

Article

A Comparative Evaluation of Geospatial Semantic Web Frameworks for Cultural Heritage

Ikrom Nishanbaev ^{1,*} , Erik Champion ^{1,2,3}  and David A. McMeekin ^{4,5}

¹ School of Media, Creative Arts, and Social Inquiry, Curtin University, Perth, WA 6845, Australia; erik.champion@curtin.edu.au

² Honorary Research Professor, CDHR, Sir Roland Wilson Building, 120 McCoy Circuit, Acton 2601, Australia

³ Honorary Research Fellow, School of Social Sciences, FABLE, University of Western Australia, 35 Stirling Highway, Perth, WA 6907, Australia

⁴ School of Earth and Planetary Sciences, Curtin University, Perth, WA 6845, Australia; david.mcmeekin@curtin.edu.au

⁵ School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, WA 6845, Australia

* Correspondence: ikrom.nishanbaev@postgrad.curtin.edu.au

Received: 14 July 2020; Accepted: 4 August 2020; Published: 12 August 2020



Abstract: Recently, many Resource Description Framework (RDF) data generation tools have been developed to convert geospatial and non-geospatial data into RDF data. Furthermore, there are several interlinking frameworks that find semantically equivalent geospatial resources in related RDF data sources. However, many existing Linked Open Data sources are currently sparsely interlinked. Also, many RDF generation and interlinking frameworks require a solid knowledge of Semantic Web and Geospatial Semantic Web concepts to successfully deploy them. This article comparatively evaluates features and functionality of the current state-of-the-art geospatial RDF generation tools and interlinking frameworks. This evaluation is specifically performed for cultural heritage researchers and professionals who have limited expertise in computer programming. Hence, a set of criteria has been defined to facilitate the selection of tools and frameworks. In addition, the article provides a methodology to generate geospatial cultural heritage RDF data and to interlink it with the related RDF data. This methodology uses a CIDOC Conceptual Reference Model (CRM) ontology and interlinks the RDF data with DBpedia. Although this methodology has been developed for cultural heritage researchers and professionals, it may also be used by other domain professionals.

Keywords: the Geospatial Semantic Web; the Semantic Web; cultural heritage; CIDOC CRM

1. Introduction

In recent years, Geographic Information Systems (GIS) have become a popular technology for cultural heritage (CH) researchers and professionals. One reason for this is the enormous possibilities of GIS, such as their ability to capture, manage, analyze, and visualize all forms of spatio-temporal data, including 3D geospatial data. Hence, more and more CH organizations, professionals, and researchers are adopting GIS frameworks and tools [1,2].

On the other hand, GIS are also undergoing marked changes in their technology stack. One example is the Geospatial Semantic Web, which was recently introduced to GIS following the evolution of the Semantic Web. The Semantic Web is a relatively recent breakthrough in web development that provides a set of best practices for publishing and interconnecting structured data on the Web. Most of today's data on the Web are designed for people to read and not for machines to understand and process. The Semantic Web aims to fill this gap by providing a set of best practices to publish data to

the Web in such a way that their meaning is well-defined to machines, and the data are interlinked with other related data [3]. In contrast, the Geospatial Semantic Web is an extension of the Semantic Web in which geospatial data has explicit meaning to machines defined by geospatial ontologies. Geospatial datasets in the Geospatial Semantic Web have links to related external data sets and can also be linked to and from external datasets [4,5].

The CH domain has successfully implemented some Geospatial Semantic Web concepts. For instance, Hiebel, et al. [6] developed a geospatial ontology called CRMgeo that provides a schema to integrate geospatial ontology GeoSPARQL with CIDOC Conceptual Reference Model (CRM), one of the most widely used CH ontologies. CRMgeo offers the necessary classes and relations to model spatio-temporal aspects of CH objects in the CIDOC CRM ontology [6]. Another example is a WarSampo—the Semantic Web project by Hyvönen, et al. [7] that has published Finnish military historical data relating to the Second World War as Linked Open Data (LOD). The project datasets include photographs, memoirs of soldiers, historical maps, and biographies including place names in the form of geolocation, while the CIDOC CRM ontology and the Simple Knowledge Organization System (SKOS) vocabulary data model were used to provide well-defined meaning to the datasets [7]. The SKOS¹ is an application of RDF that can be used to represent concept schemas such as thesauri and classification schemes as an RDF graph.

However, for many domain professionals, including CH, generating geospatial LOD and interlinking it with related RDF data is a challenging task that requires extensive knowledge of computer programming, the Semantic Web, and the Geospatial Semantic Web, among others. Furthermore, many existing LOD sources are sparsely interlinked. According to recent research findings by Schmachtenberg, et al. [8], who analyzed the adoption of LOD best practices, including interlinking, in different domains such as media and life sciences, 44% of published LOD datasets are not linked to any other datasets.

This article provides a comparative evaluation of features and functionality of the current state-of-the-art geospatial RDF data generation tools and interlinking frameworks. The former allows users to convert geospatial data into RDF data, while the latter enables users to interlink generated RDF data with related RDF datasets. This evaluation is specifically performed for CH researchers and professionals who do not have considerable expertise in computer programming. Thus, the paper does not attempt to conduct performance benchmarking of the tools, neither does it consider tools that require computer programming such as RDFLib², ontospy³, or similar. Instead, the geospatial RDF generation tools and interlinking frameworks are selected based on pre-defined criteria, which are described in Section 3.

Secondly, this article presents a methodology demonstrating how the evaluated tools and frameworks can be applied to generate geospatial Linked Data. To demonstrate the applicability of the methodology, it was applied in a sample use case that uses geospatial CH data relating to Western Australian CH places. These data are mapped into the CIDOC CRM ontology, which is then interlinked with the related RDF data from DBpedia. Although this methodology has been developed for CH researchers and professionals, it may also be used by other domain professionals.

Thirdly, the article provides a discussion on some key challenges and limitations that CH researchers and professionals may encounter when using the evaluated tools and frameworks, including this methodology.

The remainder of the paper has the following structure: in Section 2, the paper provides background literature on the Semantic Web, the Geospatial Semantic Web, and some of the successful applications of the Geospatial Semantic Web in the CH domain. Section 3 provides information about the criteria

¹ <https://www.w3.org/TR/swbp-skos-core-spec/> (last accessed on 8 June 2020)

² <https://github.com/RDFLib/rdfliib> (last accessed on 8 June 2020)

³ <https://pypi.org/project/ontospy/> (last accessed on 8 June 2020)

that have been used to select RDF generation tools and interlinking frameworks. Section 4 presents a comparative evaluation of the selected RDF generation tools and interlinking frameworks. Section 5 proposes a methodology based on the evaluated tools and frameworks to demonstrate the workflow for generating linked geospatial RDF data. In Section 6, the article discusses some of the key challenges that CH researchers and professionals may encounter when using the evaluated tools and frameworks. Section 7 presents a concluding summary of the article.

2. Background Literature

2.1. Semantic Web and the Geospatial Semantic Web

In recent years, GIS has become an important technology that has transformed the way CH professionals conduct research and perform applied projects. Since GIS offers enormous possibilities for collection, analysis, management, and visualization of spatio-temporal data, including 3D geospatial data, it has been applied successfully in many CH use cases such as 3D GIS for CH [1,9], GIS for analysis and visualization of CH [10,11], to name just a few.

