A Computational Model of Place on the Linked Data Web

Alia I Abdelmoty, Khalid O. Al-Muzaini

School of Computer Science & Informatics, Cardiff University, Wales, UK
Email: {A.I.Abdelmoty, Almuzainiko}@cs.cf.ac.uk

***Abstract*—The Linked Data Web (LDW) is an evolution of the traditional Web from a global information space of linked documents to one where both documents and data are linked. A significant amount of geographic information about places is currently being published on this LDW. These are used to qualify the location of other types of datasets. This paper examines the limitations in the nature of location representation in some typical examples of these Resource Description Framework (RDF) resources, primarily resulting from the simplified geometric representation of location and the incomplete and random use of spatial relationships to link place information. The paper proposes a qualitative model of place location that enforces an ordered representation of relative spatial relationships between places. This work further explores how semantic properties of place can be included to derive meaningful location expressions. The model facilitates the application of qualitative spatial reasoning on places to extract a potentially large percentage of implicit links between place resources, thus allowing place information to be linked and to be explored more fully and more consistently than what is currently possible. The paper describes the model and presents experimental results demonstrating the effectiveness of the model on realistic examples of geospatial RDF resources.**

Keywords–qualitative place models; spatial reasoning; geospatial web.

I. INTRODUCTION

One of the 'Linked Data Principles' is to include links to connect the data to allow the discovery of related things. However, identifying links between data items remains a considerable challenge that needs to be addressed [1], [2], [3]. A key research task in this respect is identity resolution, i.e., to recognise when two things denoted by two URIs are the same and when they are not. Automatic linking can easily create inadequate links, and manual linking is often too time consuming [4]. Geo-referencing data on the LDW can address this problem [5], whereby links can be inferred between data items by tracing their spatial (and temporal) footprints. For example, the BBC uses RDF place gazetteers as an anchor to relate information on weather, travel and local news [6].

Yet, for geospatial linked data to serve its purpose, links within and amongst the geographic RDF resources need themselves to be resolved. That is to allow place resources to be uniquely identified and thus a place description in one dataset can be matched to another describing the same place in a different dataset. A scheme that allows such links between place resources to be discovered would be a valuable step towards the realisation of the LDW as a whole.

In this paper, location is used as a key identifier for place resources and the question to be addressed is how location can be used to define a *linked* place model that is sufficient to enable place resources to be uniquely identified on the LDW. Several challenges need to be addressed, namely, 1) location representation of RDF place resources is simple; defined as point coordinates in some resources, detailed; defined with extended geometries in others, and sometimes missing all together, 2) coordinates of locations may not match exactly across data sources, where volunteered data mapped by individuals is mashed up with authoritative map datasets, 3) non-standardised vocabularies for expressing relative location is used in most datasets, e.g., in DBpedia, properties such as *dbp:location*, *dbp-ont:region* and *dbp-ont:principalarea* are used to indicate that the subject place lies inside the object place.

Towards addressing this problem, a linked place model is proposed that uses qualitative spatial relationships to describe unique place location profiles, as presented in [1]. The profiles don't rely on the provision of exact geometries and hence can be used homogeneously with different types of place resources. They can be expressed as RDF statements and can thus be integrated directly with the resource descriptions. The rationale behind the choice of links to be modelled is primarily twofold: to allow for a sensible unique description of place location and to support qualitative spatial reasoning over place resources.

The model is further adapted to consider semantic aspects of place location definition. In particular, the notion of salience of place is used to scope the type of relationships used in the location expressions in the defined place profiles. It is shown how the proposed representation scheme is flexible to allow for the encoding of relevant location expressions, whilst also retaining the power of spatial reasoning within the framework proposed.

The value of the linked place model is illustrated by measuring its ability to make the underlying RDF graph of geographic place resources browsable. Samples of realistic geographic linked datasets are used in the experiments presented and results demonstrate significant potential value of the methods proposed.

The paper is structured as follows. An overview of related work on the representation and manipulation of place resources on the LDW is given in section II, In section III the proposed relative location model, as well its adaptation to include semantic aspects of place definition, are described. In section IV, application of the models proposed is evaluated on two different realistic datasets. Conclusions and an overview of future work is given in section VI.

II. RELATED WORK

Here related work on the topics of representing place resources and reasoning with them on the LDW are reviewed.

2016, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

A. Representing RDF place Resources on the LDW

Sources of geographic data on the LDW are either volunteered (crowdsourced) resources, henceforth denoted Volunteered Geographic Information (VGI), created by individuals with only informal procedures for validating the content, or authoritative resources produced by mapping organizations, henceforth denoted Authoritative Geographic Information (AGI). Example of VGIs are DBpedia (dbpedia.org), GeoNames (geonames.org), and OpenSreetMaps (linkedgeodata.org) [7] and examples of AGIs are the Ordnance Survey linked data [8] and the Spanish linked data [9]. Data collected from users on the Social Web, e.g., on Twitter and Foursquare, can also be considered as VGIs [10].

The volume of VGI resources is increasing steadily, providing a wealth of information on geographic places and creating detailed maps of the world. DBpedia contains hundreds of thousands of place entities, whose locations are represented as point geometry. GeoNames is a gazetteer that collects both spatial and thematic information for various place names around the world. In both datasets, place location is represented by a single point coordinates. While DBpedia does not enforce any constraints on the definition of place location (e.g., coordinates may be missing in place resources), reference to some relative spatial relationships, and in particular to represent containment within a geographic region, is normally maintained. A detailed analysis of the spatial data content of DBpedia can be found in [11], [12]. GeoNames places are also interlinked with each other by defining associated parent places.

In [13], the LinkedGeoData effort is described where OpenStreetMap (OSM) data is transformed into RDF and made available on the Web. OSM data is represented with a relatively simple data model that captures the underlying geometry of the features. It comprises three basic types, nodes (representing points on Earth and have longitude and latitude values), ways (ordered sequences of nodes that form a polyline or a polygon) and relations (groupings of multiple nodes and/or ways). Furthermore, [14] presented methods to determine links between map features in OSM and equivalent instances documented in DBpedia, as well as between OSM and Geonames. Their matching is based on a combination of the Jaro-Winkler string distance between the text of the respective place names and the geographic distance between the entities. Example of other work on linking geodata on the Semantic Web is [15], which employs the Hausdorff distance to establish similarity between spatially extensive linear or polygonal features.

In contrast to VGI resources that manages geographic resource as points (represented by a coordinate of latitude and longitude), AGI resources deal with more complex geometries as well, such as line strings. AGIs tend to utilise well-defined standards and ontologies for representing geographic features and geometries. Ordnance Survey linked data also demonstrates the use of qualitative spatial relations to describe spatial relationships in its datasets. Two ontologies, the Geometry Ontology and the Spatial Relations Ontology, are used to provide geospatial vocabulary. These ontologies describe abstract geometries and topological relations (equivalent to RCC8 [16]) respectively.

In summary, the spatial representation of place resources in VGI datasets is generally limited to point representation, and is managed within simple ontologies that encode non-spatial semantics and in some cases limited spatial relationships. On the other hand, place data provided as AGI tend to present more structured and detailed spatial representations, but is also limited to specific types and scales of representation. Use of some qualitative spatial relationships has been demonstrated for capturing the spatial structure in some example datasets. The model proposed in this paper offers a systematic and homogenous representation of place location that can be consistently applied to VGIs or AGIs and demonstrates the value of heterogenous qualitative spatial relations in representing place information on the LDW.

