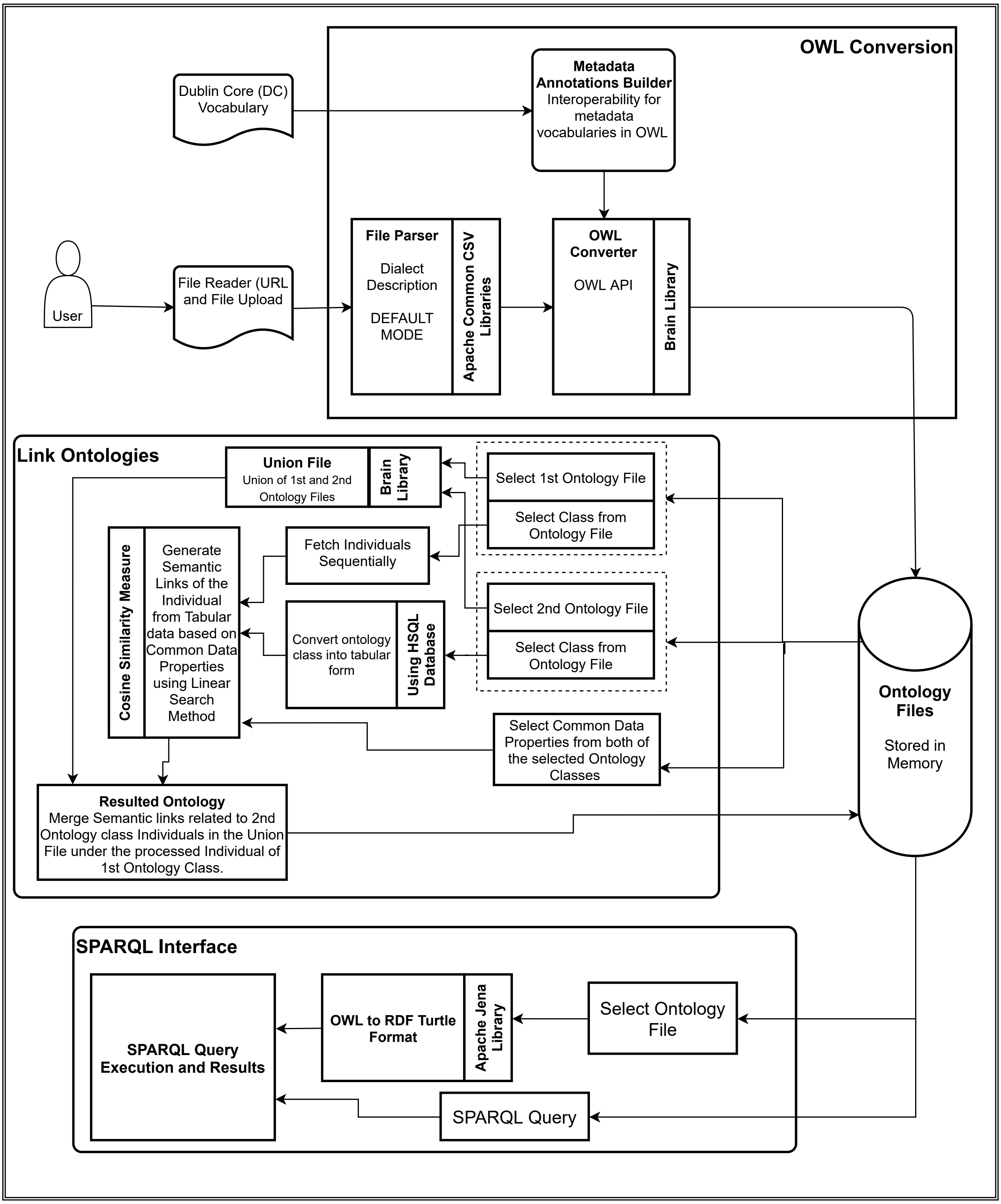


Fig. 5: The Proposed Architecture

TABLE IV: Provides the complete information related to the work done in linked open data

Year	Author	Datasets Used	Evaluation Method	Techniques
2009	Hassanzadeh et al.	Freebase, OMDB, DBPedia movies, Rotten tomatoes.com, Stanford movie database	String matching and precion recall	Weighted Jaccard, Jaccard, Edit Similarity BM25, HMM (Hidden Markov Model), Cosine w/tf-idf
2010	Parundekar et al.	LINKEDGEODATA, GEONAMES, DBPEDIA, GEOSPECIES, MGI, GENEID	Empricial Evalaution	Alignment algorithm
2010	Thanh Le et al.	US Census, GeoNames, DBPedia, World Factbook	Experimental Evaluation	Graph Mining techniques is used in the Algorithm to detect hidden relationship by comparing Naïve algorithm
2011	Kalampkois et al.	Moraitis School, The 2nd local Directorate of Secondary Education of Athens, The ministry of Education	Prototype implementation	Silk framework, D2R server, SPARQL interface, RESTful APIs
2011	Liu et al.	Energy, Population, Economic, Environment (Rainfall and temperature data)	Prototype based Case Study	RPI data conversion tool, TWC LOGD portal, cell based conversion
2011	Haslhofer et al.	Libraries, Archives, Museums	Europeana LOD Prototype	Open Link Virtuoso, Dereferencing HTTP URIs, RDF Mapping, SPARQL Queries
2011	Kontokosras et al.	Greek Dbpedia	Dbpedia Informtion Extraction Framework	Infobox mapping and properties, Inter language link extractor, Inter Dbpedia linking, IRI Serialization, Transparent Content negotiation rule
2012	Hoxha et al.	OGD of Albania	ODA ontology	XLWrap Wrapper, Ajax, SPARQL, Google Visualization API, Spark (Javascript library)
2012	Breitman et al.	OGD of Brazil related to Population	DataMashups	Triplify, stdTrip
2012	Vafopoulos et al.	Public spending datasets	Ontology	Virtuoso Jena Provider, Python libraries, Diavgia API, Java libraries, XML static instances
2013	Galiotou et al.	ERMIS Government portal for Public Administration	Case Study	Jena Framework, Open linkVituoso, SPARQL Endpoints
2014	Fragkou et al.	ERMIS Government portal for Public Administration	Enrichment of e-GIF ontology	Protégé
2014	Theocharis et al.	Public Administration Greek	Ontology development	Protégé