GIS is also undergoing marked changes in its technology stack. The Geospatial Semantic Web is a major change that evolved after the introduction of the Semantic Web. The Semantic Web provides a set of best practices for publishing and interlinking structured data on the Web, also known as Linked Data [3]. Some fundamental concepts of the Semantic Web are an RDF data model, the RDF Vocabulary Definition Language (RDFS), the Web Ontology Language (OWL), a semantic query language (SPARQL), and a database to store Linked Data, also known as a triple store. An RDF data model is an abstract data model that represents web resources in the Semantic Web. It structures data in the form of subject, predicate, and object. This structure is also known as triples. The RDFS and the OWL provide a modelling language to develop ontologies and vocabularies that can be used in the Semantic Web to describe entities in the world and the relationships between these entities [3,12]. “SPARQL Protocol and RDF Query Language” (SPARQL)⁴ is a query language for RDF data, whereas a triple store is a database to store RDF data. SPARQL is an official W3C Recommendation used to query RDF data stored in triple stores.

The Semantic Web is the World Wide Web Consortium’s (W3C)⁵ vision of the Web of Linked Data, while the Geospatial Semantic Web is an extension of the Semantic Web especially catering to geospatial Linked Data. The Geospatial Semantic Web incorporates geospatial data and Semantic Web concepts. Geospatial data have distinct features such as geometry; a coordinate reference system; the topology of geometries, which require special components such as geospatial ontology; a geospatial query language; and a geospatial triple store to express, store and query geospatial data in the Geospatial Semantic Web. A geospatial query language called GeoSPARQL has been developed as one of the components of the Geospatial Semantic Web. It is a small vocabulary to describe geospatial information and a query language used in the Geospatial Semantic Web [13,14]. It is a standard created by the OGC⁶-non-profit organization that implements open standards for the global geospatial community. Strabon⁷ and Parliament⁸ (when paired with Apache Jena⁹) are some of the well-known examples of geospatial triple stores that store geospatial RDF data, providing endpoints for sending GeoSPARQL queries.

⁴ <https://www.w3.org/TR/sparql11-query/> (last accessed on 8 June 2020)

⁵ <https://www.w3.org/standards/semanticweb/> (last accessed on 8 June 2020)

⁶ <https://www.ogc.org/> (last accessed on 8 June 2020)

⁷ <http://www.strabon.di.uoa.gr/> (last accessed on 8 June 2020)

⁸ <https://github.com/SemWebCentral/parliament> (last accessed on 8 June 2020)

⁹ <https://jena.apache.org/> (last accessed on 8 June 2020)

2.2. The Geospatial Semantic Web and Cultural Heritage

CIDOC CRM is an ISO standardized ontology developed by the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM) for CH organizations and institutions. It enables heterogeneous CH information to be integrated and interchanged, which is achieved by providing some of the necessary classes and properties to define concepts and relationships in CH data. The ontology follows an object-oriented data model and can be extended if needed [15]. One example of a CIDOC CRM extension is the previously mentioned CRMgeo that integrates the GeoSPARQL with CIDOC CRM to encode the spatio-temporal aspects of CH data [6]. Even though the Geospatial Semantic web is a relatively new technology, it has been successfully applied in several research projects in the CH domain. For instance, Pelagios [16,17], WarSampo [7], Omeka¹⁰, and Getty Thesaurus of Geographic Names (TGN)¹¹, among others, have all successfully implemented geospatial semantic web concepts.

3. Methodology

3.1. Criteria for Selecting RDF Generation Tools

In the last two decades, many RDF generation tools have been developed, such as RMLMapper and XSPARQL. These enable users to convert various types of data such as structured, semi-structured, and unstructured data to RDF data. Many of these tools require considerable expertise in Semantic Web and Geospatial Semantic Web concepts to be used successfully. However, as mentioned earlier, in this article we aim to evaluate RDF data-generating tools that support geospatial data and do not require computer programming. Hence, to select suitable tools, the flowchart illustrated in Figure 1 has been developed to facilitate the selection of RDF generation tools. The first process in the diagram is to identify potential RDF generation tools, which are found in three different resources. The first resource is the official “RDF Generation Tools¹²” list by the World Wide Web Consortium (W3C). The second resource is scholarly research articles that incorporated Semantic Web and Geospatial Semantic Web technologies as well as their applications in the CH domain. Finally, the third resource includes search results returned from different search engines. The second process in the diagram is to identify whether the tool requires computer programming or not. If it does not, the tool is included in the evaluation, otherwise it is excluded. The third process in the diagram is to find out if the tool supports geospatial data, such as vector data or raster data, stored in geospatial file formats or geospatial databases.

3.2. Criteria for Selecting RDF Linking Frameworks

Generally, linking RDF data in many interlinking frameworks includes 3 steps: preprocessing, linking configuration, and postprocessing. In the first step, the data to be interlinked are cleaned and transformed as in many cases datasets may contain irregular characters. After this step linking configurations must be specified, such as properties of the RDF data based on which matching is compiled, and a similarity comparator must be selected (e.g., Jaccard similarity measure, Jaro-Winkler similarity measure, Levenshtein similarity measure). The postprocessing step includes selecting the best match with the highest confidence value. This step may involve domain experts who verify the provided match results. Currently, there are several interlinking frameworks available that allow finding links between related RDF datasets. The main objective of these tools is to discover semantically equivalent resources in related RDF datasets and link them using a relation property such as owl:sameAs [18]. The owl:sameAs¹³ property indicates that two Uniform Resource Identifier (URI)

¹⁰ <https://omeka.org/s/> (last accessed on 8 June 2020)

¹¹ <https://www.getty.edu/research/tools/vocabularies/tgn/> (last accessed on 8 June 2020)

¹² https://www.w3.org/2001/sw/wiki/Category:RDF_Generator (last accessed on 8 June 2020)

¹³ <https://www.w3.org/TR/owl-ref/#sameAs-def> (last accessed on 8 June 2020)

references refer to the same thing. In the article, we have compared three state-of-the-art interlinking frameworks, namely LIMES, Silk, and OpenRefine with an RDF extension. This is mainly because other frameworks such as SLINT+¹⁴, SERIMI¹⁵, KNOFUSS¹⁶ either do not have a graphical user interface or require semantic web domain experts to use them. Furthermore, some are not publicly available to download.

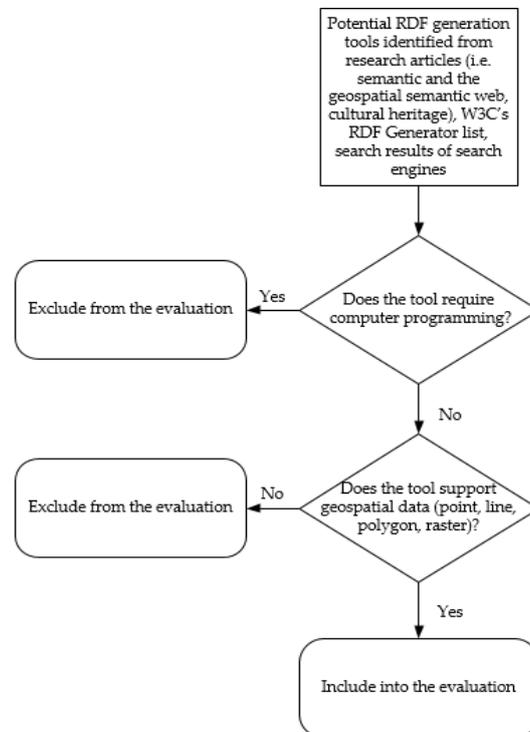


Figure 1. A flowchart diagram for selecting geospatial Resource Description Framework (RDF) generation tools.