B. Manipulating and Querying RDF place resources on the LDW

Recently, much work has been done on extending RDF for representing geospatial information through defining and utilising appropriate vocabularies encoded in ontologies to represent space and time. The work capitalises on specification of standards, defined by the Open Geospatial Consortium (OGC)(opengeospatial.org), for modeling core concepts related to geospatial data. Prominent examples are the geographic query language for RDF (GeoSPARQL), an OGC standard [17] and stRDF/stSPARQL (st stands for spatiotemporal) [18]. Both proposals provide vocabulary (classes, properties, and functions) that can be used in RDF graphs and SPARQL queries to represent and query geospatial data, for example *geo:SpatialObject*, which has as instances everything that can have a spatial representation and *geo:Geometry* as the superclass of all geometry classes. In addition, geometric functions and topological functions are offered for performing computations, such as *geof:distance* and for asserting topological relations between spatial objects, e.g., *dbpedia:Cardiff geo:sfWithin dbpedia:Wales*.

Qualitative spatial representation and reasoning (QSRR) are established areas of research [19], [20], whose results have influenced the definition of models of spatial relationships in international standards, e.g., the OGC models, and commercial spatial database systems (for example, in the Oracle DB system). $RCC8$, a QSRR model, has been recently adopted by GeoSPARQL [17], and there is an ever increasing interest in coupling QSR techniques with Linked Geospatial Data that are constantly being made available [18]. On the other hand, Semantic Web reasoning engines have been extended to support qualitative spatial relations, e.g., Racerpro [21] and PelletSpatial [22]. Scalability of the spatial reasoning is recognised and reported challenge. Scalable implementations of constraint network algorithms for qualitative and quantitative spatial constraints are needed, as RDF stores supporting Linked Geospatial Data are expected to scale to billions of triples [18]. Lately, promising results have been reported by [23], who proposed an approach for removing redundancy in RCC8 networks and by [24], who examined graph-partitioning techniques as a method for coping with large networks; in both cases leading to more effective application of spatial reasoning mechanisms. Finally, qualitative methods were used to complement existing quantitative methods for representing the geometry of spatial locations. In [25], heterogenous reasoning methods are proposed that combine calls between a spatial database system and a spatial reasoning engine implemented in *OWL2 RL* to check the consistency of place ontologies.

In [26], Younis et al described query plans that make use of a combination of qualitative spatial relationships associated with place resources in DBpedia and detailed representations of geometry maintained in a spatially indexed database for answering complex queries. In both cases, qualitative reasoning was limited by the fragmented and scarce availability of spatial relationships to work on. The qualitative scheme of representation of place location proposed in this paper addresses this issue and provides a novel method for defining spatial relationships that is designed to support and facilitate the effective use of qualitative spatial reasoning on the LDW.

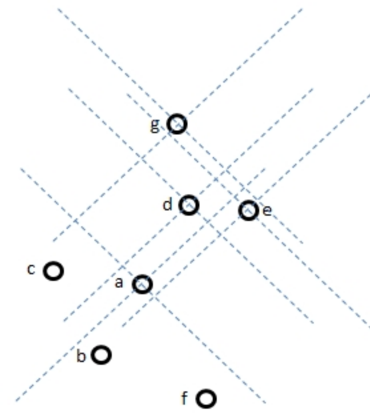
III. A Linked Place Model for the Linked Data Web

A Relative Location model (*RelLoc*) is proposed here to capture a qualitative representation of the spatial structure of place location. Two types of spatial relations are used as follows.

1) Containment relationships, to record that a parent place directly contains a child place; i.e., one step hierarchy. For example, for three places representing a district, a city and a country, the model will explicitly record the relationships: inside(district, city) and inside(city, country), but not inside(district, country).
2) Direction-proximity relationships, to record for every place the relative direction location of its nearest neighbour places. The direction frame of reference can be selected as appropriate. For example, for a 4-cardinal direction frame of reference, a place will record its relative direction relation with its nearest neighbour in four directions.

For a given set of places Pl, let $DirPr$ be the set of all direction-proximity relations between instances of places in Pl as defined above, and let Con be the set of containment relations between instances of places in Pl as defined above. Then, $RelLoc(Pl)$ is defined as a tuple $RelLoc(Pl) := (Pl, D, C)$, where: $D \in DirPr$ and $C \in Con$. $R_{nn}(x, y)$ is used to denote that x is the nearest neighbour from the direction R to object y. For example, $N_{nn}(pl_1, pl_2)$ indicates that pl_1 is the nearest neighbour from the north direction to pl_2, etc.

To illustrate the model, consider the scene in Figure 1 that consists of a set of places, a to f, with a 4-cardinal direction frame of reference overlaid for some places in the scenes. A representative point is used to define the place location. It is further known that places represented as points a, b, c, e are inside d and places d, f are inside g. The full set of relationships used to model the scene are given in the table in Figure 1(b). Note that in some cases, no relation can be found, e.g., there are no neighbours for object c from the west direction in Figure 1(a).



(a)

Set of spatial relations to model relative location
$N_{nn}(d,a)$, $S_{nn}(b,a)$, $W_{nn}(c,a)$, $E_{nn}(e,a)$
$N_{nn}(g,d)$, $S_{nn}(a,d)$, $W_{nn}(c,d)$, $E_{nn}(e,d)$
$N_{nn}(g,c)$, $S_{nn}(b,c)$, $E_{nn}(a,c)$
$N_{nn}(a,b)$, $E_{nn}(f,b)$
$N_{nn}(a,f)$, $W_{nn}(b,f)$
$N_{nn}(g,e)$, $S_{nn}(b,e)$, $W_{nn}(d,e)$
$S_{nn}(d,g)$
$in(a,d)$, $in(b,d)$, $in(c,d)$, $in(e,d)$,
$in(d,g)$, $in(f,g)$

(b)

Figure 1. (a) An example map scene with a set of places represented as points.(b) Set of direction, proximity and containment relations chosen to representative relative location in the proposed model.

TABLE I. COMPOSITION TABLE FOR 4-CARDINAL DIRECTION RELATIONSHIPS.

	N	E	S	W
N	N	$N \vee E$	All	$N \vee W$
E	$N \vee E$	E	$S \vee E$	All
S	All	$S \vee E$	S	$S \vee W$
W	$W \vee N$	All	$W \vee S$	W

A. Spatial Reasoning with the Relative Location Model

We can reason over the relative location model to infer more of the implicit spatial structure of place location. Qualitative spatial reasoning (QSR) tools can be utilised to propagate the defined relationships and derive new ones between places in the scene. QSR takes advantage of the transitive nature of the partial or total ordering of the quantity space in order to infer new information from the raw information presented. In particular, the transitive nature of some spatial relationships can be used to directly infer spatial hierarchies, for example, containment and cardinal direction relations. The scope of the model is deliberately focussed on general containment relationships and ignores other possible topological relations, such as overlap or touch. Hence, building containment hierarchies is straightforward using the transitivity rules: $inside(a, b) \wedge inside(b, c) \rightarrow inside(a, c)$ and $contains(a, b) \wedge contains(b, c) \rightarrow contains(a, c)$.

