School of Electrical Engineering, Computing and Mathematical Sciences

A Web GIS-based Integration of 3D Digital Models with Linked Open Data for Cultural Heritage Exploration

Ikrom Nishanbaev

0000-0003-0032-3283

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University

December 2021

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Human Research Ethics Approval

The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number HRE2021-0155.

Ikrom Nishanbaev

December 2021

Abstract

The amount of digital cultural heritage is growing rapidly as cultural heritage sites and collections are being digitized at exponentially accelerating rates by cultural heritage projects and institutions around the globe. Digital cultural heritage repositories have therefore become an efficient and effective way to disseminate and exploit digital cultural heritage data including 3D digital models. However, cultural heritage data produced by many cultural heritage projects and institutions is heterogeneous and sparsely interlinked, which means these cultural heritage datasets may involve different data types including geospatial data, formats, semantic heterogeneity issues, and sparse interlinking of related datasets. Publishing cultural heritage using semantic web concepts, also known as linked open data, is currently being adopted by many cultural heritage institutions and projects to address these problems.

Nevertheless, it is still a challenging task for many cultural heritage professionals and researchers to generate and interlink geospatial linked open data as many linked open data generation tools and interlinking frameworks require solid expertise in computer programming to be deployed successfully. Another main reason for that is a lack of clear methodology and guidelines.

Another challenge is the long-term archiving of 3D digital cultural heritage models including interlinking them with knowledge bases such as DBpedia and GeoNames. Long-term archiving of 3D digital cultural heritage models is imperative in digital cultural heritage conservation, while interlinking them with knowledge bases may enhance digital cultural heritage exploration and offer new dimensions of interaction and exploration of digital cultural heritage.

Thus, this research project explored how geospatial semantic web concepts, 3D web-based visualisation, digital interactive map, and cloud computing concepts could be integrated to enhance digital cultural heritage exploration; to offer long-term archiving and dissemination of 3D digital cultural heritage models; to better interlink heterogeneous and sparse cultural heritage data.

The project started by conducting a survey of the geospatial semantic web concepts pertinent to cultural heritage. This included geospatial semantic web tools, frameworks, cultural heritage domain-specific ontologies, free and open-source purpose-built semantic web, and the geospatial semantic web platforms which can be used to implement geospatial semantic web cultural heritage repositories. This survey concluded with a discussion on the major limitations of the geospatial semantic web, which are limiting its use in cultural heritage.

Next, a complete methodology and guidelines were developed for generating geospatial linked open data including interlinking the resulting linked open data to knowledge bases such as DBpedia. Since this methodology was developed specifically for cultural heritage researchers and professionals, who are usually non-programmers, it does not involve computer programming or scripting and excludes all commercial tools and frameworks, thus facilitating easy and cost-free replication.

Afterwards, the project proposed a reusable, completely extendable novel web repository as well as methodology based on free and open-source, easy-to-implement frameworks for long-term archiving and dissemination of 3D digital cultural heritage models. This methodology and web repository were then integrated with cloud computing concepts as cloud computing offers enormous capabilities such as almost unlimited storage space and computing capacity among others, which are advantageous to long-term archiving and dissemination of 3D cultural heritage models.

Finally, the above-mentioned web repository and methodology were further extended to enhance digital cultural heritage exploration by integrating cloud computing concepts, digital interactive map, linked open data from DBpedia and GeoNames, and 3D digital cultural heritage models into a novel web-GIS- based cloud architecture. The architecture was validated by applying it to specific case studies of Australian cultural heritage and seeking expert feedback on the system, its benefits, and scope for improvement in the near future.

The research findings in the survey can equip cultural heritage researchers and professionals with an overview of the geospatial semantic web concepts pertinent to cultural heritage, whereas a complete methodology and guidelines on generating geospatial linked open data can help them better interlink cultural heritage data using geospatial semantic web concepts, without requiring computer programming or scripting.

A novel web-GIS architecture including methodology can be used for long-term archiving and dissemination of 3D digital cultural heritage models as well as for enhancing digital cultural heritage exploration.