Continued on next page

TABLE IV – continued from previous page

Year	Author	Datasets Used	Evaluation Method	Technique
2014	Zhao et al.	Dbpedia, Geonames, NYTimes, LinkedMDB	Graph based ontology Analysis	SameAs Graphs extraction Algorithm, Related classes and properties grouping, Aggregation of all integrated classes and properties
2014	Gonzalez et al.	Mexican federal budget, investment Portfolio	Theoretical framework for Data Visualization	Tableau software Public version 8.1
2016	Galiotou et al.	Agriculture, Wholesale and retail, Transportation, Accommodation and food services, Education, Arts, entertainment and recreation, Construction, Mining and Quarrying	interface implementation using Jena	Jena Framework, FUSEKI, SPARQL Endpoints
2016	Azevedo et al.	ANA, ANEEL, IGAM, CPRM, CEMIG, Transparency portal for MG, IBGE, Health Portal, PNUD, Open data Portal, Geonames, DBPedia, Generic data sources (such as Google, Wikipedia, sciencedirect etc.)	case study-based experiment	D2RQ platform, GIS for Visualization of Map and SPARQL Queries, DB4Trading
2017	Martynas Mockus, Monica Palmirani	European Union (EU) legal framework of reuse of Public Sector Information (PSI), the EU Database Directive and copyright framework and other legal sources (e.g., licenses, legal notices, terms of use)	Open Government Data Licenses Framework for a Mash-up Model	MeLOn, RDF, SPARQL Queries, OWL reasoner
2019	Shanshan Jiang, Thomas F. Hagelien, Marit Natvig,Jingyue Li	Transport domain datasets	A prototype search engine for manual and automatic linking of the concepts in the ontology	Semantic, NLP (Natural Langauge Processing), Semantic and Machine learning techniques
2020	Pilar Escobar, Maria del Mar Roldan-Garcia, Jesus Peral, Gustavo Candela and Jose Garcia-Nieto	Water Supply management of Valencia (Spain)	An Ontology Framework to generate semantically enriched linked data	RDF, OWL, SPARQL, Data Mapping, Preprosesing, Data Modelling, storage and Exploitation.

For testing purposes, we applied the following SPARQL queries to the resulting ontologies:

1. Find the rainfall rate of wellington region in 2007 and what type of farming activities are carried out in that year?

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT DISTINCT ?ind1 ?ind2 ?Rainfall ?farm_type
WHERE
{
?ind1 rdf:type owl:NamedIndividual .
?ind2 rdf:type owl:NamedIndividual .
?ind1 ?linked ?ind2 .
?ind1 rdf:type <http://www.ontogen.org/Rainfall/Rainfall>.
?ind1 <http://www.ontogen.org/Rainfall/Year>2007 .
?ind2 rdf:type <http://www.ontogen.org/Agriculture/
Agriculture>.
?ind2 <http://www.ontogen.org/Agriculture/year>2007 .
?ind2 <http://www.ontogen.org/Agriculture/region>
”Wellington Region”.
?ind1 <http://www.ontogen.org/Rainfall/r95ptot>?Rainfall.
?ind2 <http://www.ontogen.org/Agriculture/farm_type>
?farm_type
}
```

2. Find the region with maximum rainfall rate (r95ptot) in year 2012. Also find the farming activities for that region

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT ?ind1 ?ind2 ?year ?site ?farming ?rainfall_rate
WHERE
{
?ind1 rdf:type owl:NamedIndividual .
?ind2 rdf:type owl:NamedIndividual .
?ind1 ?linked ?ind2 .
?ind1 rdf:type <http://www.ontogen.org/Agriculture/
Agriculture>.
?ind2 rdf:type <http://www.ontogen.org/Rainfall/Rainfall>.
?ind2 <http://www.ontogen.org/Rainfall/site>?site .
?ind2 <http://www.ontogen.org/Rainfall/r95ptot>
?rainfall_rate .
?ind1 <http://www.ontogen.org/Agriculture/farm_type>
?farming.
FILTER (?rainfall_rate =?maximumrainfall) .
?ind2 <http://www.ontogen.org/Rainfall/Year>?year .
{
SELECT (MAX(?rainfall) AS ?maximumrainfall)
WHERE
{
?ind3 rdf:type <http://www.ontogen.org/Rainfall/Rainfall>.
?ind3 <http://www.ontogen.org/Rainfall/Year>?year.
?ind3 <http://www.ontogen.org/Rainfall/r95ptot>?rainfall.
FILTER(?rainfall != ”NA” && ?year =2012)
}
}
}
```

Figure 8 and 9 highlights the results for the query one and two, where semantically linked ontology of Rainfall and Agriculture is used to impose the SPARQL queries so that useful information can be extracted. A manual analysis is undertaken for accuracy and consistency. Following a manual review, it was discovered that the system's results were accurate.