In the case of direction relationships, more detailed spatial reasoning can be applied using composition tables. Table I shows the composition table for a 4-cardinal direction frame of reference between point representations of spatial objects. In considering the entries of the composition tables, some of those entries provide definite conclusions of the composition operation, i.e., the composition result is only one relationship (emboldened in the table), other entries are indefinite and result in a disjunctive set of possible relationships, e.g., the composition: $N(a, b) \wedge E(b, c) \rightarrow N(a, c) \vee E(a, c)$.

Spatial reasoning can be applied on the linked place model

TABLE II. RESULT OF REASONING WITH CARDINAL RELATIONS FOR THE PLACE MODEL IN FIGURE 1.

	a	b	c	d	e	f	g
a	-	**N**	**E**	**S**	W	**N**	S
b	**S**	-	**S**	S	**S**	**W**	S
c	**W**	W	-	**W**	W	$N \vee W$	$S \vee W$
d	**N**	N	E	-	**W**	N	**S**
e	**E**	N	E	**E**	-	N	S
f	S	**E**	$S \vee E$	S	S	-	S
g	N	N	**N**	**N**	**N**	N	-

using different strategies. The most straightforward is through deriving the algebraic closure, i.e., completing the scene by deriving all possible missing relationships between objects. Path-consistency algorithms for deriving the algebraic closure has been been implemented in various tools, e.g., in the SparQ spatial reasoning engine [27]. Table II shows the result of this operation for the example scene in Figure 1. Explicit relations are shown in bold and the remaining relation are inferred by spatial reasoning. As can be seen in the table, using the 19 relationships defined for the model in Figure 1(b), reasoning was able to derive a further 19 definite relationships, completing over 90% of the possible relations in the scene.

B. Applying the Relative Location Place Model on the LDW

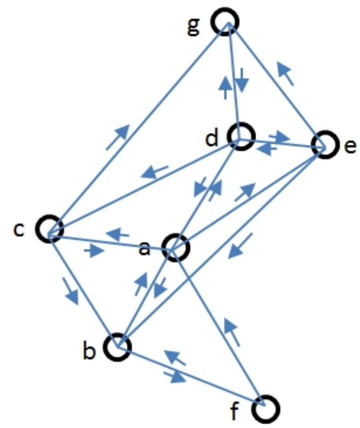
The underlying structure of any expression in RDF is a collection of triples, each consisting of a subject, a predicate and an object. A set of such triples is called an RDF graph, in which each triple is represented as a node-arc-node link and each triple represents a statement of a relationship between the subjects and objects, denoted by the nodes, that it links. The meaning of an RDF graph is the conjunction (logical AND) of the statements corresponding to all the triples it contains.

The *RelLoc* place model can be interpreted as a simple connected graph with nodes representing place resources and edges representing the spatial relationships between places. Thus a realisation of the place model for a specific RDF document of place resources is a subgraph of the RDF graph of the document. The *RelLoc* RDF graph is completely defined if RDF statements are used to represent all spatial relationships defined in the model, e.g., for the scene in Figure 1, 25 RDF statements are needed to encode the cardinal (19) and containment (6) relationships in the table in Figure 1(b).

Let Pl be a finite set of place class resources defined in an RDF data store and $DirPr(Pl)$ defines cardinal direction relations between members of Pl and $Con(Pl)$ describes the containment relations between members of Pl as defined by the relative location model above.

A *RelLoc* subgraph $\mathbb{G}_{\mathbb{L}} = (V_{\mathbb{L}}, E_{\mathbb{L}})$ is a simple connected graph that models Pl, where: $V_{\mathbb{L}} = Pl$ is the set of nodes, $E_{\mathbb{L}} = \{DirPr(Pl) \cup Con(Pl)\}$ is the set of edges labelled with the corresponding direction and containment relationships.

Note that there exists a subgraph of $\mathbb{G}_{\mathbb{L}}$ for every place $pl \in Pl$, which represents the subset of direction-proximity and containment relationships that completely define the relative location of pl. Thus, a *location profile* for a particular place $pl \in Pl$ can be defined as $\mathbb{L}_{pl} = \{(DirPr_{pl}, Con_{pl}\}$. $\mathbb{L}_{pl}$ is the restriction of $\mathbb{L}$ to pl, where $DirPr_{pl}$ and Con_{pl} defines direction proximity and containment relations respectively between pl and other places in Pl, as specified by our model.



(a)

	a	b	c	d	e	f	g
a	-	1	1	1	1	0	0
b	1	-	0	0	0	1	0
c	1	1	-	0	0	0	1
d	1	0	1	-	1	0	1
e	0	1	0	1	-	0	1
f	1	1	0	0	0	-	0
g	0	0	0	1	0	0	-

(b)

	a	b	c	d	e	f	g
a	-	N	E	S	0	N	0
b	S	-	S	0	S	W	0
c	W	0	-	W	0	0	0
d	N	0	0	-	W	0	N
e	E	0	0	E	-	0	0
f	0	E	0	0	0	-	0
g	0	0	N	N	N	0	-

(c)

Figure 2. (a) A graph representing the sample map scene from Figure 1. (b) Adjacency matrix for the location graph representing nearest neighbour relationships. (c) Adjacency-orientation matrix representing nearest neighbour and direction relationships.

For example the location profile for place a in Figure 1 is the set of statements describing the relations: $N(d, a), S(b, a), W(c, a), E(e, a), in(a, d)$.

The *RelLoc* graph can be represented by a matrix to register the adjacency relationship between the place and its nearest neighbours. The scene in Figure 1 is shown as a graph with nodes and edges in Figure 2(a) and its corresponding adjacency matrix is shown in (b). The fact that two places are neighbours is represented by a value (1) in the matrix and by a value (0) otherwise. Values of (1) in the matrix can be replaced by the relative orientation relationship between the corresponding places as shown in Figure 2(c) and the resulting structure is denoted *Adjacency-Orientation Matrix*.

C. A Semantic Place Model

So far, the *RelLoc* place model considers distance and direction relationships as the primary factors for specifying place location. The importance of a place or its *salience* is another factor that is useful to consider. Salience of a place can be described from a personal or from an absolute point of view.

On a personal level, many factors can influence the importance of place to an individual [28], including, a) place dependence; how far the place satisfies the individuals behavioural goals as compared to other alternatives (e.g., [29], [30]), b) place affect; reflecting the emotional or affective bond between an individual and a place (e.g., [30], [31]), and c) place social bonding; reflecting the importance of social relationships and the context within which they occur. The specific settings of the place share the meanings attributed to them by the individuals social environment (e.g., [32], [30]).

On an absolute level, salience of a place can be defined irrelevant of attachment to specific individuals. For example, Hall, Smart and Jones [33] considered salience as a factor in defining the place location when devising methods for automatic caption generation for images (or photographs). In their work, the equation that determines the set of relative places to choose from in a particular image caption is a combination of an equal number of "ways" (highways, roads, paths, $\cdots$) and other places, ordered by their relative salience. A salience value is, in turn, a measure of how close the location of a place is to the image (i.e., its distance from the image), and its popularity (i.e., how well-known the place is). The later factor can be derived automatically from the Web, for example, from the counts of place mentions on Flickr, Wikipedia and web pages [34].