4. Comparative Evaluation

4.1. A Comparative Evaluation of RDF Data Generation Tools

In the next paragraphs, the geospatial RDF tools Karma Data Integration Tool, GeoTriples, TripleGeo, and OpenRefine with an RDF extension are comparatively evaluated. These tools were selected based on the criteria denoted in Figure 1. Firstly, a comparative evaluation of RDF data generation tools is provided. Afterwards, a detailed discussion of each tool is presented.

The above-mentioned tools are geospatial RDF data generation tools that can convert geospatial and non-geospatial data to RDF data. To compare the technical features of these RDF generation tools, a feature comparison table was developed (see Table 1) that describes the technical features provided by each RDF generation tool. The technical features include supported input file formats and types, output file formats, built-in GeoSPARQL ontology compatible export option, ontology import option, and others. It can be seen from Table 1 that KARMA Data Integration Tool can take as an input data geospatial database PostGIS and JSON file format, which enables users to convert vector geospatial data to RDF data. GeoTriples and TripleGeo are purpose-built geospatial RDF data generation tools. Hence, they support many geospatial file formats such as ESRI Shapefiles, GML,

¹⁴ <http://ri-www.nii.ac.jp/SLINT/index.html> (last accessed on 8 June 2020)

¹⁵ <https://github.com/samuraujo/SERIMI-RDF-Interlinking> (last accessed on 8 June 2020)

¹⁶ <https://code.google.com/archive/p/knofuss/downloads> (last accessed on 8 June 2020)

KML, and geospatial databases such as PostGIS. OpenRefine is not a purpose-built geospatial RDF data generation tool. Nevertheless, it supports converting vector geospatial data stored in JSON, CSV, TSV file formats to RDF data. Regarding the supported output file formats, all of these RDF generation tools support export option to RDF. However, Karma Data Integration Tool also provides an export option to JSON-LD. As for the built-in GeoSPARQL ontology export option, purpose-built geospatial tools GeoTriples and TripleGeo support this feature, but the other two do not. Conversely, an ontology import option is provided by Karma Data Integration Tool and OpenRefine, whereas GeoTriples and TripleGeo do not support this feature out-of-the-box.

Table 1. Feature comparison of RDF generation tools.

	Karma Data Integration Tool	GeoTriples	TripleGeo	OpenRefine Version 3.3 with an RDF Refine Extension
Developed by	Center on Knowledge Graphs in Information Sciences Institute, the University of Southern California	EU Projects LEO and Melodies, National and Kapodistrian University of Athens	EU Project GeoKnow, the Athena Research Center	Freebase, then Google, now open source community
Supported Input File Formats and Types	<ul style="list-style-type: none"> A database table SQL query sent to the databases: SQL Server, MySQL, Oracle, PostGIS, Sybase Web APIs CSV and other delimited text files JSON, KML and XML MS Excel files 	<ul style="list-style-type: none"> ESRI Shapefiles XML GML KML JSON and GeoJSON CSV Databases: PostGIS and MonetDB 	<ul style="list-style-type: none"> ESRI Shapefiles GML KML Databases: Oracle Spatial, PostGIS, MySQL, IBM DB2 with Spatial extender 	<ul style="list-style-type: none"> JSON TSV CSV XML RDF various syntaxes MS Excel files Google Data Document Supports additional formats by extensions Databases: PostgreSQL, MySQL, MariaDB
Supported Output File Formats	RDF—Turtle syntax, JSON-LD	RDF—N-Triples syntax	RDF/XML, N-Triples, N3, Turtle	RDF/XML, Turtle
Built-in GeoSPARQL ontology compatible export option	No	Yes	Yes	No
Ontology import option	Yes	No (a user needs to edit the R2RML/RML mapping document to achieve that)	No (Supports exporting according to GeoSPARQL ontology, WGS84 RDF Geoposition Vocabulary and the Virtuoso RDF Vocabulary)	Yes
Origin	Karma is written in a Java programming language	D2RQ Engine (http://d2rq.org/)	Geometry2RDF (https://github.com/boricles/geometry2rdf/tree/master/Geometry2RDF)	OpenRefine is written in a Java programming language
Licence	Apache Licence 2.0	Apache Licence 2.0	GPL-3.0	BSD 3-Clause “New” or “Revised” Licence
Source Code	https://github.com/usc-isi-i2/Web-Karma	https://github.com/LinkedEOData/GeoTriples	https://github.com/GeoKnow/TripleGeo	https://github.com/OpenRefine/OpenRefine
Official Website	https://usc-isi-i2.github.io/karma/	http://geotriples.di.uoa.gr/	http://geoknow.eu/Project.html	https://openrefine.org/
Other	Supports importing and exporting R2RML models	Built-in support for the geospatial ontology stSPARQL	Built-in support for the geospatial ontologies WGS84 RDF Geoposition Vocabulary and Virtuoso RDF Vocabulary for point features	Offers reconciliation/linking two RDF datasets based on a SPARQL endpoint, RDF file, and Apache Standbol’s Entity Hub

One remarkable tool for generating geospatial RDF data is Karma¹⁷, which is free and open-source software (Apache Licence 2.0) and enables users to produce RDF data from several data sources such

¹⁷ <https://usc-isi-i2.github.io/karma/> (last accessed on 8 June 2020)

as spreadsheets, JSON documents, XML, KML file formats and Web APIs. The tool has been developed by the Center on Knowledge Graphs in Information Sciences Institute, the University of Southern California, and applied in some use cases such as integration of bioinformatics data, geospatial data, and Smithsonian Art Museum data. It offers a graphical user interface to upload an ontology and map the data according to the uploaded ontology. In other words, the mapped RDF data will be compatible with the defined ontology [19]. As it is possible to import ontologies in RDFS/RDF formats, CIDOC CRM can easily be imported to the Karma Data Integration Tool. The latest version of the CIDOC CRM with the “.rdf” file extension can be found at the official website of the CIDOC CRM¹⁸. However, the tool does not support some of the extensively used GIS file formats such as ESRI Shapefiles, and raster file formats.

GeoTriples is a tool that makes it easy to generate geospatial RDF data from geospatial data stored in file formats such as ESRI Shapefiles or spatially-enabled relational databases. It has been developed by the European projects LEO and Melodies. The tool is free and open-source (Apache Licence 2.0) and consists of two main components, the mapping generator, and the mapping processor, as shown in Figure 2. The mapping generator takes, as input, the geospatial data source along with other configurations. Based on the defined input and configurations, the tool creates so-called R2RML or RML mapping, which includes rules for generating RDF data. The rules in the mapping describe how the data should be represented in the RDF data model. Once the mapping is ready, the mapping processor executes the mapping to generate RDF data. By default, the tool makes RDF data compatible with the GeoSPARQL ontology. It is also possible to use a different ontology in the tool, for instance, the CIDOC CRM ontology. For this, a user should edit the mapping. However, the tool does not provide the possibility to edit the mapping out-of-the-box. Hence, it must be completed manually. To perform the above-mentioned actions, a user can use a command line or the graphical user interface provided in the tool [20]. The tool also supports Ontology-Based Data Access Engine, which facilitates on-the-fly GeoSPARQL to SQL translations. To accomplish this, a user provides, as input, a geospatial database such as PostGIS or Oracle Spatial and creates a mapping. Once the mapping is created, it can be used on another tool called Ontop-spatial¹⁹, which performs the translation from GeoSPARQL to SQL. This approach allows for the accessing of geospatial data stored in the geospatial databases as linked geospatial data using GeoSPARQL queries [21].