As a result, our SPARQL interface interacts with ontologies that are semantically interconnected and captures the desired results. At this point, we're using SPARQL queries to evaluate our results. We want to perform an expert evaluation of the proposed model in the future to ensure its performance and robustness. The individuals highlighted in Figure 8 can be accessed by clicking on the blue hyperlink. To access the data properties, one needs to click the individuals either under ind1 or ind2, corresponding data properties, will be highlighted. As Figure 10 highlights all data properties corresponding to the 3rd individual under ind1 of Figure 8.

IX. Discussion

Currently, the open data is gaining momentum as all countries are looking for approaches and machenisms to make this data knolwedgebale to the general public. An ontology framework for New Zealand open data will contribute in the research by opening up path for data encoding and semantic queries. However, designing such an ontology is quite challenging because of heterogeneous data sources and multiple formats used to publish them. It is difficult to analyse big data sets and discover relevant data fields to generate the triples. For an initial experiment, data sets from agriculture, land and Rainfall sectors have encoded and designed in the form of ontology. This preliminary design would be used as a prototype to encode and link other sectors of the government. Moreover, other countries can take advantage by selecting this design and can semantically link their Open government data initiatives. However, it would be difficult for non-English countries to use this framework if they don't have English version of the data sources to be linked.

X. Conclusion

This paper showed that Linked open data has substantially increased in momentum in recent years with an increasing number of countries opening up their data for public consumption. For instance, UK government is pioneering to create a web of linked open government data which has opened up the data for the developers so that they can use it for economic and other benefits. Furthermore, in this paper we have also discussed the growth of various OGD initiatives, benefits and limitations of Linked Open Data. While, Open Data has several benefits, it also entails numerous barriers in the fields of technology, legislation, use and complexity. Just by itself, open data had has no value, as it only becomes valuable when utilized in an application, primarily for decision support.

There is an abundance of misconceptions about the linked open data, such as, all information should be published unrestrictedly and publishing of open data will automatically bring transparency. These misconceptions are used a lot to convince data providers to open up their data to the public however, it ignores the fact that there are numerous