Our basic *RelLoc* place model can be adapted to handle different possible semantics of place, such as, place type, activities carried out in a place or place salience. The adapted model will, henceforth, be denoted Semantic Relative Location model, or *SemRelLoc*.

Hence, in *SemRelLoc*, a layer of salient places is first extracted from the base map layer and this acts as the anchor for place location definition. Thus, the algorithm for defining the relative location model is applied between, a) all places on the salient feature layer, and b) every place in the remaining set of places in the base map layer and the salient place layer only.

Consider the example schematic maps in Figure 3. In Figure 3(a), places on the map are not distinguished by any specific property and relationships between them are defined using *RelLoc*. In (b), a salient place layer is filtered out and used as a basis for the *SemRelLoc* model. The selection of places in this layer can be chosen to serve the application in context, for example, as a selection of particular place types, or specific place instances with high popularity, or even those of relevance to a particular individual. Note that in (b) spatial relationships are defined only with reference to the salient place instances and no relationships are defined amongst the remaining places on the base map layer, as will be described below.

Let $SalientPl$ be a finite set of place resources defined as a subset of all places Pl in an RDF data store, and $\overline{SalientPl}$ be the rest of places remaining on the base layer (i.e., $\{Pl \backslash SalientPl\}$.

A semantic relative location *SemRelLoc* subgraph $\mathbb{G}_{\mathbb{L}} = (V_{\mathbb{L}}, E_{\mathbb{L}})$ is a simple connected graph that models Pl, where: $V_{\mathbb{L}} = SalientPl$ is the set of nodes, $E_{\mathbb{L}} =$

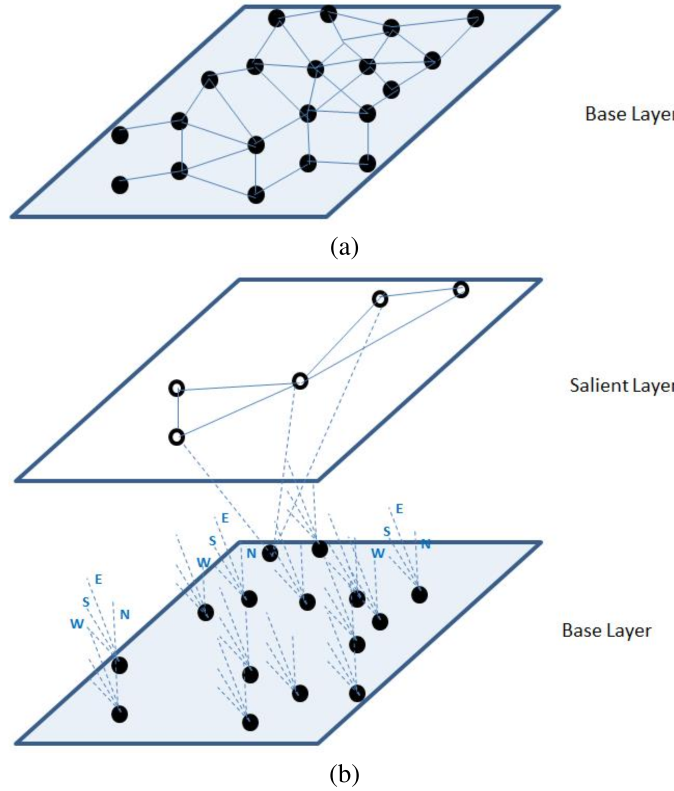


Figure 3. (a) Schematic of sample base map layer. (b) Salient place layer filtered out.

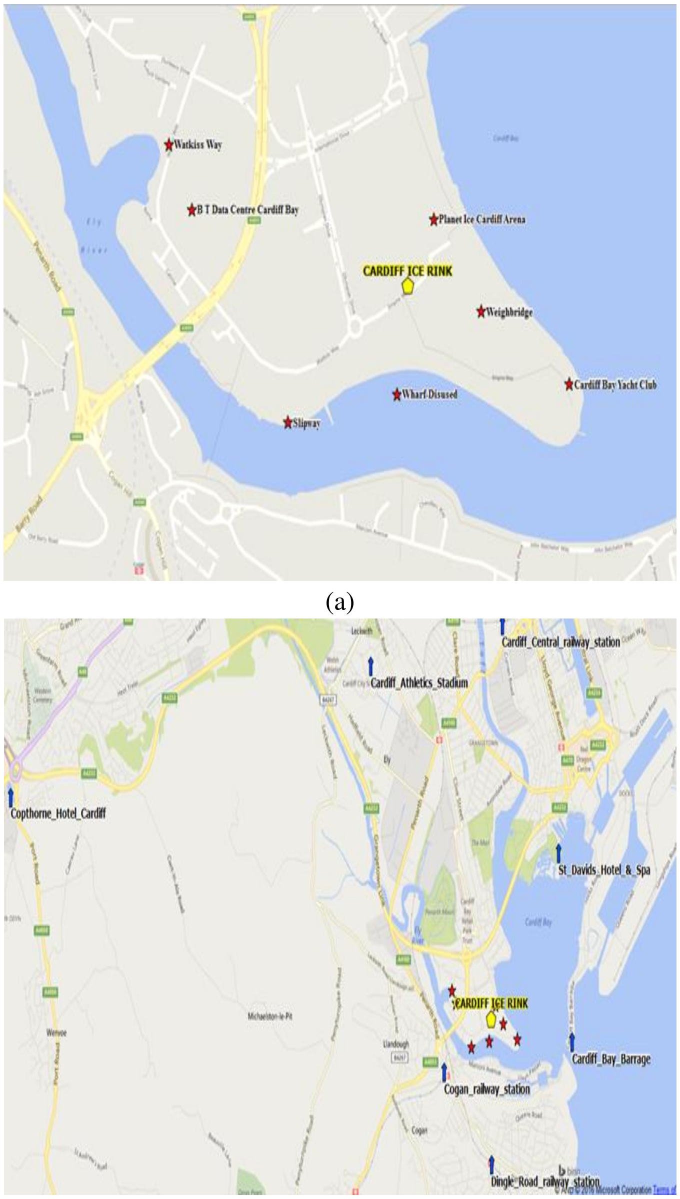
$\{DirPr(SalientPl) \cup DirPr(\overline{SalientPl}) \cup Con(Pl)\}$ is the set of edges labelled with the corresponding direction and containment relationships. $\{DirPr(SalientPl)$ is the set of direction-proximity relationships between places on the salient feature layer. $DirPr(\overline{SalientPl})$ is the set of direction-proximity relationships between the rest of places on the base layer with places on the salient layer. Hence, no inter-relationships are defined between places on the base layer itself.

Note that there exists a subgraph of $\mathbb{G}_{\mathbb{L}}$ for every place $pl \in Pl$, which represents the subset of direction-proximity and containment relationships that completely define the relative location of pl. Thus, a *Semantic location profile* for a particular place $pl \in Pl$ can be defined as follows.

$$\mathbb{L}_{pl} = \left\{ \begin{array}{ll} \{DirPr_{Salientpl}, Con_{pl}\} & ,\text{if } pl \in SalientPl \\ \{DirPr_{\overline{SalientPl}}, Con_{pl}\} & ,\forall pl \in \overline{SalientPl} \end{array} \right\}$$

Figure 4(a) shows a section of the Cardiff Bay area in Cardiff, Wales. A set of places are shown around the place: 'Cardiff Ice Rink'. Figure 4(a) shows the set of places chosen to describe the location with the original *RelLoc* model, while in 4(b) the set of some selected salient features (hotels, museums, railway stations, etc.) around the place are shown. These are used to describe the location with *SemRelLoc*. Table III lists the set of location expressions defined by both models. While both are topologically correct, the location expressions of the *SemRelLoc* model can be considered more meaningful and useful for general contexts.

