Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

LINKED OPEN SPATIAL DATA FOR EVALUATION OF LAND SUITABILITY

CHARITAS FIBRIANI[1], AHMAD ASHARI[*2], MARDHANI RIASETIAWAN[3]

[1]Universitas Kristen Satya Wacana, Indonesia

[2,3]Department of Computer Science and Electronics Universitas Gadjah Mada, Indonesia

E-mail: [1]charitas.fibriani@mail.ugm.ac.id, [*2]ashari@ugm.ac.id, [3]mardhani@ugm.ac.id

ABSTRACT

The agricultural land evaluation analysis is obtained from several factors, such as: Nutrients, Erosion hazard, temperature, flood hazard, and root media, in which the data are from various sources. The principle of Linked Open Data (LOD) is data can be accessed by anyone and from anywhere. The spatial data are needed for the evaluation of land suitability are connected using the longitude and latitude, information will be obtained related to soil type, andesite material, texture, relief, landform, slope, rock condition in the soil, drainage conditions, soil depth, water absorption on the soil surface and soil consistency at that point which comes from 7 different data sources. The contribution of this research is to identify the vocabulary data used for land suitability that came from several sources and connecting the Uniform Resource Name (URI) is owned by these sources using LOD in semantic web.

Keywords: *Linked Open Data, Spatial Data, Information Intelligent, Semantic Web, Land Suitability*

1. INTRODUCTION

An evaluation of land suitability is necessary to get maximum results in harvesting. The type of data used in land evaluation is spatial data. The agricultural land evaluation analysis is obtained from several factors, such as: (1) Nutrients; (2) Erosion hazard, (3) temperature, (4) flood hazard, (5) root media. The spatial data must be put together in one source before being used for the land suitability analysis process. The analysis is carried out to find alternative solutions to increase agricultural production [1].

The development of the technology has created various data types, in which the data is from various sources [2]. Tim Berners Lee has presented the development of five-star ratings on open data technologies [3]. The characteristic of open data is free to access, modify, and share to others for any purpose [4]. Even though it is characterized as open data, in fact the data on the open data portal is not connected so that there haven't been any connection between the data. To solve the problem, we need analysis by utilizing a combination of several data [5]. Therefore, data on open data portals can be interconnected. Then, it can be used to answer various questions as a decision support. Linked Open Data (LOD) technology is one of concept on semantic web, it is define that anyone can use any data [6], because LOD creates connections between data from various sources even though the data has different business processes [7]. The principle of linked data is to be able to bridge data related to spatial data used for the evaluation of land suitability stored at various sources. Therefore, with the interrelationship of the data, the data can be utilized without having to store it first in one source. The condition for linking data sources in LOD concept is the data has Uniform Resource Name (URI). URI is a character to identify the location of the data source.

The problem of this research is how to use information contains in the spatial data for these different sources. Usually, the use of spatial for the analysis process requires the unification of these data stored in one source. However the spatial data in this research is taken from various sources without the need to store them into one source first to conduct analysis for evaluation of land suitability.

The aim of this study is to link data sources using their URI characters. The novelty in this study is to structure the relationship between spatial data from several sources that can be used in the process of evaluating land suitability.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. LITERATURE REVIEW

There has been some research which discussed linked design between dataset, such as research conducted by [8], [9], [10], [11] and [12]. The research focused on linked design between dataset used and the research hasn't covered the area of architecture model for the linked open data. Meanwhile [13], [14], and [15] has discussed architectural models for LOD. The data level standard used in the research is at the level standard (*) to level standard (***) and the type of data used is the same, means that the used data hasn't been on a stage of open data. Whereas the data of this research already are at level (****). It means that the data souces can be accessed openly and already has Uniform Resource Name (URI) that allows it to be linked to other sources.

Research conducted by [16] produced applications that provide information based on proximity to the user's location. This means, in this study, spatial data has been utilized as a strength. This information is obtained by utilizing user's geographical coordinates of and then connecting with data that has been stored from various sources and has been interconnected using datasets from DBpedia. Whereas this research using datasets from several data sources.

The research carried out by [17] produced a linked open data design that connects several sources related to drug names, indications and contraindications of a drug. The concept used is Natural Language Programming by detecting the similarity of names. Meanwhile, research by [18], [19] and [20] focused on Resource Description Framework (RDF) vocabulary design in which this stage is the initiating stage to be used at future LOD design. Even though the data used has the same type of data at level (*) to level (***). While this research focuses on utilizing URI in each data sources to make bridge between that can be use for land suitability analysis.