TripleGeo is another tool for converting geospatial data into RDF data. This tool is free and open-source (GPL-3.0 Licence) and was developed by the European project GeoKnow: Making the Web an Exploratory for Geospatial Knowledge. It can take as input geospatial data in file formats such as ESRI shapefiles, KML, GML, and geospatial databases and convert them into RDF data. The tool is based on the utility called Geometry2RDF²¹ and provides a command-line interface [22]. A user of the tool should provide the geospatial data source and configuration in the command line.

OpenRefine (previously Google Refine) is a free and open-source tool (BSD Licence) that allows users to work with data, including cleaning and transforming data from one format into another. OpenRefine does not support the capability to convert data into RDF data out-of-the-box. However, it has an RDF extension that allows datasets to be converted into RDF data. Another feature this extension provides is a function to link two RDF datasets. In other words, it can identify equivalent resources in two RDF datasets by comparing entities in the datasets. Once the tool processes entity matching, it is possible to generate RDF data. The tool offers this service in three ways, which are linking based on a SPARQL endpoint, RDF file, and Apache Standbol Entity Hub.

¹⁸ <http://www.cidoc-crm.org/Version/version-5.0.4> (last accessed on 8 June 2020)

¹⁹ <http://ontop-spatial.di.uoa.gr/> (last accessed on 8 June 2020)

²¹ <https://github.com/boricles/geometry2rdf/tree/master/Geometry2RDF> (last accessed on 8 June 2020)

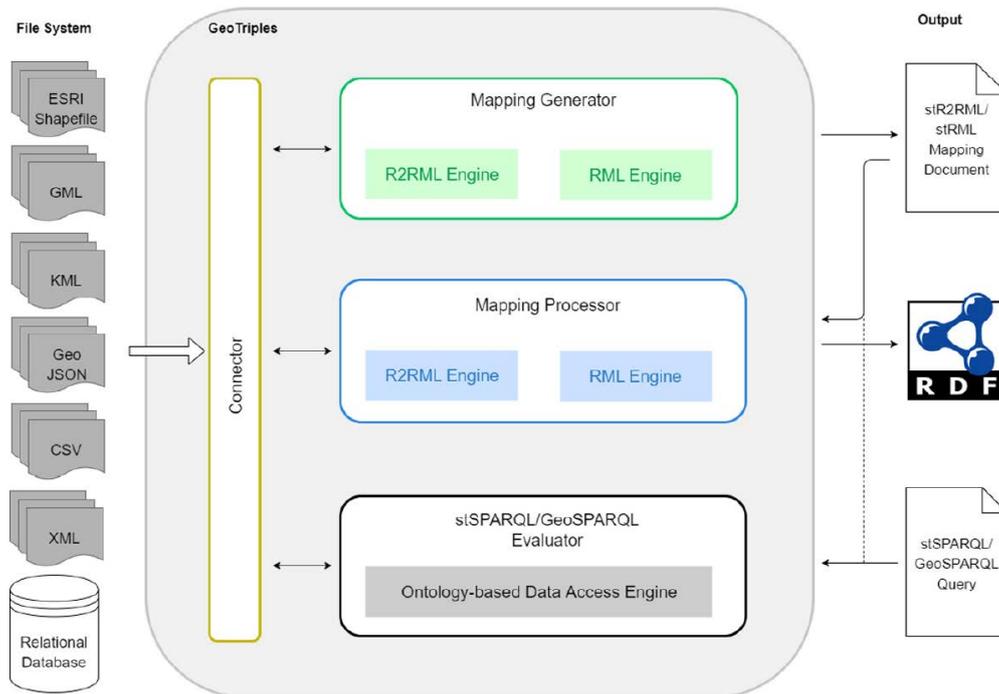


Figure 2. A workflow for generating RDF in GeoTriples²⁰.

4.2. A Comparative Evaluation of RDF Linking Frameworks

The next paragraphs present a comparative evaluation of RDF linking frameworks. As mentioned previously, this article comparatively evaluates three state-of-the-art interlinking frameworks, namely LIMES, Silk, and OpenRefine with an RDF extension. There are also other RDF linking frameworks such as SLINT+, SERIMI, KNOFUSS. These RDF linking frameworks were not included in the article as they either do not have a graphical user interface or require semantic web domain experts to use them. Furthermore, some are not publicly available for downloading. Next, a comparative evaluation of RDF linking frameworks is provided. Then, a detailed discussion of each framework is presented.

LIMES, Silk, and OpenRefine with an RDF extension support various technical features such as input formats, output formats, matching techniques to find links in two RDF data sources, and pre-processing functions. To compare the supported features of RDF linking frameworks, Table 2 was developed. As can be seen from Table 2, all of these frameworks accept as input data various RDF syntaxes. However, only LIMES and Silk frameworks can fetch RDF data stored in a SPARQL endpoint. In respect to supported output formats, all these frameworks support at least two different RDF syntaxes, while only LIMES supports tab-separated values (TSV) and comma-separated values (CSV) in addition to RDF syntaxes. Concerning matching techniques, LIMES and Silk frameworks support various matching techniques such as string and numeric, whereas OpenRefine with an RDF extension only provides a string-based matching technique.

²⁰ Reprinted from Journal of Web Semantics, Volumes 52–53, Kyzirakos et al, GeoTriples: Transforming geospatial data into RDF graphs using R2RML and RML mappings, Pages 16–32, Copyright 2020, with permission from Elsevier.

Table 2. Feature comparison of RDF linking frameworks.

	LIMES Version 1.7.1	Silk Version 3.0.0	OpenRefine Version 3.3 with RDF Refine Extension
Developed by	Agile Knowledge Engineering and Semantic Web (AKSW), Leipzig University	University of Mannheim	Freebase, then Google, now open source community
Supported data input formats	<ul style="list-style-type: none"> N-Triples (N3) Turtle (TTL) Notation3 (N3) Tab Separated Values (TAB) Comma Separated Values (CSV) SPARQL endpoint 	<ul style="list-style-type: none"> Alignment Comma Separated Values (CSV) JavaScript Object Notation (JSON) RDF XML SPARQL endpoint 	<ul style="list-style-type: none"> JSON TSV CSV XML RDF various syntaxes MS Excel files Google Data Document Supports additional formats by extensions Databases: PostgreSQL, MySQL, MariaDB
Supported output formats	Turtle (TTL) N-Triples (N3) Tab Separated Values (TAB) Comma Separated Values (CSV)	N-Triples (N3) Alignment (http://alignapi.gforge.inria.fr/format.html)	RDF/XML Turtle (TTL)
Matching technique/measures	String Vector Space Point-set Topological Temporal Resource-set Edge-counting semantic	Asian Character-based Equality Numeric Token-based	String
Licence	Gnu Affero General Public Licence	Apache Licence, Version 2	BSD 3-Clause “New” or “Revised” Licence
Source Code	https://github.com/dice-group/LIMES	https://github.com/silk-framework/silk/releases	https://github.com/OpenRefine/OpenRefine
Official Website	http://aksw.org/Projects/LIMES.html	http://silkframework.org/	https://openrefine.org/
Required software to execute the framework	Java SDK 12 (or later) and Maven 3.6.2 (or later)	Java JDK 8 Simple Build Tool (sbt) Yarn dependency management tool	None
Pre-processing functions	Yes Supports various types of pre-processing functions such as converting a string to lowercase, uppercase, replacing a string character, etc.	Yes Supports many types of pre-processing functions such as string to lowercase, removing whitespaces in strings, replacing a string character, format a number according to a user-defined pattern, etc.	Yes Supports many types of pre-processing functions such as replacing a string character, transforming strings to uppercase, lowercase, etc.