2016, © Copyright by authors, Published under agreement with IARIA - www.iaria.org



(a)

(b)

Figure 4. Sample map scene with places defining the location of "Cardiff Ice Rink": a) with *RelLoc* model, and b) with *SemRelLoc*.

$SemRelLoc$ offers two potential advantages over $RelLoc$; a) more meaningful place location expressions, using selected relevant place instances, and b) potentially a more economical data model to manage and reason with. The number of pre-defined relationships remains constant, as every place will have a set of statements defining its proximity and direction relationships. However, spatial reasoning with the semantic location graph can be more efficient with the reduction of the variety of modelled edges between places. In the following section, the effectiveness of spatial reasoning with $SemRelLoc$ will be compared against the basic $RelLoc$ model.

IV. Application and Evaluation

The main goals of the Linked Place model is to provide a representation of place location on the LDW that allows for place information to be linked effectively and consistently. The effectiveness of the proposed model can be evaluated with respect to two main aspects; whether it provides a sound definition of place location, that is to test the correctness of the place location profiles, and whether it provides a complete definition of place location, that is whether a *complete* relative location graph can be derived using the individual place location profiles.

The soundness of the location profiles is assumed as it essentially relies on the validity of the computation of the spatial relationships. Issues related to the complexity of this process are discussed in the next section.

Here, we evaluate the completeness aspect of the model. An individual place location profile defined using the model represents a finite set of spatial relationships between a place and its nearest neighbours and direct parent. Completeness of the model can be defined as the degree to which these individual profiles can be used to derive implicit links between places not defined by the model. The model is entirely complete if a full set of links between places can be derived using automatic spatial reasoning, i.e., the model can produce a complete graph, where there is a defined spatial relationship between every place in the dataset and every other place.

A system was developed that implements the Linked Place model and further builds an enriched model using spatial reasoning for evaluation purposes as shown in Figure 5.

TABLE III. Location expressions defining the place "Cardiff Ice Rink" in both the *RelLoc* and *SemRelLoc* models

RelLoc Model
Wharf Disused **N** Cardiff Ice Rink
Slipway **NE** Cardiff Ice Rink
BT Data Centre Cardiff Bay **E** Cardiff Ice Rink
Watkiss Way **SE** Cardiff Ice Rink
Planet Ice Cardiff Arena **SW** Cardiff Ice Rink
Weighbridge **W** Cardiff Ice Rink
Cardiff Bay Yacht Club **NW** Cardiff Ice Rink
SemRelLoc Model
Dingle Road railway station **N** Cardiff Ice Rink
Cogan railway station **NE** Cardiff Ice Rink
Copthorne Hotel Cardiff **E** Cardiff Ice Rink
Cardiff Athletics Stadium **SE** Cardiff Ice Rink
Cardiff Central railway station **S** Cardiff Ice Rink
St Davids Hotel and Spa **SW** Cardiff Ice Rink
Cardiff Bay Barrage **W** Cardiff Ice Rink

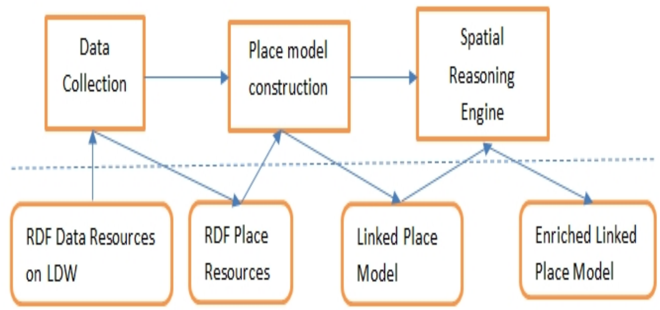


Figure 5. Components of the developed system to implement the linked place model.

```
prefix d: <http://dbpedia.org/ontology/>
prefix :<http://dbpedia.org/resource/>
prefix prop: <http://dbpedia.org/property/>
prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>

select ?place  (MAX(?lat) as ?lat)(MAX(?long) as ?long)
where{
?place ?ontology ?resource.
?place a d:Place.
?place geo:lat ?lat.
?place geo:long ?long.
filter ( ?resource = :Wales or ?resource = "Wales"@en )
}
group by ?place
order by ?place
```

Figure 6. SparQL query used to extract place data from DBpedia.

TABLE IV. RESULTS OF REASONING APPLIED ON THE DBPEDIA DATASET.

Defined	Definite	2-Relations	3-Relations	Others
2751	50340	63148	28	136
2.36%	43.24%	54.22%	0.02%	0.12%

A. Evaluation of the Relative Location Place Model

Two datasets were used in this experiment, DBPedia and the Ordnance Survey open data [8]. These were chosen as they exhibit different representations of place resources on the LDW and are typical of VGIs and AGIs respectively. A description of the datasets used is presented below, along with the results of the application of spatial reasoning over the constructed linked place models.

DBpedia DataSet

A sample dataset containing all Places in Wales, UK, has been downloaded from DBpedia using the SPARQL query in Figure 6.

A total of 489 places were used, for which a relative location graph of 2751 direction-proximity relations was constructed. Completing the graph resulted in 116403 total number of relations, out of which 50340 relations are definite (defining only one possible relationship).

Note that of the indefinite relationships some are a disjunction of 2 relations, e.g., $\{N, NW\}$ or $\{E, SE\}$ and some are a disjunction of 3 relations, e.g., $\{N, NE, NW\}$ or $\{NE, E, SE\}$. In both cases, relations can be generalised to a "coarser" direction relation, for example, $\{NE, E, SE\}$ can be generalised to general $East$ relationship. These results are considered useful and thus are filtered out in the presentation. The remaining results are disjunctions of unrelated directions, e.g., $\{N, NE, E\}$, and are thus considered to be ambiguous. A summary of the results is shown in Table IV. Using the Linked Place model we are able to describe nearly half the possible relations precisely (45.6%), as well as almost all of the rest of the scene (54.22%) with some useful generalised direction relations.

Ordnance Survey DataSet

The Boundary-line RDF dataset for Wales was downloaded from the Ordnance Survey open data web site [8]. The data gives a range of local government administrative and electoral boundaries.

Figure 7 shows the relative location graph constructed for the Unitary Authority dataset for Wales. Dashed edges are

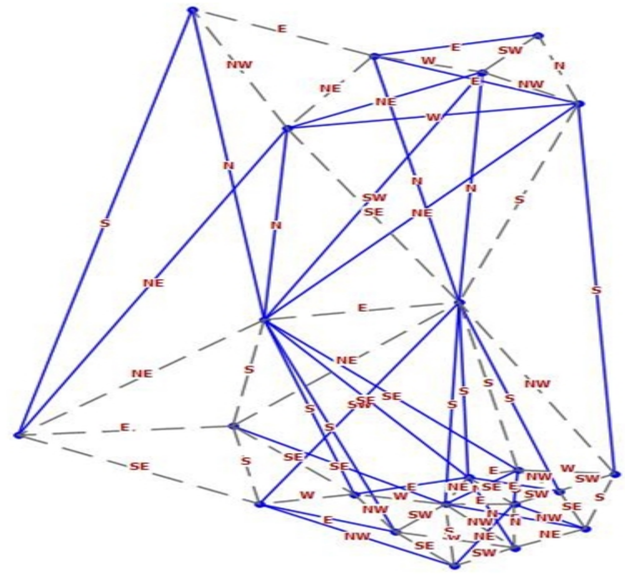


Figure 7. (a) Linked Place Graph for the Unitary Authorities in Wales from the Ordnance Survey dataset.