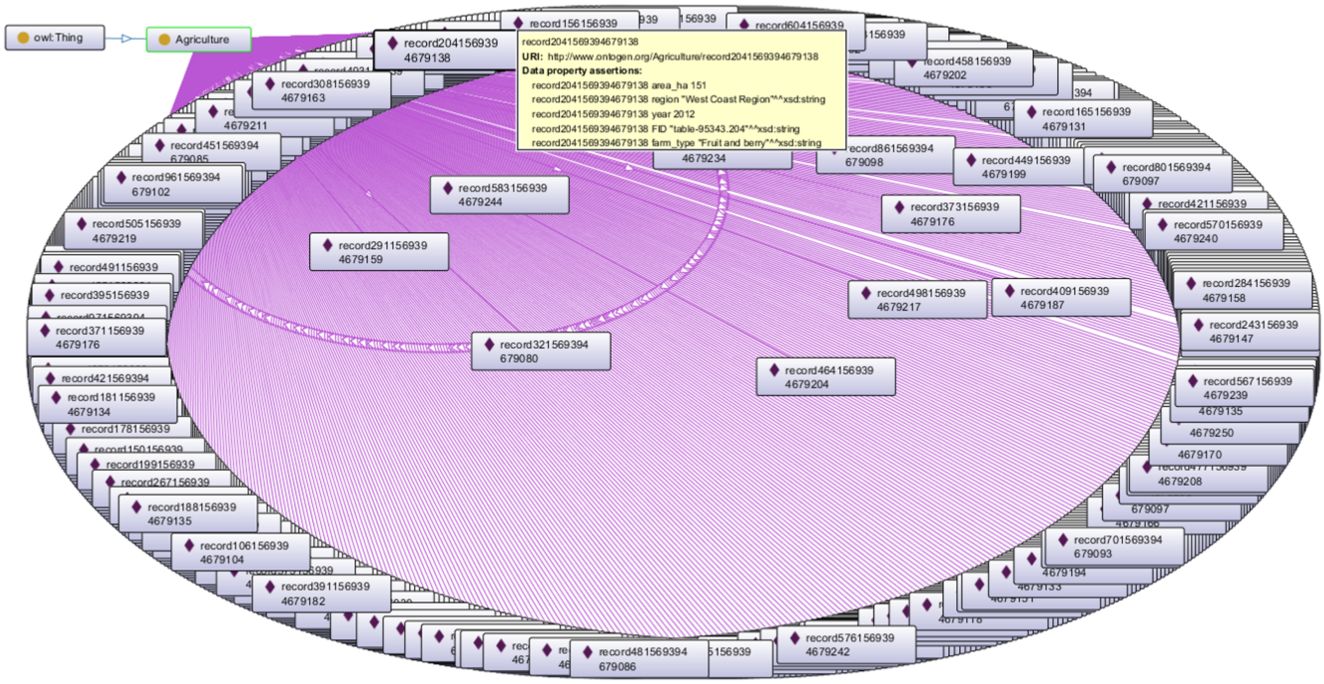


Fig. 6: The Agriculture Ontology

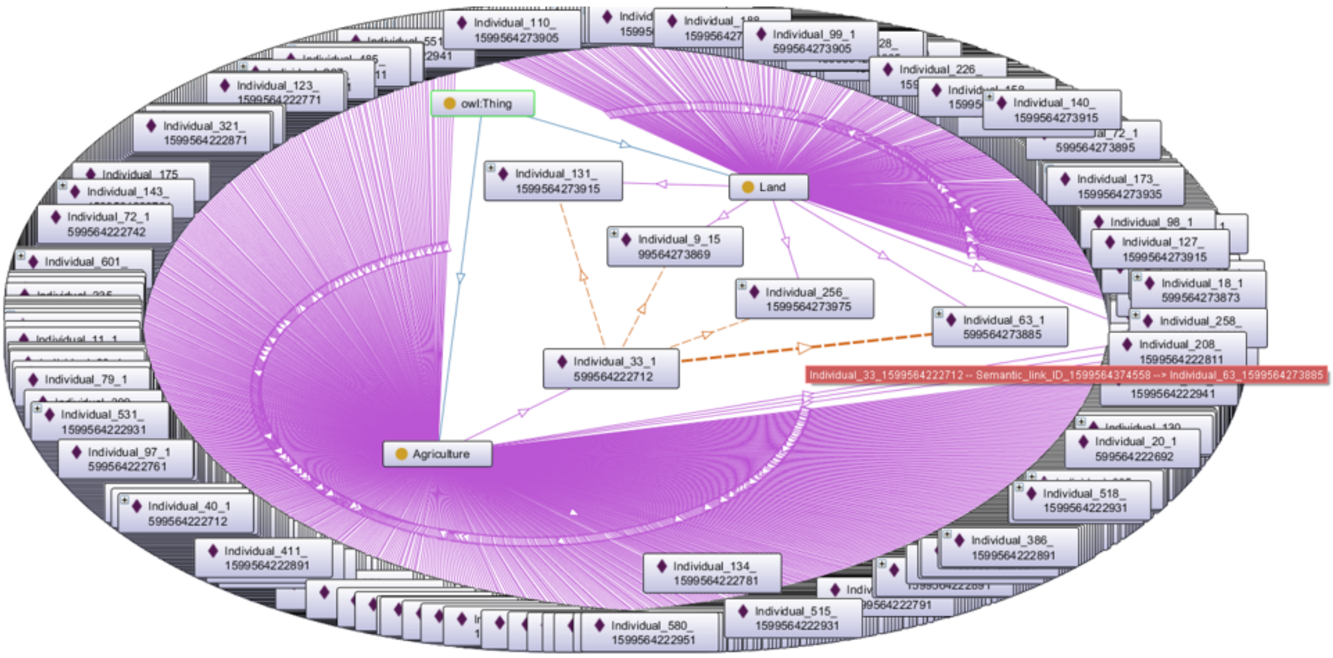


Fig. 7: The Semantically Linked Ontology of Agriculture and Land

limitations due to the fact of the heterogeneous nature of the open data. Our analysis of literature showed that OGD supports creative use of government data which in turn increases governance transparency. Identification of potential OGD datasets also helps public bodies to better understand and manage their own data. It also provides a transparent way of notifying the results of the governmental regulations. Currently the main focus is the supply of the data, however, the success of linked open data depends on the quality and use of the data. It is of utmost importance for governments to work on policies and procedures towards the usage and access of linked open data. Governments have to take a broader perspective of having such infrastructure which would help users to access the data and use it for productive purposes. Such policies and infrastructure would promote the current level of engagement of ordinary citizens by increasing transparency and trust. In order to get full benefits from linked open data, it is vital to put data and information together in a context which would create novel knowledge and enable other potential useful applications and services. Linked open data facilitates knowledge and innovation of the nexuses in the data, which is the prime mechanism for information integration and Management.

ind1	ind2	Rainfall	farm_type
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record4411599560713434	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Other Livestock
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record1201599560713126	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Forestry
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record251599560713011	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Dairy
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record1871599560713195	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Fruit and berry
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record5801599560713547	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Vegetable growing
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record2511599560713268	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Grain growing
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record3141599560713335	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Nursery and turf
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record3761599560713382	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Other
http://www.ontogen.org/Rainfall/record15301569394651113	http://www.ontogen.org/Agriculture/record5081599560713484	16.29671789^^http://www.w3.org/2001/XMLSchema#double	Sheep and Beef