LIMES (Link Discovery Framework for Metric Spaces) is a free and open-source interlinking framework for the Semantic Web. It can discover links between entities in Linked Data sources such as two related RDF files. It can also find links between an RDF file and existing published RDF data sources such as DBpedia (via a SPARQL endpoint). The framework also implements some machine-learning algorithms to semi-automatically learn interlinking specifications. It provides a command-line interface and graphical user interface (GUI) that allow users to specify interlinking configurations and to execute the interlinking process. LIMES framework supports many types of interlinking techniques, called similarity measures, that can be used in various linking cases. For example, it supports string measures such as ExactMatch (compares two strings to determine if they are identical), RatcliffObershelp (calculated by dividing the matching characters of two strings by the total number of characters in those strings), and vector space measures such as Euclidean and Manhattan distance, to name just a few. In terms of geospatial RDF data, it supports point-set measures such as Geo_Max (maximum distance between pairwise points of the two input geometries), Geo_Min (minimum distance between pairwise points of the two input geometries), Geo_Mean, Geo_Avg, etc.; topological measures such as Top_Contains, Top_Covers, Top_Crosses, Top_Equals, Top_Intersects,

Top_Overlaps, etc.; and temporal measures including Tmp_After, Tmp_Before, Tmp_Overlaps, etc. These geospatial measures can be used in many use cases such as interlinking geospatial RDF data based on topological relations, temporal (time-based) relations, and geographical distance.

Silk is another free and open-source RDF data interlinking framework developed by the University of Mannheim. This framework can be used to generate links in two related RDF data sources. It also allows for the interlinking of RDF data sources with a published RDF data source using a SPARQL endpoint. This feature is beneficial when RDF data need to be interlinked with already published large RDF data sources such as DBpedia or LinkedGeodata. The framework is shipped together with a graphical user interface called the Silk Workbench that allows users to easily create link specifications and execute a link discovery process. It supports many types of link discovery measures called comparators in this framework. Jaccard, String Equality, and Numeric similarity measures are some of the examples of supported link discovery measures in the framework. It also supports some discovery measures that can be used on geospatial RDF data. For instance, a geographical distance measure (computes the geographical distance between two points) or an inside numeric interval measure (checks if a number is contained inside a numeric interval such as 1900 to 2000) can be used on geospatial RDF data. However, more complex geospatial link discovery measures such as topological measures (e.g., spatially overlap, within) are not supported in the framework.

OpenRefine, as mentioned previously, is a data cleaning and transformation tool that has an extension to convert data sources into an RDF data model. This extension allows data cleaning and conversion to an RDF data model, generating links in two related RDF data sources. It can also generate links between a local RDF data source and a published data source stored in a triple store. In this case, the published data source should be reachable from a SPARQL endpoint. The tool supports string-based link discovery measures, however, the geospatial link discovery measures such as overlap, within, etc. are not supported.

5. Methodology for Producing Linked Geospatial Cultural Heritage Data

A methodology for producing linked geospatial CH data consists of five steps, as shown in Figure 3. The first step is data preparation. In many cases, text in the data includes whitespaces and other irregular characters. Hence, in this step, the data are cleaned by removing whitespaces and making them compatible with a UTF-8 encoding standard. This procedure can also be performed by RDF generation tools as many support various types of data transformations. The second step is to use an RDF mapping generation tool to map geospatial CH data into the preferred ontology. Once the mapping is complete, RDF data can be exported and fed into Step 3. In the third step, an interlinking framework should be used to discover links in two related datasets. Afterwards, the interlinked RDF data can be stored in a triple store, and query languages such as SPARQL or GeoSPARQL can be used to query the data. In the last step, a Geospatial Semantic Web application can be built with the interlinked RDF data stored in a triple store. Furthermore, the RDF data can be submitted to the LOD Cloud²² that stores the collection of RDF data accessible to people and machines.

To demonstrate the applicability of this methodology, we mapped sample geospatial CH data and interlinked them with the related RDF data from DBpedia. The sample data are about CH places located in Western Australia. These data are freely available on the website of the Government of Western Australia²³. The attributes of the data include ID, name, address, and geolocation of the CH places among others. As an ontology, CIDOC CRM was used, while the mapping was achieved using the KARMA Data Integration Tool. Attributes of the data and mapping are illustrated in Figure 4. As discussed previously, in the second step the data should be interlinked. The SILK Interlinking Framework was selected as an RDF interlinking framework. The framework requires the data source

²² <https://lod-cloud.net/> (last accessed on 8 June 2020)

²³ <https://catalogue.data.wa.gov.au/dataset/heritage-council-wa-state-register> (last accessed on 8 June 2020)

and target source to be specified. In this case, the former is the Western Australian CH places data, while the latter is related to CH data from DBpedia. The SILK Interlinking Framework accepts data in several formats. For instance, data can be provided as a local file, URL, or SPARQL endpoint. We have provided both data source and target source as a local file. DBpedia data were included by querying the DBpedia SPARQL endpoint and downloading the query result. The SPARQL endpoint of the DBpedia provides query results in several RDF syntaxes such as RDF/XML, JSON, and XML+XSLT. The query and results are illustrated in Figure 5.

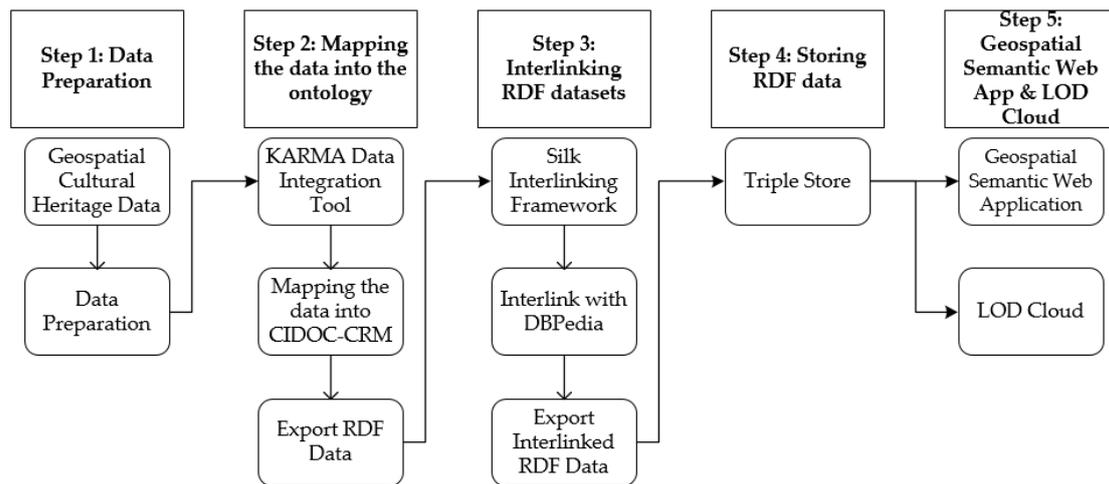


Figure 3. From geospatial cultural heritage data to linked geospatial cultural heritage data.