TABLE V. RESULTS OF REASONING APPLIED ON THE ORDNANCE SURVEY DATASET.

Defined	Definite	2-Relations	3-Relations	Others
73	94	64	0	0
31.6%	40.69%	27.7%	0	0

used to indicate that relationships (and inverses) are defined both ways between the respective nodes, but only one relation is used to label the edge in the Linked Place model. The set contains 22 regions, for which 73 direction-proximity relations were computed. Reasoning applied on this set of relations produces the results shown in Table V.

We can use the above results to describe the effectiveness of the linked place model in terms of the information content it was able to deduce using the ratio of the number of defined relations to the number of deduced relations. A summary is presented in Table VI.

B. Evaluation of the Semantic Place Model

The value of the $SemRelLoc$ model is primarily in its ability to deliver flexible and meaningful place location expressions. Here, we also evaluate its effectiveness with respect to spatial reasoning. An experiment is carried out with a sample point of interest dataset obtained from the Ordnance Survey, that records information on places and place types in the city of Cardiff, Wales, UK. A set of approximately 300 places were chosen in 5 unitary authorities in South Wales; (Cardiff, Newport, Caerphilly, Vale of Glamorgan and

TABLE VI. SUMMARY OF THE EXPERIMENT RESULTS.

	$\frac{Defined}{Definite}$	$\frac{Defined}{Useful}$
$DBpedia$	0.054	0.024
OS	0.78	0.32

2016, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

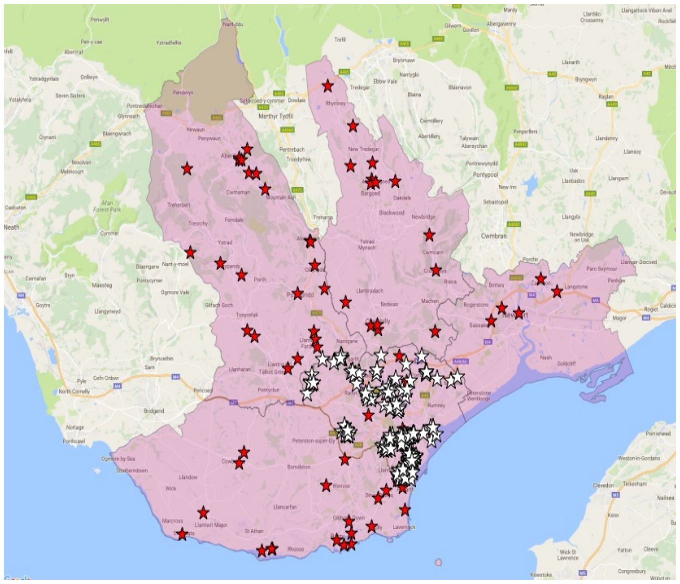


Figure 8. A map scene with a sample set of point of interest places in South Wales, UK. Red/dark stars represent salient features and white stars are all other places.

TABLE VII. RESULTS OF REASONING APPLIED ON THE SALIENT FEATURE LAYER ONLY.

Defined	Definite	2-Relations	3-Relations
538	3261	2515	2
8.52%	51.63%	39.81%	0.031%

Rhondda). Salient features were chosen on the basis of popular place types, including hotels, museums, hospitals, castles and railway stations. A map of the area chosen is shown in Figure 8 with the salient (red/dark) and other places (white) highlighted.

Table VII shows the result of applying spatial reasoning on the complete graph on the salient feature layer only. A total of 108 places were used, for which a relative location graph of 538 direction-proximity relations was defined. Completing the graph resulted in 5778 total number of relations, out of which 3261 (56%) relations are definite (defining only one possible relationship) and a further 2515 (44%) are useful 2-relations. Thus, using $RelLoc$ on the salient feature layer, and defining only 8% of relations, we were able to derive almost all of the scene with useful location expressions.

With the $SemRelLoc$ model, no relations have to be predefined between the base layer places. Every place on the base layer is instead linked to places on the salient feature layer. Thus, for the complete map of places in Figure 8, a further 181 other places were added to the scene and a total of 1331 pre-defined proximity direction relations are defined by the model. Completing the graph resulted in 29403, out of which 12939 (44%) relations are definite and a further 15454 (53%) are useful 2-relations. Thus, using $SemRelLoc$ model, and defining only 5% of possible relationships between places in the map scene, we were able to complete the whole graph and derive over 96% of all possible relationships between all places.

The result demonstrates that the application of spatial

TABLE VIII. RESULTS OF REASONING APPLIED ON THE WHOLE MAP SCENE with $SemRelLoc$.

Defined	Definite	2-Relations	3-Relations or more
1331	12939	15454	1010
4.52%	44%	52.56%	3.44%

reasoning on the adapted semantic model is as effective as in the case of the basic model. Further research can now be directed at the scalability of the framework with respect to both representation and reasoning on the Linked Data Web.

C. Discussion

The proposed approach to place representation and reasoning can be adopted on the LDW in different scenarios as follows.

- The spatial integrity of the linked web resources can be checked. Here, spatial reasoning can be applied on the whole resource to determine which predicates are contradictory [25]. Scalability of the reasoning engine can be an issue and methods for managing large resources need to be considered [18].
- Enriching linked web resources by the basic relative location model allows uniform and complete representation of place across resources, which in turn supports effective retrieval of place information.
- Using the reasoner to build a parallel, complementary resource of the complete set of possible spatial information that can be inferred from the basic relative location model, as was done in the experiments above. The inferred resource will need to be updated regularly to reflect the current state of the original resource and scalability of the reasoning method will also be an issue that needs to be addressed.
- Spatial reasoning can be applied 'on the fly' to support a defined set of query plans on the basic relative location structure. A possible framework to implement this scenario is given in Figure 9. Possible query plans that can be supported by this framework includes finding relationships directly supported by the model (namely, containment, nearest neighbour and direction queries) as well as those derived by spatial inference using spatial reasoning. Detailed specification of these queries are the subject of future research.

V. Conclusions

Data on geographic places are considered to be very useful on the LDW. Individuals and organisations are volunteering data to build global base maps enriched with different types of traditional and non-traditional semantics reflecting people's views of geographic space and place. In addition, geographic references to place can be used to link different types of datasets, thus enhancing the utility of these datasets on the LDW. This work explores the challenges introduced when representing place data using the simple model of RDF, with different geometries to represent location and different non-standardised vocabularies to represent spatial relationships between locations.