Keywords

3D cultural heritage models, cultural heritage, linked open data, the semantic web, the geospatial semantic web, web-GIS, GIS.

Publications during PhD candidature

Research article 1:

Nishanbaev, I.; Champion, E.; McMeekin, D. A. (2021). A Web GIS-Based Integration of 3D Digital Models with Linked Open Data for Cultural Heritage Exploration. ISPRS International Journal of Geo-Information, 10(10), 684. https://doi.org/10.3390/ijgi10100684

Research article 2:

Nishanbaev, I. (2020). A Cloud Architecture for Processing and Visualization of Geo-located 3D Digital Cultural Heritage Models. In the proceedings of 7th International Conference on Geographical Information Systems Theory, Applications and Management (pp. 51-61).
https://doi.org/10.5220/0009341500510061

(This research article received "Best Student Paper Award" http://www.gistam.org/PreviousAwards.aspx)

Research article 3:

Nishanbaev, I. (2020). A web repository for geo-located 3D digital cultural heritage models. *Digital Applications in Archaeology and Cultural Heritage*, *16*, e00139.
https://doi.org/10.1016/j.daach.2020.e00139

Research article 4:

Nishanbaev, I., Champion, E., & McMeekin, D. A. (2020). A Comparative Evaluation of Geospatial Semantic Web Frameworks for Cultural Heritage. *Heritage*, *3*(3), 875-890.
https://doi.org/10.3390/heritage3030048

Research article 5:

Nishanbaev, I., Champion, E., & McMeekin, D. A. (2019). A survey of geospatial semantic web for cultural heritage. *Heritage*, *2*(2), 1471-1498. https://doi.org/10.3390/heritage2020093

Supervisors:

21.08.2020 to completion

Dr David A. McMeekin (Primary supervisor)

Professor Erik Champion (Secondary supervisor)

11.12.2017 to 21.08.2020

Professor Erik Champion (Primary supervisor)

Dr David A. McMeekin (Secondary supervisor)

Author Contributions

Co-author attribution - Ikrom Nishanbaev

	Research article 1	**Research article 2**	**Research article 3**	**Research article 4**	**Research article 5**
Conceptualization	100%	100%	100%	100%	100%
Methodology	100%	100%	100%	100%	100%
Software implementation / other implementation	100%	100%	100%	100%	100%
Survey implementation	75%	Not applicable	Not applicable	Not applicable	Not applicable
Writing - review and editing	85%	100%	100%	70%	70%

I acknowledge that these represent my contribution to the above research output

Signature Date

Co-author attribution - Professor Erik Champion

	Research article 1	**Research article 2**	**Research article 3**	**Research article 4**	**Research article 5**
Conceptualization	X	X	X	X	X
Methodology	X	X	X	X	X
Software implementation / other implementation	X	X	X	X	X
Survey implementation	15%	Not applicable	Not applicable	Not applicable	Not applicable
Writing - review and editing	15%	X	X	15%	15%

I acknowledge that these represent my contribution to the above research output

Signature Date

Co-author attribution - Dr David A. McMeekin

	Research article 1	**Research article 2**	**Research article 3**	**Research article 4**	**Research article 5**
Conceptualization	X	X	X	X	X
Methodology	X	X	X	X	X
Software implementation / other implementation	X	X	X	X	X
Survey implementation	10%	Not applicable	Not applicable	Not applicable	Not applicable
Writing - review and editing	X	X	X	15%	15%

I acknowledge that these represent my contribution to the above research output

Signature Date

Acknowledgements

I would like to firstly acknowledge the essential funding support I received from the Curtin International Postgraduate Research Scholarship (CIPRS) and UNESCO Chair of Cultural Heritage and Visualisation project.

A huge thank you to my supervisors Professor Erik Champion and Dr. David McMeekin for their high-quality supervision and continuous academic support throughout the PhD studies.

Thank you to my wife, Adiba, who supported me all the way through this journey. I would have never completed this research work without her support. Thank you also to my daughter, Sadia, who helped me to overcome the challenges in my PhD studies and in life.

I would like to acknowledge and thank Digital Technology and Solutions at Curtin University for providing IT support and access to Amazon Web Services cloud services.