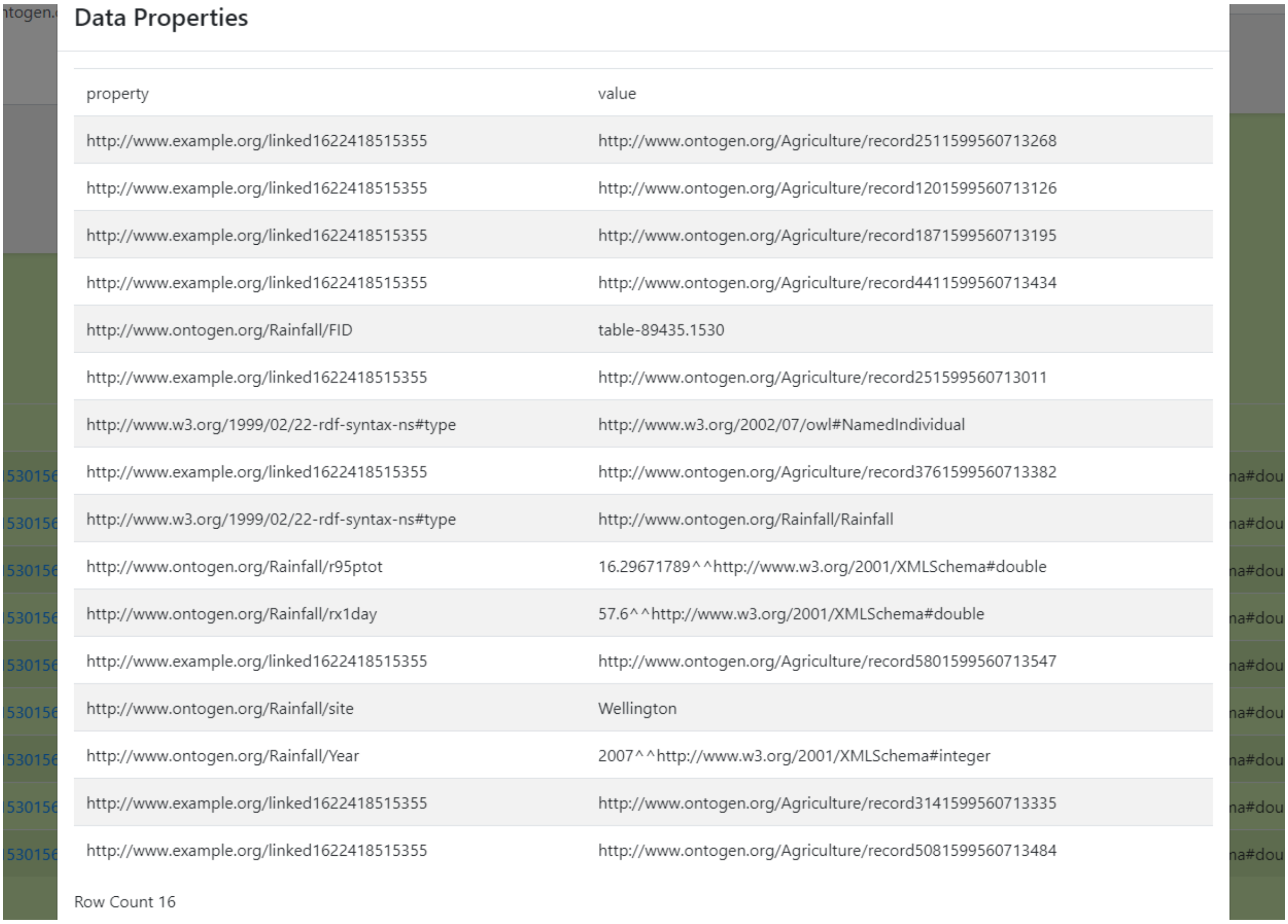
Row Count 9

Fig. 8: Result of 1st Query for Rainfall and Agriculture Ontology

ind1	ind2	year	site	farming	rainfall_rate
http://www.ontogen.org/Agriculture/record3891569394679181	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Other	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record5931569394679248	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Vegetable growing	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record2631569394679153	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Grain growing	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record371569394679082	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Dairy	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record5211569394679224	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Sheep and Beef	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record4541569394679200	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Other Livestock	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record3261569394679167	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Nursery and turf	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record1331569394679123	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Forestry	42.49050167^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Agriculture/record1991569394679137	http://www.ontogen.org/Rainfall/record3381569394650800	2012^^http://www.w3.org/2001/XMLSchema#integer	Gisborne	Fruit and berry	42.49050167^^http://www.w3.org/2001/XMLSchema#double

Row Count 9

Fig. 9: Result of 2nd Query for Rainfall and Agriculture Ontology



Data Properties

property	value
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record2511599560713268
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record1201599560713126
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record1871599560713195
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record4411599560713434
http://www.ontogen.org/Rainfall/FID	table-89435.1530
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record251599560713011
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#NamedIndividual
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record3761599560713382
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.ontogen.org/Rainfall/Rainfall
http://www.ontogen.org/Rainfall/r95ptot	16.29671789^^http://www.w3.org/2001/XMLSchema#double
http://www.ontogen.org/Rainfall/rx1day	57.6^^http://www.w3.org/2001/XMLSchema#double
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record5801599560713547
http://www.ontogen.org/Rainfall/site	Wellington
http://www.ontogen.org/Rainfall/Year	2007^^http://www.w3.org/2001/XMLSchema#integer
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record3141599560713335
http://www.example.org/linked1622418515355	http://www.ontogen.org/Agriculture/record5081599560713484