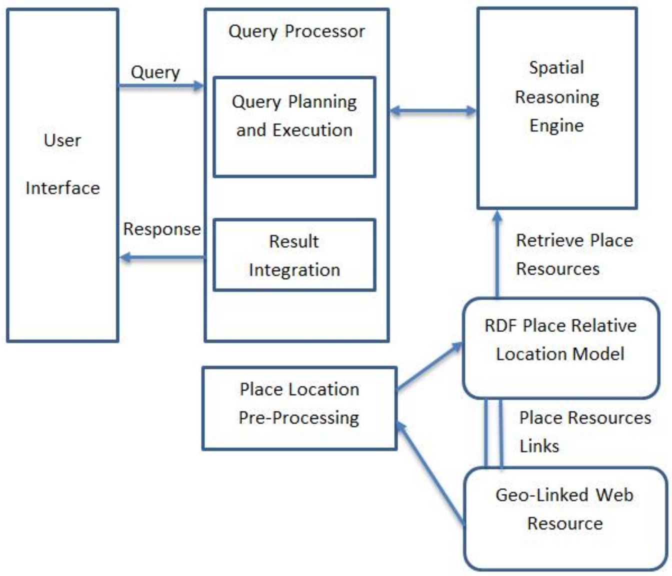


Figure 9. A framework for the implementation of the proposed approach.

A linked place model is proposed that injects certain types of spatial semantics into the RDF graph underlying the place data. Specific types of spatial relationships between place nodes are added to the graph to allow the creation of individual place location profiles that fully describe the relative spatial location of a place. It is further shown how the enriched relative location graph can allow spatial reasoning to be applied to derive implicit spatial links to produce even more richer place descriptions. Salience of place is introduced as a means of scoping out relevant and meaningful place location expressions. The representation scheme is adapted to allow for the flexible choice of place instances to be used in the model.

The results obtained from the evaluation experiments demonstrate possible significant value in the proposed model. Further work need to be done to explore the potential utility of the proposal. Some of the interesting issues that we aim to explore in the future are described below.

- Simple methods and assumptions were used to compute the direction relationships between places. Further study need to be carried out to evaluate whether more involved representations are useful [35].
- Applications of spatial reasoning need to be considered further. Describing the complete graph is not a practical (nor a useful) option. Can spatial reasoning be selectively applied, for example, as part of query processing on the location graph?
- The application of the approach on other types of data sets on the LDW as individual as well as combined resources.

APPENDIX

Follows is a sample set of relationships defining the location of "Techniquest": an educational charity in The Cardiff Bay area, Wales, UK, resulting from the application of spatial reasoning on the salient features in the map in Figure 8.

```
(Techniquest (sw) Bastion_Road)
(Techniquest (sw) Barry_Power_Station)
(Techniquest (sw) Barry_Docks_railway_station)
(Techniquest (sw) Barry_Island_railway_station)
(Techniquest (sw) Barry_Dock_Lifeboat_Station)
(Techniquest (sw w) Cowbridge_railway_station)
(Techniquest (sw w) Llantwit_Major_Roman_Villa)
(Techniquest (sw) Barry_railway_station)
(Techniquest (sw w) Aberthaw_power_stations)
(Techniquest (sw w) Aberthaw_High_Level_railway_station)
(Techniquest (sw w) Aberthaw_Low_Level_railway_station)
(Techniquest (ne) Celtic_Manor_Resort)
(Techniquest (ne) Kingsway_Shopping_Centre)
(Techniquest (ne) Allt-yr-yn)
(Techniquest (ne) Caerleon_railway_station)
(Techniquest (n nw) Caerphilly_Castle)
(Techniquest (n nw) Argoed_railway_station)
(Techniquest (n nw) Brithdir_railway_station)
(Techniquest (n nw) Aberbargoed_Hospital)
(Techniquest (n nw) Aber_Bargoed_railway_station)
(Techniquest (n ne) Crosskeys_railway_station)
(Techniquest (sw w) Atlantic_College_Lifeboat_Station)
(Techniquest (nw) Coed_Ely_railway_station)
(Techniquest (n nw) Bargoed_railway_station)
(Techniquest (n nw) Cefn_Eglwysilan)
(Techniquest (nw) Church_Village)
(Techniquest (nw) Church_Village_Halt_railway_station)
(Techniquest (nw) Sardis_Road)
(Techniquest (nw) Cross_Inn_railway_station)
(Techniquest (nw) Aberdare_Low_Level_railway_station)
(Techniquest (nw) Aberdare_railway_station)
(Techniquest (nw) Coed-Ely)
(Techniquest (n nw) Cilfynydd)
(Techniquest (n nw) Abercynon_railway_station)
(Techniquest (n nw) Abertysswg_railway_station)
(Techniquest (n nw) Darran_and_Deri_railway_station)
(Techniquest (nw) Dinas_Rhondda_railway_station)
(Techniquest (n nw) Abercynon_North_railway_station)
(Techniquest (n nw) Bute_Town)
(Techniquest (nw) Mynydd_William_Meyrick)
(Techniquest (nw) Clydach_Vale)
(Techniquest (nw) Lluest-wen_Reservoir)
(Techniquest (n nw) Abercwmboi_Halt_railway_station)
(Techniquest (n nw) Abernant_railway_station)
(Techniquest (nw) Athletic_Ground_Aberdare)
(Techniquest (nw) Aberaman_railway_station)
(Techniquest (n nw) Cwmbach_railway_station)
(Techniquest (nw) Beddau_Halt_railway_station)
(Techniquest (n) Abercarn_railway_station)
(Techniquest (s) Alberta_Place_Halt_railway_station)
(Techniquest (w) St_Fagans_National_History_Museum)
(Techniquest (sw) Dinas_Powys_railway_station)
(Techniquest (e) Pierhead_Building)
(Techniquest (e) Mermaid_Quay)
(Techniquest (n nw) Caerphilly_railway_station)
(Techniquest (n ne) Ruperra_Castle)
(Techniquest (n nw) Birchgrove_railway_station)
(Techniquest (n nw) National_Museum_Cardiff)
(Techniquest (ne) Bassaleg_Junction_railway_station)
(Techniquest (ne) RAF_Pengam_Moors)
(Techniquest (n) Childrens_Hospital_for_Wales)
(Techniquest (w) Aberthin_Platform_railway_station)
(Techniquest (n nw) Abertridwr_railway_station)
(Techniquest (n) Cefn_Onn_Halt_railway_station)
(Techniquest (nw) Creigiau_railway_station)
(Techniquest (nw) Efail_Isaf_railway_station)
(Techniquest (n nw) Aber_railway_station)
(Techniquest (nw) Castell_Coch)
(Techniquest (nw) Whitchurch_Hospital)
(Techniquest (n nw) Hilton_Cardiff)
(Techniquest (w) St_Fagans_Castle)
(Techniquest (sw) Eastbrook_railway_station)
(Techniquest (w) Cardiff_International_Sports_Stadium)
(Techniquest (s) Cardiff_Bay_Barrage)
(Techniquest (sw w) Dyffryn_Gardens)
(Techniquest (n) University_Hospital_of_Wales)
```

REFERENCES

[1] A. Abdelmoty and K. Al-Muzaini, "Reasoning with place information on the linked data web," in *The Second International Conference on Big Data, Small Data, Linked Data and Open Data- ALLDATA 2016*, V. Gudivada, D. Roman, M. di Buono, and M. Monteleone, Eds., 2016, pp. 48–53.

2016, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[2] T. Berners-Lee. Linked data. http://www.w3.org/DesignIssues/LinkedData.html. Accessed: 2016-11-30.