Finally, I would like to thank cultural heritage researchers and professionals, who participated in the research survey, for their constructive feedback and recommendations. Also, thank you to academic journal reviewers, who peer-reviewed my research articles and provided constructive feedback on them.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Acronym	Definition
3D	3D Geographic Information System
AAT	Art and Architecture Thesaurus
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
AR	Augmented Reality
ARIADNE	Advanced Research Infrastructure for Archaeological Dataset Networking in Europe
AWS	Amazon Web Services
BIM	Building Information Modelling
CAA	Computer Applications and Quantitative Methods in Archaeology
CH	Cultural Heritage
CIDOC-CRM	CIDOC Conceptual Reference Model (CRM)
CIPA	The International Committee for Documentation of Cultural Heritage
CIPA	International Committee of Architectural Photogrammetry
CIPRS	Curtin International Postgraduate Research Scholarship
CMS	Content Management System
COLLADA	Collaborative Design Activity
CRS	Coordinate Reference System
CSV	Comma-Separated Values
DBMS	Database Management Systems
DCAT	Data Catalog Vocabulary
EBS	Amazon Elastic Block Store
ENVI	Environment for Visualizing Images
ESRI	Environmental Systems Research Institute
EU	European Union
EuroMed	The European Mediterranean Conferences
FAIR	Findable, Accessible, Interoperable, and Reusable
FOSS	Free and Open-Source Software
GeoSPARQL	A Geographic Query Language for RDF Data
GIS	Geographic Information System
GML	Geographic Markup Language

H-BIM	Heritage Building Information Modelling
HTML	The HyperText Markup Language
IBM	International Business Machines Corporation
ICOMOS	International Council of Monuments and Sites
INCEPTION	Inclusive Cultural Heritage in Europe through 3D Semantic Modelling
ISO	The International Organization for Standardization
ISPRS	International Journal of Geo-Information
IT	Information technology
JSON	JavaScript Object Notation
JSON-LD	JavaScript Object Notation for Linked Data
KML	Keyhole Markup Language
LD	Linked Data
LOD	Linked Open Data
MIT	Massachusetts Institute of Technology
MTL	Material Template Library
OGC	Open Geospatial Consortium
OWL	Web Ontology Language
PARTHENOS	Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies
PC	Personal Computer
PhD	Doctor of Philosophy
PHP	Hypertext Preprocessor
QGIS	Quantum Geographic Information System
RDBMS	Relational Data Base System
RDF	Resource Description Framework
SCEA	Shared Source Licence
SOCH	Share Our Cultural Heritage
SPARQL	SPARQL Protocol and RDF Query Language
SQL	Structured Query Language
TGN	Getty Thesaurus of Geographic Names
TSV	Tab-Separated Values
Turtle	Terse RDF Triple Language
UAV	Unmanned Aerial Vehicles
UNESCO	The United Nations Educational, Scientific and Cultural

	Organization
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
VR	Virtual Reality
VRE	Virtual Research Environment
VRML	Virtual Reality Modelling Language
VSMM	The International Society for Virtual Systems and Multimedia
WCMS	Web Content Management Systems
Web-GIS	Web-Geographic Information System
WGS_84	World Geodetic System 1984
WissKI	Wissenschaftliche KommunikationsInfrastruktur
WKT	Well-Known Text
XML	Extensible Markup Language

Chapter 1 Introduction

1.1 Cultural heritage and web-GIS

Tangible cultural heritage (e.g., buildings, historical maps, landscapes, artifacts,) is often exposed to different natural disasters such as flooding and, earthquakes as well as man-made hazards such as fires or human negligence. Thus, digital preservation of tangible cultural heritage is vitally important to ensure that the shape and texture of the cultural heritage are not lost in case of damage by natural or man-made disasters [1]. Hence, in the last 20 years, considerable efforts have been made by cultural heritage organizations around the world to digitize cultural heritage artifacts, sites, and historical maps, etc. for digital preservation and online representation [2,3]. On the other hand, many research projects and studies have been published that demonstrate the high potential of web-Geographic Information Systems (web-GIS) in the dissemination and online representation of cultural heritage data [4,5].