Row Count 16

Fig. 10: Data Properties of 3rd individual under ind1

REFERENCES

[1] Rita Fleiner, "Linking of Open Government Data", *IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pp1-5, 2018
[2] S. K. Sowe and K. Zettsu, "Towards an Open Data Development Model for Linking Heterogeneous Data Sources", *Seventh International Conference on Knowledge and Systems Engineering (KSE)*, pp344-347, 2015, doi:10.1109/KSE.2015.56
[3] Peter R. Orszag, Open Government Directive, https://obamawhitehouse.archives.gov/open/documents/open-government-directive,2009, [Online; accessed 23-Feb-2018]
[4] Gov UK. Open Government Partnership. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf, 2013, [online; accessed 20-Oct-2018]
[5] DC Office Washington. Open Data Charter. https://www.opengovpartnership.org/, 2011, [online; accessed 20-Dec-2018]
[6] Gov NZ. Open Government Data Programme. https://www.data.govt.nz/standards-and-guidance/open-data/open-data-nz/, 2015, [online; accessed 20-Oct-2018]
[7] Christian Bizer, Tom Heath, and Tim Berners-Lee. "Linked Data-the Story so far", *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, pp205-227, 2011, IGI Global
[8] Tom Heath, and Christian Bizer. "Linked Data: Evolving the Web into a Global Data Space", *Synthesis Lectures on the Semantic Web: Theory and Technology*, vol.1, no.1, pp1-136, 2011, Morgan & Claypool Publishers
[9] Francis Zinke, "Relational Database: A Practical Foundation", 2009, Citeseer
[10] Hitzler, Pascal and Krotzsch, Markus and Rudolph, Sebastian, *Foundations of Semantic Web Technologies* (Book style), New York, Chapman and Hall/CRC Press, 2010
[11] Molood Barati, Quan Bai, and Qing Liu, "Mining Semantic Association Rules from RDF Data", *Knowledge-Based Systems*, vol.133, pp183-196, 2017
[12] W3C, What is RDF Triplestore?, https://ontotext.com/knowledgehub/fundamentals/what-is-rdf-triplestore/, [Online; accessed 20-Sep-2017]
[13] Jan Kucera, and Dusan Chlapek, "Benefits and Risks of Open Government Data", *Journal of Systems Integration*, vol.5, no.1, pp30-41, 2014
[14] Barbara Ubaldi, "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", *OECD Working Papers on Public Governance*, no.22, pp60, OECD Publishing, 2013, https://doi.org/10.1787/19934351
[15] Oktie Hassanzadeh, "Introduction to Semantic Web Technologies & Linked Data", *University of Toronto*, 2011
[16] Florian Bauer, and Martin Kaltenbock, "Linked Open Data: The Essentials", *Edition mono/monochrom, Vienna*, vol.710, 2011
[17] Axel-Cyrille Ngonga Ngomo, Soren Auer, Jens Lehmann, and Amrapali Zaveri, " Introduction to Linked Data and its Lifecycle on the Web", *Reasoning Web International Summer School*, pp1-99, 2014, Springer
[18] Nicola Guarino, Daniel Oberle, and Steffen Staab, " What is an Ontology?", In *Handbook on Ontologies*, pp1-17, 2009, Springer
[19] A. Taychatanompong, and W. Vatanawood, "Sales Forecasting using Ontology", *Proceedings of the International MultiConference of Engineers and Computer Scientists*,pp416-420, 2019, Hong Kong
[20] Nansu Zong, Sejin Nam, Jae-Hong Eom, Jinhyun Ahn, Hyunwhan Joe, and Hong-Gee Kim, " Aligning Ontologies with Subsumption and Equivalence Relations in Linked Data", *Knowledge-Based Systems*, vol.76, pp30-41, 2015
[21] Vicknair, Chad and Macias, Michael and Zhao, Zhendong and Nan, Xiaofei and Chen, Yixin and Wilkins, Dawn, "A Comparison of a Graph Database and a Relational Database: A Data Provenance Perspective", *Proceedings of the 48th Annual SouthEast Regional Conference*, pp1-6, 2010
[22] Oktie Hassanzadeh and Mariano Consens, " Linked Movie Database", *Linked Data on the Web (LDOW2009)*, 2009
[23] Rahul Parundekar, Craig A Knoblock, and Jose Luis Ambite, " Linking and Building Ontologies of Linked Data", In *International Semantic Web Conference*, pp598-614, 2010, Springer
[24] Ngoc-Thanh Le, Ryutaro Ichise, and Hoai-Bac Le, " Detecting hidden relations in geographic data", In *Proceedings of the 4th International Conference on Advances in Semantic Processing*, pp61-68, 2010, Citeseer
[25] Evangelos Kalampokis, Efthimios Tambouris, and Konstantinos Tarabanis, "A classification Scheme for Open Government Data: Towards linking Decentralised Data", *International Journal of Web Engineering and Technology*, vol.6, no.3, pp266-285, 2011
[26] Qing Liu, Quan Bai, Li Ding, Huong Pho, Yun Chen, Corne Kloppers, Deborah L. McGuinness, David Lemon, Paulo deSouza, Peter Fitch, and Peter Fox, " Linking Australian Government Data for Sustainability Science- A Case Study", *Linking Government Data*, pp181-204, 2011, Springer