Based on previous research, this research will focus on connecting data from different sources using LOD concept. In addition, URI on the data sources are used as an identifier to create a bridge between data sources.

Based on previous research related to linked open data, there are several notes, namely that the research is still using level standard (*) to level (***) and an analysis of the existing dataset has not been carried out because it only collecting data. This journal will discuss the relationship between datasets that have a standard data level (****) that reside in various sources and have different data types [21].

The development of open data technology makes data published openly on the internet [4]. However, to obtain relevant information from these data, it requires more effort. A semantic layer is added to the internet in order for machines to understand web documents, while other function is to get this data to be linked or interconnected [22]. When these data can be interconnected, we will get deeper information [23] such as name, description, location and so on. The concept of linked data refers to a way to link data structures on the internet so that users can integrate all existing data and information and mine data from various sources.

The difficulties in connecting data are caused by differences in kind of data, data types and data formats of each. The solution is to make the machine smarter. There are 2 approaches to make machine smarter in processing. An artificial intelligence approach is to make machine understand the structure of the web and the semantic web approach is to make the data easy for machines to find [24]. On Figure 1, we can see semantic web activities [25].

There are rules for uploading data to the internet so that the data can be easily found by machines, such as:

a. Use Uniform Resource Identifier (URI) as the identity of the object, where this URI must be unique because it can be used for the description of a concept. HTTP URI describes real-world entities and their relationships [7].

b. Connecting data including metadata with the Resource Description Framework (RDF), connecting each data from various sources with a connected graph from these data. It is important to use a standard format when uploading data to the internet. When uploading liked open data on the internet, the Resource Description Framework (RDF) is used because RDF provides a data model that is adapted to the web architecture. The most common RDF formats used to upload linked data are RDF/XML and RDF [7].

c. Using an ontology for the data used, it can be from an existing ontology or create a new ontology.

d. Consider aspects of rules, queries, transformations, and deployments.

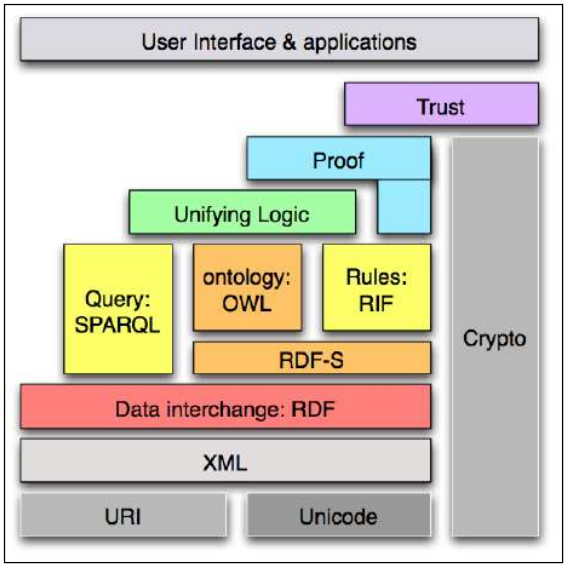

Figure 1: Semantic Web Activity [25]

3. RESEARCH METHODOLGY

The research stages in this paper consist of several stages which can be seen in Figure 2. The first stage is to formulate a research problem. Based on the description of the background, it can be seen that the spatial data needed for land suitability comes from different sources. The problem is how to connect these data sources so the information that needed can be used without having to collect the data in one storage location.

The second stage is to identify the data sources to be used. In addition to identifying data sources, at this stage identification of spatial data needed for land suitability analysis is also carried out.

The third stage is to create a relationship between the data sources used. Where relations use the URI owned by each data source.

The next stage is the process of reading spatial data information on each data source. The information collected is variable information used for land suitability analysis.

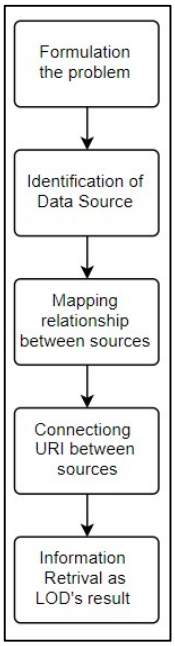

Figure 2: Research Stage

4. DATA SOURCES OF EVALUATION OF LAND SUITABLITY

Land evaluation is a follow-up activity, where there are 3 activities that must be conducted previously, such as: (1) Land Use Type, (2) Land Characteristics, (3) Land Quality. The data for these activities are obtained from various sources. The data used comes from the Yogyakarta Agricultural Technology Research Center (BPTP), where these activities data are found in 7 different sources, namely: Nutrient, Erosion Hazard, Temperature, Land Preparation, Flood Hazard, Rooting Media and Region. The explanation of the data sources used for land evaluation can be seen in Table 1.