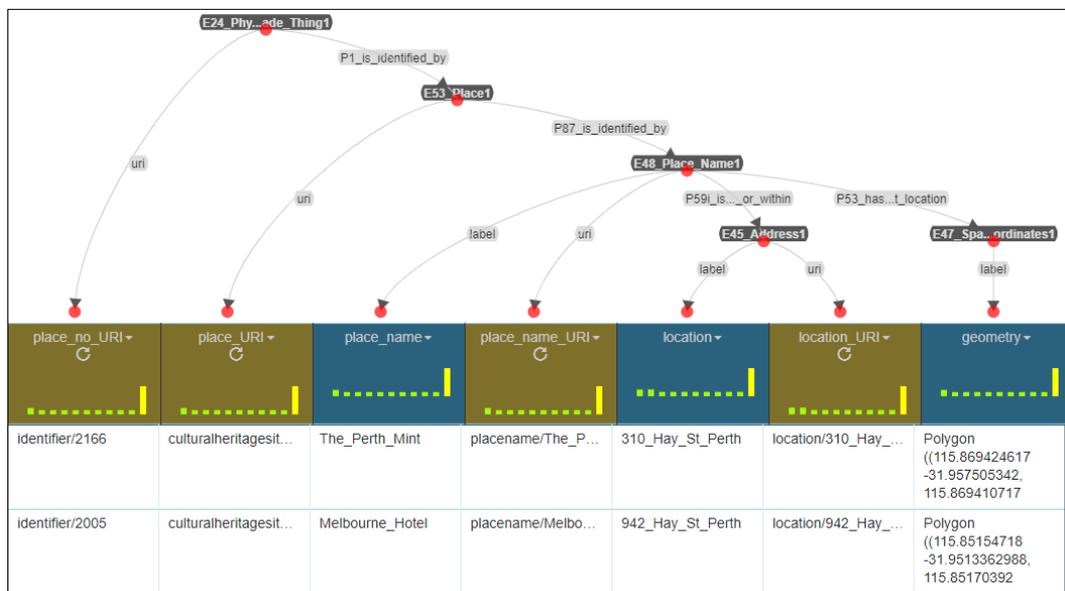


Figure 4. Mapping cultural heritage (CH) data into CIDOC CRM ontology in Karma Data Integration Tool. Copyright © 2020 the University of Southern Carolina. Apache Licence 2.0.

SPARQL Explorer for <http://dbpedia.org/sparql>

SPARQL:

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbpedia2: <http://dbpedia.org/property/>
PREFIX dbpedia: <http://dbpedia.org/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?property ?value
WHERE {
  { <http://dbpedia.org/resource/Category:Heritage_places_in_Perth,_Western_Australia> ?property ?isValueof }
  UNION
  { ?value ?property <http://dbpedia.org/resource/Category:Heritage_places_in_Perth,_Western_Australia> }
}

```

Results:

SPARQL results:

property	value
rdf:type	-
rdfs:label	-
owl:sameAs	-
owl:sameAs	-
dbpedia:ontology/wikiPageID	-
dbpedia:ontology/wikiPageRevisionID	-
skos:broader	-
skos:broader	-
skos:prefLabel	-
<http://www.w3.org/ns/prov#wasDerivedFrom>	-
<http://purl.org/dc/terms/subject>	:Palace_Hotel,_Perth
<http://purl.org/dc/terms/subject>	:St_Mary's_Cathedral,_Perth
<http://purl.org/dc/terms/subject>	:Cygnet_Cinema
<http://purl.org/dc/terms/subject>	:The_Cloisters,_Perth
<http://purl.org/dc/terms/subject>	:Plaza_Theatre,_Perth
<http://purl.org/dc/terms/subject>	:Walsh's_Building
<http://purl.org/dc/terms/subject>	:Perth_railway_station
<http://purl.org/dc/terms/subject>	:Tranby_House
<http://purl.org/dc/terms/subject>	:Luna_Leederville

Figure 5. Cultural heritage places of Western Australia in DBpedia SPARQL endpoint. Copyright © 2020 DBpedia. CC BY-SA 3.0.

Once the datasets are provided to the SILK Interlinking Framework, the exact property path for the interlinking entity should be specified. Then, the framework retrieves all values of the specified entity. The entity values often contain underscores or other characters that may require transformation before the interlinking process. In this case, the entity values include underscores that were replaced with spaces as shown in Figure 6. Afterwards, the entity values were changed to lowercase letters. As an interlinking comparator method, qGrams was used with a threshold value of 0.65, and the value of q was set to two. A threshold usually accepts a value between 0 and 1 and represents a confidence value. If the threshold has a greater value, the similarity measure provides a greater number of potential links but may involve more incorrect links. By contrast, if the threshold has a smaller value, the similarity measure finds a fewer number of potential links, but the result also contains fewer incorrect links. qGrams is a similarity measure that also accepts a value for q . Based on the value of q , a string is divided into a set of q -Grams. In this case, the value of q is set to two, which means q is replaced with two. As a result, it becomes 2-Grams. Hence, a string is divided into a set of two-character grams. For instance, a string “semantic” in 2-Grams is divided as follows (‘se’, ‘em’, ‘ma’, ‘an’, ‘nt’, ‘ti’, ‘ic’). Next, the measure calculates the similarity of the two input strings by counting the number of grams they share. For a more detailed discussion on similarity measures including qGrams, we recommend referring to a research article by Gali, et al. [23].