[27] Bernhard Haslhofer, and Antoine Isaac, "data.europeana.eu: The Europeana Linked Open Data Pilot", In *International Conference on Dublin Core and Metadata Applications*, pp94-104, 2011
[28] Stamatios A Theocharis, and George A Tsihrintzis, " Open Data for E-government the Greek Case", In *Information, Intelligence, Systems and Applications (IISA), Fourth International Conference on*, pp1-6, 2013, doi: 10.1109/IISA.2013.6623722
[29] Dimitris Kontokostas, Charalampos Bratsas, Soren Auer, Sebastian Hellmann, Ioannis Antoniou, and George Metakides, "Towards Linked Data Internationalization-Realizing the Greek Dbpedia", pp1-4, 2011
[30] Charalampos Alexopoulos, Lefkothea Spiliotopoulou, and Yannis Charalabidis, "Open data Movement in Greece: A Case Study on Open Government Data Sources", *Proceedings of the 17th Panhellenic Conference on Informatics*, pp279-286, 2013
[31] Juan Carlos Gonzalez, Jaime Garcia, Federico Cortes, and David Carpy, "Government 2.0: a conceptual framework and a case study using Mexican data for assessing the evolution towards open governments", In *Proceedings of the 15th Annual International Conference on Digital Government Research*, pp124-136, New York, USA, 2014, ACM Press
[32] Judie Attard, Fabrizio Orlandi, Simon Scerri, and Soren Auer, "A Systematic Review of Open Government Data Initiatives", *Government Information Quarterly*, vol.32, no.4, pp399-418, 2015, Elsevier
[33] Pavlina Fragkou, Nikolaos Kritikos, and Eleni Galiotou, " Querying Greek Governmental site using Sparql", In *Proceedings of the 20th Pan-Hellenic Conference on Informatics*, pp80, 2016, ACM
[34] Pavlina Fragkou, Eleni Galiotou, and Michalis Matsakas, "Enriching the e-GIF Ontology for an Improved Application of Linking Data Technologies to Greek Open Government Data", *Procedia - Social and Behavioral Sciences*, vol. 147, pp167-174, 2014, Elsevier
[35] Li Baiyang, and Huang Ruhua, "A Study on the Approaches of Value Realization of Open Government Data", 2016
[36] Patricia Carolina Neves Azevedo, Vitor Afonso Pinto, Guilherme Sousa Bastos, and Fernando Silva Parreiras, "Using Linked Open Data in Geographical Information Systems", pp152-166, 2016, Springer
[37] Dmitrij Koznov, Olga Andreeva, Uolevi Nikula, Andrey Maglyas, Dmitry Muromtsev, and Irina Radchenko, "A Survey of Open Government Data in Russian Federation", *arXiv preprint arXiv:1612.05164*, 2016
[38] Sergio Consoli, Valentina Presutti, Diego Reforgiato Recupero, Andrea G Nuzzolese, Silvio Peroni, Aldo Gangemi, "Producing Linked Data for Smart Cities: The Case of Catania", *Big Data Research*, vol.7, pp1-15, 2017, Elsevier
[39] Sweety Agrawal, Jayati Deshmukh, Srinath Srinivasa, Chinmay Jog, KS Bhavaani, and Rahul Dhek, " A Survey of Indian Open Data", In *Proceedings of the 5th IBM Collaborative Academia Research Exchange Workshop*, pp2, 2013, ACM
[40] Leonida N Mutuku, and Jessica Colaco, "Increasing Kenyan Open Data Consumption: A Design Thinking Approach", In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, pp18-21, 2012, ACM
[41] Julia Hoxha, and Armand Brahaj, "open.data.al - Increasing the Utilization of Government Data in Albania", *Proceedings of the 7th International Conference on Semantic Systems*, pp237-240, 2011
[42] Karin Breitman, Percy Salas, Daniel Saraiva, Vinícius Gama, Marco A Casanova, Jose Viterbo, Ednylton Franzosi, Regis Pires, and Miriam Chaves, "Open Government Data in Brazil", *IEEE Intelligent Systems*, vol.27, no.3, pp45-49, 2012, Institute of Electrical and Electronics Engineers, Inc., USA United States
[43] Alia O Bahanshal, and Hend S Al-Khalifa, "Toward Recipes for Arabic Dbpedia", In *Proceedings of International Conference on Information Integration and Web-based Applications & Services*, pp331, 2013, ACM
[44] Michalis N Vafopoulos, Marios Meimaris, Agis Papantoniou, Ioannis Anagnostopoulos, Giorgos Alexiou, Ioannis Avraam, Ioannis Xidias, Giorgos Vafeiadis, and Vasilis Loumos, "Public spending: Interconnecting and Visualizing Greek Public Expenditure Following Linked Open Data Directives", *Available at SSRN 2064517*, 2012
[45] Stamatios Theocharis, and George A Tsihrintzis, "Ontology Development to Support the Open Public Data-the Greek Case", In *The 5th International Conference on Information, Intelligence, Systems and Applications, IISA*, pp385-390, 2014, IEEE
[46] Eleni Galiotou, and Pavlina Fragkou, "Applying Linked Data Technologies to Greek Open Government Data: A Case Study", *Procedia - Social and Behavioral Sciences*, vol.73, pp479-486, 2013, Elsevier
[47] Ricardo Matheus, Manuella Maia Ribeiro, and Jose Carlos Vaz, "New Perspectives for Electronic Government in Brazil: the adoption of Open