Table 1: Land Evaluation Data Sources

No	Data Sources	Note
1	Nutrient	Information related to data on the type of soil at a location, including the material in it and the texture of the soil.
2	Erosion Hazards	Information related to the level of erosion of a location, which contains relief data of a location including the landform of the area
3	Temperature	Information related to the temperature of a location which contains slope data of a location

4	Land Preparation	Information related to the condition of rocks in a soil
5	Flood Hazard	Information related to the flood hazard level of a location, contains drainage data and soil depth.
6	Rooting Media	Information related to roots water absorption level in a location
7	Region	Area map of a location

Nutrients have soil data, andesite material data and texture data. Erosion Hazard has texture data, relief data and landscape data. Temperature has relief data, minimum slope data and maximum slope data. Land Preparation has andesite material data, and rock data. Flood Hazard has drainage data, soil depth data and surface data. Rooting media has surface data, consistency data, and texture data. While the Region has data on the longitude, latitude and administrative information of a region. Because the data used is spatial, the data from these sources also has longitude and latitude. The availability of data on each source can be seen in Table 2.

Table 2: Availability of Information Data Sources

	Region	Nutrient	Erosion Hazards	Temp
Longitude	√	√	√	√
Latitude	√	√	√	√
Soil		√		
Andesite		√		
Texture		√	√	
Relief			√	√
Landform			√	
Slope Min				√
Max Tilt				√
Stone				
Drainage				
Deep				
Surface				
Consistency				

	Field Preparation	Flood Hazard	Rooting Media
Longitude	√	√	√
Latitude	√	√	√
Soil			
Andesite	√		
Texture			√
Relief			
Landform			
Slope Min			
Max Tilt			
Stone	√		
Drainage		√	
Deep		√	
Surface		√	√
Consistency			√

4. LINKED OPEN DATA FOR LAND SUITABILITY

The linked open model of land suitability for rice crops at a location in this research used 14 datasets located in 7 different sources. Table 2 can be seen the availability of data on each source, where the same data is stored on several different sources. The similarity of data stored in several sources is what creates a relationship that can be used as a link between these data. The linked mapping of these data sources can be seen in Figure 3.

The interrelationship of data will create relationships between the data. Interconnected data can be used as an evaluation of agricultural land. The rules created under Figure 3 are:

(1) A data set that has coordinate (longitude and latitude) of an area will be connected with data from other sources with the same coordinate location.

(2) Data sets that have similar texture attributes will have similar landform values.

(3) Data sets that have similar andesite material attributes will have the same value for the amount of rock at ground level.

(4) Data sets that have similar relief attributes will have similar slope values.

(5) Data sets that have similar surface attributes will have similar drainage values.

(6) Data sets that have similar texture attributes will have similar consistency values

Where based on the rules formed, a link is created related to the suitability of the land. Relationships formed based on linked rules can be seen in Table 3.

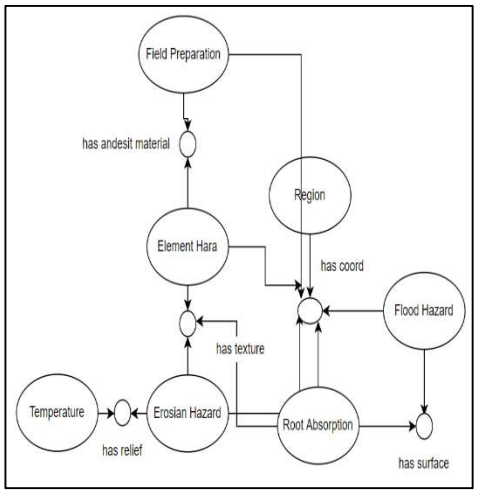

Figure 3: Linked Data Connected Architecture for Land Suitability

Table 3: Characteristic Relationship by Rules

Source	Destination	Relation	Rule
Region	Land Preparation	has a coordinate	1
Region	Nutrients	has a coordinate	1
Region	Erosion Hazards	has a coordinate	1
Region	Temperature	has a coordinate	1
Region	Flood Hazard	has a coordinate	1
Region	Root Absorption	has a coordinate	1
Nutrients	Erosion Hazards	has a texture	2
Erosion Hazards	Flood Hazard	has a texture	6
Temperature	Erosion Hazards	has relief	4
Flood Hazard	Root Absorption	has a surface	5
Land Preparation	Nutrients	has andesite material	3

Spatial data on the source Region is the parent data that stores the coordinates of the location, where the latitude and longitude data in this data are used as the origin node for other sources in obtaining the coordinate location of a region. The process of initializing spatial data used to introduce the coordinate position of a region can be seen in Figure 4.