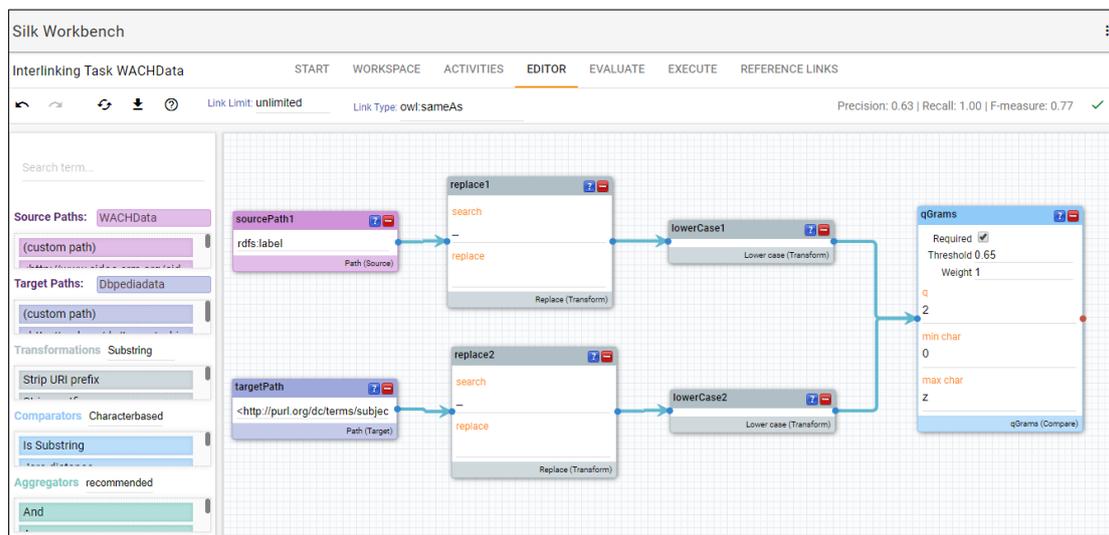


Figure 6. Configuration of the interlinking datasets in the Silk Interlinking Framework. Copyright © 2020 University of Mannheim. Apache Licence 2.0.

Then, the interlinking process was performed, as illustrated in Figure 7. This step usually involves verification of the correctness of the computed links, which is performed by people. In this interlinking method with a specified threshold, we identified that the computed link with above eighteen percent was correct while below eleven percent was incorrect. The last procedure in step three is to export the RDF set of links (owl:sameAs), which can be seen in Figure 8. As mentioned previously, in step four interlinking RDF data can be stored in a specialized RDF store called a triple store. The RDF data can then be accessed using SPARQL or GeoSPARQL query languages, and the result of the query can be used in Geospatial Semantic Web applications. Finally, it is also possible to publish the RDF data to the LOD Cloud.

Silk Workbench

Interlinking Task WACHData START WORKSPACE ACTIVITIES EDITOR **EVALUATE** EXECUTE REFERENCE LINKS

Finished in 1.439s

Filter:

Source: WACHData	Target: DbpediaData	Score	Correct
80/source/placename/Swanbourne_Hospital_Conservation_Area	http://dbpedia.org/resource/Swanbourne_Hospital	23.1%	✓ ? ✗
http://localhost:8080/source/placename/Tranby_House	http://dbpedia.org/resource/Tranby_House	100.0%	✓ ? ✗
http://localhost:8080/source/placename/Astor_Theatre	http://dbpedia.org/resource/Astor_Theatre_Perth	41.4%	✓ ? ✗
http://localhost:8080/source/placename/Palace_Hotel	http://dbpedia.org/resource/Palace_Hotel_Perth	34.1%	✓ ? ✗
http://localhost:8080/source/placename/Perth_Observatory	http://dbpedia.org/resource/Perth_Observatory	100.0%	✓ ? ✗
http://localhost:8080/source/placename/The_Perth_Mint	http://dbpedia.org/resource/Perth_Mint	48.7%	✓ ? ✗
http://localhost:8080/source/placename/Windsor_Hall	http://dbpedia.org/resource/Windsor_Cinema	10.9%	✓ ? ✗
!/source/placename/St_Georges_Hall_Portico_and_Facade_Perth	http://dbpedia.org/resource/St_Georges_Hall_Perth	25.0%	✓ ? ✗
http://localhost:8080/source/placename/Old_Mill	http://dbpedia.org/resource/Old_Mill_Perth	18.6%	✓ ? ✗
http://localhost:8080/source/placename/Old_Mill_Theatre	http://dbpedia.org/resource/Old_Mill_Perth	3.8%	✓ ? ✗
http://localhost:8080/source/placename/The_Weld_Club	http://dbpedia.org/resource/Weld_Club	38.5%	✓ ? ✗
http://localhost:8080/source/placename/Cygnets_Cinema	http://dbpedia.org/resource/Cygnets_Cinema	100.0%	✓ ? ✗
http://localhost:8080/source/placename/The_Cliffe	http://dbpedia.org/resource/The_Cliffe	100.0%	✓ ? ✗
http://localhost:8080/source/placename/The_Cloisters	http://dbpedia.org/resource/The_Cloisters_Perth	46.2%	✓ ? ✗
http://localhost:8080/source/placename/Melbourne_Hotel	http://dbpedia.org/resource/Melbourne_Hotel	100.0%	✓ ? ✗
!/localhost:8080/source/placename/Mooras_Post_Office_Quarters	http://dbpedia.org/resource/General_Post_Office_Perth	2.8%	✓ ? ✗

Figure 7. Interlinking process in Silk Interlinking Framework. Copyright © 2020 University of Mannheim. Apache Licence 2.0.