Government Data in National and Subnational Governments of Brazil", In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, pp22-29, 2012, ACM
[48] James Hendler, Jeanne Holm, Chris Musialek, and George Thomas, "US Government Linked Open Data: Semantic.data.gov", *IEEE Intelligent Systems*, vol.27, no.3, pp25-31, 2012, IEEE Computer Society
[49] Lihua Zhao, and Ryutaro Ichise, "Ontology Integration for Linked Data", *Journal on Data Semantics*, vol.3, no.4, pp237-254, 2014, Springer
[50] Hend S. Al-Khalifa, "A Lightweight Approach to Semantify Saudi Open Government Data", In *2013 16th International Conference on Network-Based Information Systems*, pp594-596, 2013, IEEE Computer Society
[51] Tung Mou Yang, Jin Lo, Hui Ju Wang, and Jing Shiang, "Open Data Development and value-added Government Information: Case Studies of Taiwan E-government", In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance - ICEGOV*, pp238-241, New York, USA, 2013, ACM Press
[52] Uldis Bojars, and Renars Liepicnvs, "The State of Open Data in Latvia: 2014", *arXiv preprint arXiv:1406.5052*, 2014
[53] Anatoliy Gruzd, and Jeffrey Roy, "Investigating Political Polarization on Twitter: A Canadian Perspective", *Policy & Internet*, vol.6, no.1, pp28-45, 2014, Wiley Online Library
[54] Andreiwid Sheffer Correa, Pedro Luiz Pizzigatti Correa, and Flavio Soares Correa da Silva, "Transparency Portals Versus Open Government Data: An Assessment of Openness in Brazilian Municipalities", In *Proceedings of the 15th Annual International Conference on Digital Government Research*, pp178-185, 2014, ACM
[55] OpenK̃nowledge International. "The Comprehensive Knowledge Archive Network", https://ckan.org/,2017, [Online; accessed 20-Feb-2017]
[56] Geoff Walsham, "Knowledge Management: The Benefits and Limitations of Computer Systems", *European Management Journal*, vol.19, no.6, pp599-608, 2001, Elsevier
[57] Anneke Zuiderwijk, and Marijn Janssen, "Open Data Policies, their Implementation and Impact: A Framework for Comparison", *Government Information Quarterly*, vol.31, no.1, pp17-29, 2014, Elsevier
[58] John C Bertot, Paul T Jaeger, and Justin M Grimes, "Using ICTs to Create a Culture of Transparency: E-government and Social Media as Openness and Anti-Corruption Tools for Societies", *Government Information Quarterly*, vol.27, no.3, pp264-271, 2010, Elsevier
[59] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk, "Benefits, Adoption Barriers and Myths of Open Data and Open Government", *Information Systems Management*, vol.29, no.4, pp258-268, 2012, Taylor & Francis
[60] Anneke Zuiderwijk, Marijn Janssen, Sunil Choenni, Ronald Meijer, and Roexsana Sheikh Alibaks, "Socio-Technical Impediments of Open Data", *Electronic Journal of E-Government*, vol.10, no.2, pp156-172, 2012
[61] Stuti Saxena, and Irfan Muhammad. "The Impact of Open Government Data on Accountability and Transparency", *Journal of Economic and Administrative Sciences*, vol.34, no.3, pp204-216, 2018, Emerald Publishing Limited, doi:10.1108/JEAS-05-2017-0044
[62] Evangelos Kalampokis, Efthimios Tambouris, and Konstantinos Tarabanis, "Linked Open Government Data Analytics", *International Conference on Electronic Government*, pp99-110, 2013, Springer
[63] Jan Kucera, Dusan Chlapek, Jakub Klimek, and Martin Necasky, "Methodologies and Best Practices for Open Data Publication", In *CEUR Workshop Proceedings*, vol.1343, pp52-64, 2015
[64] Monica Palmirani, Michele Martoni, and Dino Girardi, "Open Government Data beyond Transparency", In *International Conference on Electronic Government and the Information Systems Perspective*, vol.8650, pp275-291, 2014, Springer
[65] Sharon S Dawes, and Natalie Helbig, "Information Strategies for Open Government: Challenges and Prospects for Deriving Public value from Government Transparency", In *International Conference on Electronic Government*, pp50-60, 2010, Springer
[66] Josune Arcelus, "Framework for Useful Transparency Websites for Citizens", In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, pp83-86, 2012, ACM
[67] Anneke Zuiderwijk, and Marijn Janssen, "Participation and Data Quality in Open Data Use: Open data Infrastructures Evaluated", In *Proceedings of the 15th European Conference on E-Government*, pp351-359, 2015
[68] Tung Mou Yang, and Yi-Jung Wu, "Examining the Socio-technical Determinants Influencing Government Agencies' Open Data Publication: A Study in Taiwan", *Government Information Quarterly*, vol.33, no.3, pp378-392, 2016, Elsevier
[69] Mockus Martynas, and Palmirani Monica, "Legal Ontology for Open Government Data Mashups", *Conference for E-Democracy and Open Government (CeDEM)*, pp113-124, 2017, IEEE
[70] Shanshan Jiang, Thomas F Hagelien, Marit Natvig, and Jingyue Li, "Ontology-based Semantic Search for Open Government Data", *IEEE 13th International Conference on Semantic Computing (ICSC)*, pp7-15, 2019, IEEE
[71] Pilar Escobar, Maria del Mar Roldan-Garcia, Jesus Peral, Gustavo Candela, and Jose Garcia-Nieto, "An Ontology-Based Framework for Publishing and Exploiting Linked Open Data: A Use Case on Water Resources Management", *Applied Sciences*, vol.10, no.3, pp779, Multidisciplinary Digital Publishing Institute, 2020
[72] D.Beckett, T.Berners-Lee, " Turtle-terse rdf triple language-w3c team submission", 2020 https://www.w3.org/TeamSubmission/2011/SUBM-turtle-20110328/, [Online; accessed 20-May-2021]