[3] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data - the story so far," *International Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.

[4] J. Goodwin, C. Dolbear, and G. Hart, "Geographical linked data: The administrative geography of great britain on the semantic web," *Transactions in GIS*, vol. 12, pp. 19–30, 2008.

[5] G. Hart and C. Dolbear, *Linked Data: A Geographic Perspective*. CRC Books, 2013.

[6] C. Henden, P. Rissen, and S. Angeletou. Linked geospatial data, and the bbc. http://www.w3.org/2014/03/lgd/papers/lgd14_submission_28. Accessed: 2016-11-30.

[7] M. Hackley, "How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets," *Environment and Planning B*, vol. 37, no. 4, pp. 682–703, 2010.

[8] Ordnance survey linked data platform. http://data.ordnancesurvey.co.uk/. Accessed: 2016-11-30.

[9] Geolinkeddata.es. http://linkeddata.es/. Accessed: 2016-11-30.

[10] S. Mohamed and A. Abdelmoty, "Uncovering user profiles in location based social networks," in *The Eighth International Conference on Advanced Geographic Information Systems, GEOProcessing 2016*, C.-P. Rckemann and Y. Doytsher, Eds., 2016, p. 1421.

[11] E. Younis, "Hybrid geo-spatial query processing on the semantic web," Ph.D. dissertation, Cardiff University, Cardiff Univeristy, 2013.

[12] E. M. Younis, C. Jones, V. Tanasescu, and A. Abdelmoty, "Hybrid geo-query methods on the semantic web with a spatially enhanced index of dbpedia," in *7th International Conference on Geographic Information Science, GIScience, 2012*, N. Xiao, M.-P. Kwan, M. Goodchild, and S. Shekhar, Eds., vol. LNCS 7478, 2012, pp. 340–353.

[13] S. Auer, J. Lehmann, and S. Hellmann, "Linkedgeodata- adding a spatial dimension to the web of data," in *Proc. of 7th International Semantic Web Conference (ISWC)*, vol. LNCS 5823. Springer, 2009, pp. 731–746.

[14] C. Stadler, J. Lehmann, K. Hffner, and S. Auer, "Linkedgeodata: A core for a web of spatial open data," *Semantic Web Journal*, vol. 3, no. 4, pp. 333–354, 2012.

[15] J. Salas and A. Harth, "Finding spatial equivalences accross multiple rdf datasets," in *Terra Cognita Workshop,co-located with ISWC*, R. Grtter, D. Kolas, M. Koubarakis, and D. Pfoser, Eds., 2011, pp. 114–126.

[16] A. Cohn, B. Bennett, J. Gooday, and N. Gotts, "Qualitative spatial representation and reasoning with the region connection calculus," *Geoinformatica*, vol. 1, pp. 1–44, 1997.

[17] Geosparql - a geographic query language for rdf data. http://www.opengeospatial.org/standards/geosparql. Accessed: 2016-11-30.

[18] M. Koubarakis, M. Karpathiotakis, K. Kyzirakos, C. Nikolaou, and M. Sioutis, "Data Models and Query Languages for Linked Geospatial Data," in *Reasoning Web. Semantic Technologies for Advanced Query Answering,Invited tutorial at the 8th Reasoning Web Summer School 2012 (RW 2012)*, T. Eiter and T. Krennwallner, Eds., vol. LNCS 7487. Springer, 2012, pp. 290–328.

[19] A. G. Cohn and J. Renz, *Qualitative Spatial Representation and Reasoning*, ser. Handbook of Knowledge Representation. Elsevier, 2008, pp. 551–596.

[20] B. Gottfried, "Reasoning about intervals in two dimensions," in *IEEE Int. Conf. on Systems, Man and Cybernetics*, W. e. a. Thissen, Ed., 2004, pp. 5324–5332.

[21] What is racerpro. http://www1.racer-systems.com/products/racerpro/index.phtml. Accessed: 2016-11-30.

[22] M. Stocker and E. Sirin, "Pelletspatial: A hybrid rcc-8 and rdf/owl reasoning and query engine," in *6th Intern. Workshop on OWL: Experiences and Directions (OWLED2009)*. Springer-Verlag, 2009, pp. 2–31.

[23] M. Sioutis, S. Li, and J.-F. Condotta, "On redundancy in linked geospatial data," in *2nd Workshop on Linked Data Quality (LDQ)*, ser. CEUR Workshop Proceedings, A. Rula, A. Zaveri, M. Knuth, and D. Kontokostas, Eds., no. 1376, 2015. [Online]. Available: http://ceur-ws.org/Vol-1376/#paper-05

[24] C. Nikolaou and M. Koubarakis, "Fast consistency checking of very large real-world rcc-8 constraint networks using graph partitioning," in *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 2014, pp. 2724–2730.

[25] A. Abdelmoty, P. Smart, and B. El-Geresy, "Spatial reasoning with place information on the semantic web," *International Journal on Artificial Intelligence Tools*, vol. 23, no. 5, p. 1450011, 2014.

[26] E. Younis, C. Jones, V. Tanasescu, and A. Abdelmoty, "Hybrid geo-spatial query methods on the semantic web with a spatially-enhanced index of dbpedia," in *Geographic Information Science*, ser. Lecture Notes in Computer Science, N. Xiao, M.-P. Kwan, M. Goodchild, and S. Shekhar, Eds., vol. 7478. Springer Berlin Heidelberg, 2012, pp. 340–353.

[27] D. Wolter and J. Wallgrn, "Qualitative spatial reasoning for applications: New challenges and the sparq toolbox," in *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions*, S. Hazarika, Ed., 2012, pp. 336–362.

[28] G. Kyle, A. Graefe, and M. R., "Testing the dimensionality of place attachment in recreational settings," *Environment and Behavior*, vol. 37, no. 2, pp. 153–177, 2005.

[29] G. Brown and C. Raymond, "The relationship between place attachment and landscape values toward mapping place attachment," *Applied Geography*, vol. 27, no. 1, pp. 89–111, 2007.

[30] G. Kyle, K. Bricker, A. Graefe, and T. Wickham, "An examination of recreationists relationships with activities and settings." *Leisure Sciences*, vol. 26, no. 2, pp. 123–142, 2004.

[31] B. Jorgensen and R. Stedman, "Sense of place as an attitude: Lakeshore owners attitudes toward their properties." *Journal of Environmental Psychology*, vol. 21, pp. 233–248, 2001.

[32] H. Gu and C. Ryan, "Place attachment, identity and community impacts of tourism the case of a beijing hutong." *Tourism Management*, vol. 29, no. 4, pp. 637–647, 2008.

[33] M. Hall, P. Smart, and C. Jones, "Interpreting spatial language in image captions," *Cognitive Processing*, vol. 12, no. 1, pp. 67–94, 2011.

[34] M. Hall, "Modelling and reasoning with quantitative representations of vague spatial language used in photographic image captions," Ph.D. dissertation, Cardiff University, Cardiff Univeristy, 2011.

[35] X. Yao and J.-C. Thill, "How far is too far? a statistical approach to context-contingent proximity modeling," *Transactions in GIS*, vol. 9, no. 2, pp. 157–178, 2005.

2016, © Copyright by authors, Published under agreement with IARIA - www.iaria.org