The process of retrieving data on nutrient sources using links that have been connected to Regional Elements can be seen in Figure 5. Figure 6 is the process of retrieving data on the Danger of Erosion using links that have been connected. Figure 7 is the process of retrieving data on Land Preparation using a link that has been connected. Figure 8 is the process of retrieving data at Temperature using a link that has been connected. Figure 9 is the process of retrieving data on Flood Hazard using links that have been linked. Figure 10 is the process of retrieving data on Flood Hazard using links that have been connected.

```
Program Initialization_Area
begin
        num  long,lat
        char a,b,c
{
        "type": "Feature",
        "geometry": {
        "type": "Polygon",
        "coordinates": [long,lat]
        }
        "properties": {
            "id": "a",
            "kode": "b",
            "nama": "c"
        }
}
```

Figure 4: Pseudocode of Initialization of Area Coordinate

```
Program Linked Nutrient_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_nutrient".b)
        if (c !=null) then
         {
                value[0]=['kode_kec']
                value[1]=['nama_kec']
                value[2]=['tanah']
                value[3]=['bahan_in']
                value[4]=['tekstur']
         }
        else {
                return
             }
}
```

Figure 5: Pseudocode of Linked of Nutrient with Region

```
Program Linked Erotion_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_erotion".b)
        if (c !=null) then
         {
                value[0]=['relief']
                value[1]=['landform']
         }
        else {
                return
             }
}
```

Figure 6: Pseudocode of Linked Erosion Hazard with Region

```
Program Linked Field_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_field".b)
        if (c !=null) then
         {
                value[0]=['batu']
         }
        else {
                return
             }
}
```

Figure 7: Pseudocode of Linked Land Preparation with Region

```
Program Linked Temp_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_temp".b)
        if (c !=null) then
         {
                value[0]=['lereng_minimal']
                value[1]=['lereng_maksimal']
         }
        else {
                return
             }
}
```

Figure 8: Pseudocode of Linked Temperature with Region

```
Program Linked Flood_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_flood.b)
        if (c !=null) then
         {
                value[0]=['drainase']
                value[1]=['kedalaman']
                value[2]=['permukaan']
         }
        else {
                return
             }
}
```

Figure 9: Pseudocode of Linked Flood Hazard with Region

```
Program Linked Root_Area
begin
        num  x
        char a,b,c
{
        x = [long,lat]
        a = ("url_address_region".x)
        b = get(a)

        c = ("url_address_root.b)
        if (c !=null) then
         {
                value[0]=['konsistensi']
         }
        else {
                return
             }
}
```

Figure 10: Pseudocode of Linked Media Rooting with Region

The process of linking between these data sources will produce complete data that can be used in the process of analyzing the suitability of agricultural land. The data consists of a combination of 7 data sources that have 14 variables, namely: soil, material, texture, relief, landform, slopemin, slopemax, rock, drainage, soil depth, surface, consistency, id and region.

This linked process will generate geometry data that contains map information and attribute data that contains information related to the value of the field owned. In Figure 11, you can see the map of the linked process, the areas used are Margodadi, Wotgaleh, Terbah, Tlogo, Serut, and Glinggang. In addition, attributes of the data will be explain using 4 tables because the limits of space. Table 4 is a table created from the linked process, consisting of ID,

region, and soil. Table 5 is a table created from the linked process, consisting of Andesite material, texture, and relief. Table 6 is a tabel created from the linked process, consisting of landform, minimal slope, maximal slope, and stone. While Table 7 is a table created from the linked process, consisting of drainage, depth, surface, and consistency. The data type of the attribute is character according to the data type that comes from the original source. In addition, all of data type are character.