```

1 <http://localhost:8080/source/placename/Cygnets_Cinema> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Cygnets_Cinema> .
2 <http://localhost:8080/source/placename/St_Georges_Hall_Portico_and_Facade_Perth> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/
  St_Georges_Hall_Perth> .
3 <http://localhost:8080/source/placename/Swanbourne_Hospital_Conservation_Area> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Swanbourne_Hospital> .
4 <http://localhost:8080/source/placename/Tranby_House> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Tranby_House> .
5 <http://localhost:8080/source/placename/Palace_Hotel> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Palace_Hotel_Perth> .
6 <http://localhost:8080/source/placename/Old_Mill> <http://www.w3.org/2002/07/owl#sameAs> <
  http://dbpedia.org/resource/Old_Mill_Perth> .
7 <http://localhost:8080/source/placename/Old_Mill_Theatre> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Old_Mill_Perth> .
8 <http://localhost:8080/source/placename/The_Weld_Club> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Weld_Club> .
9 <http://localhost:8080/source/placename/Astor_Theatre> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Astor_Theatre_Perth> .
10 <http://localhost:8080/source/placename/Perth_Observatory> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Perth_Observatory> .
11 <http://localhost:8080/source/placename/The_Perth_Mint> <
  http://www.w3.org/2002/07/owl#sameAs> <http://dbpedia.org/resource/Perth_Mint> .
  
```

Figure 8. Results of the interlinking process.

6. Discussion

As discussed previously, in recent years there have been many research projects that developed RDF generation tools and interlinking frameworks. Many of these tools support a graphical user interface and do not require programming, thus enabling non-technical domain experts to employ them in research projects. However, the Semantic Web and the Geospatial Semantic Web are still relatively new and emerging technologies. Many challenges, therefore, need to be resolved before these tools and frameworks become widely used technology by non-semantic web professionals. For instance, as previously discussed, RDF generation tools do not require programming. However, they still require knowledge of the selected ontology. This means the user needs to have a solid knowledge of the classes and properties of the ontology as well as needing to know how to define the relationship between classes. On the other hand, the interlinking frameworks discussed in the article do not require

programming either. Some of them employ novel machine learning algorithms that ease the process of interlinking a great deal. However, in most cases, the interlinking frameworks cannot find the related data automatically without human intervention. They require someone to specify the path or class and property for the datasets being interlinked. As a result, they require users to have a solid knowledge of the ontology of the datasets to specify the correct path. This becomes especially cumbersome if the task is to interlink RDF data with a large knowledge base such as DBpedia. Furthermore, there are some other technical challenges such as representing raster data in an RDF data model as well as querying raster RDF data, 3D RDF data representation and query, storing, and representing big geospatial RDF data, among others. For a more detailed discussion on the above-mentioned concepts, we recommend referring to our previous research article [24].

7. Conclusions

In recent years, several RDF-generation tools and interlinking frameworks have been developed. However, many of them require a solid knowledge of Semantic Web and Geospatial Semantic Web concepts to be deployed successfully. Furthermore, according to recent research findings by Schmachtenberg, Bizer, and Paulheim [8], who analyzed the adoption of LOD best practices including interlinking in different domains such as media, life sciences, geographic, etc., 44% of published LOD datasets are not linked to other datasets at all.

This article did not attempt to conduct performance benchmarking of the RDF generation tools and interlinking frameworks. Instead, this article provided a comparative evaluation of features and functionality of the current state-of-the-art RDF generation tools and interlinking frameworks. This evaluation was specifically performed for CH researchers and professionals who do not have considerable expertise in computer programming. Hence, the geospatial RDF generation tools and interlinking frameworks were selected based on a pre-defined set of criteria.

Furthermore, the article presented a methodology to demonstrate how the evaluated tools and frameworks can be applied to generate geospatial Linked Data. To demonstrate the applicability of the methodology, it was applied in a sample use case that uses geospatial CH data relating to Western Australian CH places. These data were mapped into the CIDOC CRM ontology, which was then interlinked with the related data from DBpedia. Although this methodology has been developed for CH researchers and professionals, it can be adopted by other domain professionals as well.

Finally, it provided a discussion of some of the key challenges and limitations that CH researchers and professionals may encounter when using the evaluated tools and frameworks, including the methodology.

Author Contributions: Conceptualization, Methodology, and Investigation: I.N.; Writing Original Draft Preparation: I.N.; Writing—Review and Editing: I.N., E.C., D.A.M.; Supervision: E.C., D.A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Campanaro, D.M.; Landeschi, G.; Dell'unto, N.; Leander Touati, A.-M. 3D GIS for cultural heritage restoration: A 'white box' workflow. *J. Cult. Herit.* **2016**, *18*, 321–332. [[CrossRef](#)]
2. Soler, F.; Melero, F.J.; Luzón, M.V. A complete 3D information system for cultural heritage documentation. *J. Cult. Herit.* **2017**, *23*, 49–57. [[CrossRef](#)]
3. Bizer, C.; Heath, T.; Berners-Lee, T. Linked data—the story so far. *Int. J. Semant. Web Inf. Syst.* **2009**, *5*, 1–22. [[CrossRef](#)]
4. Egenhofer, M.J. Toward the semantic geospatial web. In Proceedings of the 10th ACM international symposium on Advances in geographic information systems, McLean, VA, USA, 8–9 November 2002; pp. 1–4.

5. Zhang, C.; Zhao, T.; Li, W. Conceptual Frameworks of Geospatial Semantic Web. In *Geospatial Semantic Web*; Springer: Cham, Switzerland, 2015; pp. 35–56.
6. Hiebel, G.; Doerr, M.; Eide, Ø. CRMgeo: A spatiotemporal extension of CIDOC-CRM. *Int. J. Digit. Libr.* **2017**, *18*, 271–279. [[CrossRef](#)]
7. Hyvönen, E.; Heino, E.; Leskinen, P.; Ikkala, E.; Koho, M.; Tamper, M.; Tuominen, J.; Mäkelä, E. WarSampo Data Service and Semantic Portal for Publishing Linked Open Data About the Second World War History. In Proceedings of the 13th International Conference, ESWC: European Semantic Web Conference, Heraklion, Greece, 29 May–2 June 2016; pp. 758–773.
8. Schmachtenberg, M.; Bizer, C.; Paulheim, H. Adoption of the linked data best practices in different topical domains. In Proceedings of the 13th International Semantic Web Conference, Riva del Garda, Italy, 19–23 October 2014; pp. 245–260.
9. Trizio, I.; Savini, F.; Giannangeli, A. Integration of three-dimensional digital models and 3d gis: The documentation of the medieval burials of amiternum (l'aquila, Italy). *Int. Arch. Photogramm.* **2018**, *42*, 1121–1128. [[CrossRef](#)]
10. Lezzerini, M.; Antonelli, F.; Columbu, S.; Gadducci, R.; Marradi, A.; Miriello, D.; Parodi, L.; Secchiari, L.; Lazzeri, A. Cultural heritage documentation and conservation: Three-dimensional (3D) laser scanning and Geographical Information System (GIS) techniques for thematic mapping of facade stonework of St. Nicholas Church (Pisa, Italy). *Int. J. Archit. Herit.* **2016**, *10*, 9–19. [[CrossRef](#)]
11. Elfadaly, A.; Lasaponara, R.; Murgante, B.; Qelichi, M.M. Cultural heritage management using analysis of satellite images and advanced GIS techniques at East Luxor, Egypt and Kangavar, Iran (a comparison case study). In Proceedings of the International Conference on Computational Science and Its Applications, Trieste, Italy, 3–6 July 2017; pp. 152–168.
12. Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. [[CrossRef](#)]
13. Battle, R.; Kolas, D. Geosparql: Enabling a geospatial semantic web. *Semant. Web J.* **2011**, *3*, 355–370. [[CrossRef](#)]
14. GeoSPARQL—A Geographic Query Language for RDF Data. Available online: <https://www.ogc.org/standards/geosparql> (accessed on 8 June 2020).
15. CIDOC CRM. Available online: <http://www.cidoc-crm.org/Version/version-5.0.4> (accessed on 8 June 2020).
16. Simon, R.; Barker, E.; Isaksen, L. Exploring Pelagios: A visual browser for geo-tagged datasets. In Proceedings of the International Workshop on Supporting Users' Exploration of Digital Libraries, Paphos, Cyprus, 23–27 September 2012.
17. Isaksen, L.; Simon, R.; Barker, E.T.; de Soto Cañamares, P. Pelagios and the emerging graph of ancient world data. In Proceedings of the 2014 ACM Conference on Web Science, Bloomington, IN, USA, 23–26 June 2014; pp. 197–201.
18. Nentwig, M.; Hartung, M.; Ngonga Ngomo, A.-C.; Rahm, E. A survey of current link discovery frameworks. *Semant. Web* **2017**, *8*, 419–436. [[CrossRef](#)]
19. Knoblock, C.A.; Szekely, P.; Ambite, J.L.; Goel, A.; Gupta, S.; Lerman, K.; Muslea, M.; Taheriyani, M.; Mallick, P. Semi-automatically mapping structured sources into the semantic web. In Proceedings of the 9th Extended Semantic Web Conference, Heraklion, Greece, 27–31 May 2012; pp. 375–390.
20. Kyzirakos, K.; Savva, D.; Vlachopoulos, I.; Vasileiou, A.; Karalis, N.; Koubarakis, M.; Manegold, S. GeoTriples: Transforming geospatial data into RDF graphs using R2RML and RML mappings. *J. Web Semant.* **2018**, *52*, 16–32. [[CrossRef](#)]
21. Bereta, K.; Xiao, G.; Koubarakis, M. Answering Geosparql Queries over Relational Data. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2017**, *42*, 43–50. [[CrossRef](#)]
22. Patroumpas, K.; Alexakis, M.; Giannopoulos, G.; Athanasiou, S. TripleGeo: An ETL Tool for Transforming Geospatial Data into RDF Triples. In Proceedings of the EDBT/ICDT Workshops, Athens, Greece, 28 March 2014; pp. 275–278.
23. Gali, N.; Marinescu-Istodor, R.; Fränti, P. Similarity measures for title matching. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 1548–1553.
24. Nishanbaev, I.; Champion, E.; McMeekin, D.A. A Survey of Geospatial Semantic Web for Cultural Heritage. *Heritage* **2019**, *2*, 1471–1498. [[CrossRef](#)]