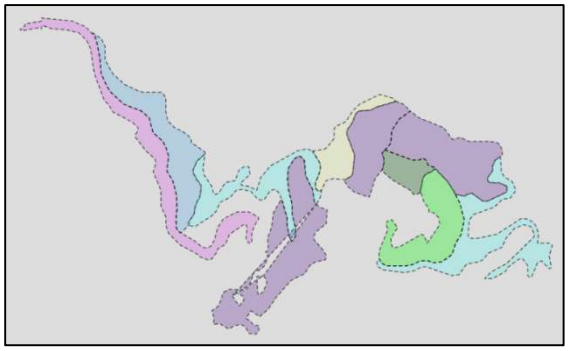

Figure 11: Map Result of Linked Data for Land Suitability

Table 4: Result of Table of Linked Open Data for Land Suitability

ID	Region	Soil
(character)	(character)	(character)
87	MARGODADI	Typic Tropaquepts
23	WOTGALEH	Lytic ustropepts
56	TERBAH	Typic Ustropepts
56	TERBAH	Typic Ustropepts
90	TLOGO	Lytic ustropepts
13	SHAVED	Typic Hapluderts
56	TERBAH	Typic Ustropepts
85	GLINGGANG	Lithic Ustorthents

Table 5: Result of Table of Linked Open Data for Land Suitability (2)

Andesite Material	Texture	Relief
(character)	(character)	(character)
Clay and sand deposits	Clay	A bit flat
Tufa dasit	Clay	Hilly
Batulanau	Clay	Steep
Batulanau	Clay	A bit steep
Batulanau	Dusty clay	Steep
Clay and sand deposits	Clay	A bit flat
Sandstone and siltstone complex	Clay	Hilly
Complex of brection and tufa dacite	Clay	Rugged

Table 6: Result of Table of Linked Open Data for Land Suitability (3)

Landform	Min Slope	Max Slope	Stone
(character)	(character)	(character)	(character)
Volcanic plain	1	3	zero or little
Monoklinal Hill	15	25	medium or many
The slopes of the force mountains	25	45	medium or many
The slopes of the force mountains	15	25	zero or little
The slopes of the force mountains	25	45	medium or many
Alluvial plain	1	3	zero or little
Folding hills	15	25	zero or little
Eskarpment	61	90	medium or many

Table 7: Result of Table of Linked Open Data for Land Suitability (4)

Drainage	Depth	Surface	Consistency
(character)	(character)	(character)	(character)
2	1	Slow	Loose
2	3	Keep	firm
1	3	Keep	firm
1	3	Keep	firm
1	2	Keep	firm
2	3	A bit slow	firm
1	3	Keep	firm
1	3	A bit fast	Loose

5. CONCLUTION

Connected data related to land evaluation allows external and internal parties to obtain the data needed to analyze the land suitability of a location. Obtaining the longitude location and latitude coordinates of a point, information will be obtained related to soil type, Andesite material, texture, relief, landform, slope, rock condition in the soil, drainage conditions, soil depth, water absorption on the soil surface and soil consistency at that point which comes from 7 different data sources. The concept of linked open data is used to design rules that are used as a reference to create relationships between these different sources. The stages are: (1) Recording the data sources to be used, (2) Recording the attributes

owned by each source, (3) Recording the relationships between the attributes in each source, (4) Designing links between the sources used as a link between these sources, (5) Data that has been connected to each other can be utilized

Connected data related to land evaluation allows external and internal parties to obtain the data needed to analyze the land suitability of a location. Obtaining the longitude location and latitude coordinates of a point, information will be obtained related to soil type, andesite material, texture, relief, landform, slope, rock condition in the soil, drainage conditions, soil depth, water absorption on the soil surface and soil consistency at that point which comes from 7 different data sources. The concept of linked open data is used to design rules that are used as a reference to create relationships between these different sources. The stages are: (1) Recording the data sources to be used, (2) Recording the attributes owned by each source, (3) Recording the relationships between the attributes in each source, (4) Designing links between the sources used as a link between these sources, (5) Data that has been connected to each other can be utilized.

Future research could include rules that are used as a reference for analysis for land suitability. Using these rules can provide information related to the suitability of a land for certain crops.

REFERENCES:

[1] A. Nada, "Service Oriented Approach for Decision Support Systems," *IEEE*, 2014.

[2] P. C. C. Choi, K. L. W. Cho, and Y. Lee, "KMLOD : linked open data service for Korean medical database," *J Supercomput*, vol. 76, pp. 7758–7776, 2020.

[3] H. Linden and J. Strale, "An Evaluation of Platforms for Open Government Data," *KTH (Royal Inst. Technol. Sch. Technol. Heal. 136 40 Handen, Sweden*, 2015.

[4] J. Debattista, J. Attard, and D. O. Sullivan, "Is the LOD cloud at risk of becoming a museum for datasets ? Looking ahead towards a fully collaborative and sustainable LOD cloud," pp. 850–858, 2019.

[5] A. Alrefae and J. Cao, "Intensional XML-enabled web-based real-time decision support system," in *Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017*, 2017, vol. 2017-Janua, pp. 1–10.

[6] P. Hitzler and K. Janowicz, "Linked Data , Big Data , and the 4th Paradigm," vol. 4, pp. 233–235, 2013.

[7] C. Bizer, *Linked Data*. 2011.

[8] A. Ashari, "Cloud-Based Processing for Data Science Visualization," vol. 7, no. 2, pp. 449–454, 2016.

[9] D. Deng, *Construction and Reuse of Linked Agriculture Data : An Experience of Taiwan Government Open Data*, vol. 1. Springer International Publishing, 2018.

[10] J. Huuskonen and T. Oksanen, "Soil sampling with drones and augmented reality in precision agriculture," *Comput. Electron. Agric.*, vol. 154, pp. 25–35, 2018.

[11] E. Folmer, W. Beek, L. Rietveld, T. L. Registry, and T. L. Registry, "LINKED DATA VIEWING AS PART OF THE SPATIAL DATA," vol. XLII, no. August, pp. 29–31, 2018.

[12] M. Lnenicka and J. Komarkova, "International Journal of Information Management Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing," *Int. J. Inf. Manage.*, vol. 46, no. June 2018, pp. 124–141, 2019.

[13] F. Cifuentes-silva, C. Sifaqui, B. Congreso, and J. E. Labra-gayo, "Towards an architecture and adoption process for Linked Data technologies in Open Government contexts A case study for the Library of Congress of Chile," pp. 79–86, 2011.

[14] C. Daniel, F. Flores, J. Luis, S. Cervantes, and G. A. Hernández, "Architecture for the integration of Linked Open Drug Data in an Augmented Reality application for mobile devices," *Springer Int. Publ. AG 2018 J. Mejia al. (eds.), Trends Appl. Softw. Eng. Adv. Intell. Syst. Comput.*, vol. 688, 2018.

[15] M. Mayumi, M. Yagui, L. Fernando, M. Passos, J. Oliveira, and A. S. Vivacqua, "Curation of physical objects in botany : Architecture and development of a linked open data-based application," pp. 889–892, 2017.

[16] C. Becker and B. Christian, "Web Semantics : Science , Services and Agents on the World Wide Web Exploring the Geospatial Semantic Web with DBpedia Mobile," vol. 7, pp. 278–286, 2009.

[17] J. Kozák, M. Neˇ, and D. Jan, "Linked Open Data for Healthcare Professionals," 2013.

[18] E. A. C. Costantini, "Exposing Vocabularies for Soil as Linked Open Data," *Inf. Process. Agric.*, 2015.

[19] C. Tian-en, C. Li-ping, G. Yunbin, and W. Yanji, "Spatial Decision Support System for Precision Farming Based on GIS Web Service," in *2009 International Forum on Information Technology and Applications*, 2009.

[20] T. Wang, "Aligning the large-scale ontologies on schema-level for weaving Chinese linked open data," *Cluster Comput.*, 2018.

[21] M. Victorino, M. T. De Holanda, E. Ishikawa, E. C. Oliveira, and S. Chhetri, "Transforming Open Data to Linked Open Data Using Ontologies for Information Organization in Big Data Environments of the Brazilian Government : the Brazilian Database Government Open Linked Data – DBgoldbr," *Knowl. Organ.*, vol. 45, no. 6, pp. 443–467, 2018.

[22] O. Debauche, S. Mahmoudi, A. L. H. Andriamandroso, P. Manneback, J. Bindelle, and F. Lebeau, "Cloud services integration for farm animals' behavior studies based on smartphones as activity sensors," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–12, 2018.

[23] T. Heath, M. Hepp, and C. Bizer, "Special Issue on Linked Data," *Int. J. Semant. Web Inf. Syst.*, vol. 5, 2009.

[24] M. Poblet, J. Suquet, A. Roig, and J. González-conejero, "Building Semantic Interoperability for European Civil Proceedings Online for the European Small Claims Procedure (ESCP) and European Order for Payment Procedure (EPO)," 2013.

[25] S. Bratt, "Semantic Web , and Other W3C Technologies to Watch," vol. 2007, no. January, 2007.