GIS is a system for collecting, storing, querying, analysing, and displaying geospatial data. GIS can be exploited for different case studies. For instance, GIS mobile applications (e.g., ArcGIS Collector, QField) can be used to precisely capture field data. Spatial databases (e.g., PostgreSQL/PostGIS) facilitate storing and querying geospatial data including asking questions relating to location and time. Furthermore, almost all spatial databases offer spatial relationships that allow comparing relationships between geometries; this cannot usually be achieved in non-GIS systems. Desktop GIS applications such as QGIS, ArcMap, ENVI, and cloud-based geospatial analysis platforms such as Google Earth Engine can be used to analyse and display geospatial data including analysing raster imagery captured by drones and satellites. Cultural heritage domain has a long history of applying GIS in a wide range of cases such as management and monitoring of cultural heritage landscapes and sites [6], documentation and conservation of cultural heritage sites [7], recording and managing cultural heritage sites [8] to name just a few.

On the other hand, web-GIS is a web-based system that performs the above-mentioned tasks on the Internet by utilizing web technologies and services. In other words, web-GIS is an online GIS that is powered by modern GIS and web technologies. As it uses web technologies, it requires a server-client architecture whereby the server is a GIS server, and the client could be a web browser, a desktop application, or a mobile application. Thus, it makes GIS more accessible and customizable, and can also remove technical obstacles such as limited hardware capacity and complicated software installations among others.

A typical web-GIS architecture is shown in Figure 1. When a client sends a geospatial request to the server, the server processes the request and sends the geospatial response back to the client. If the requested information is not available on the server, it informs the client accordingly.

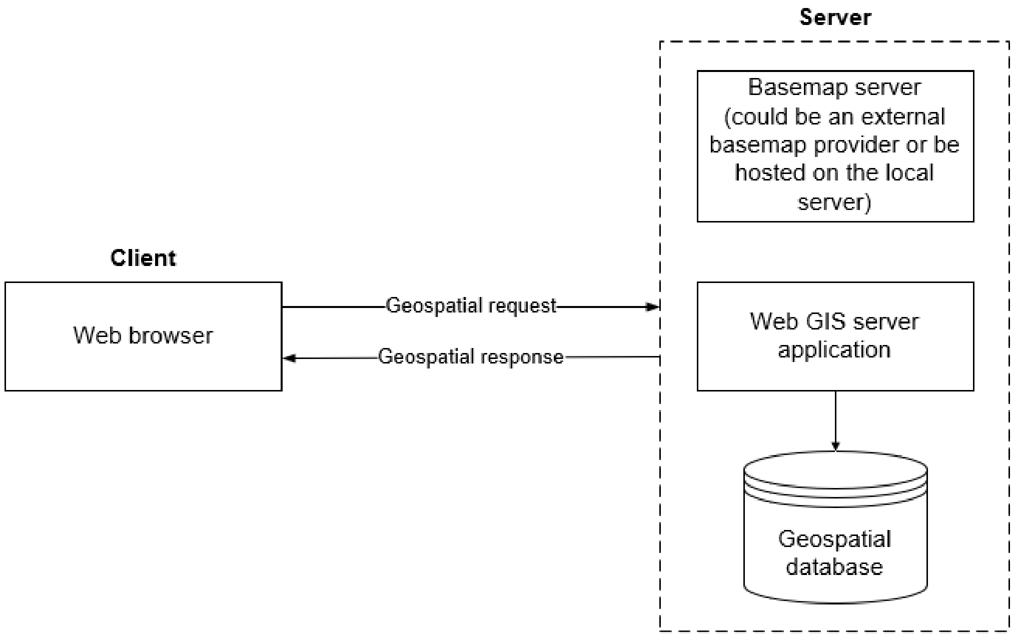


Figure 1. A typical web-GIS architecture.

From its early inception, the web-GIS has found its applications in cultural heritage as it can be used to manage geospatial data irrespective of the domain. Thus, web-GIS can be applied to archaeological [9], historical [10], environmental data [11], or geospatial data from any other domain. However, a minimum prerequisite for the data is it needs to have geolocation information so that a web-GIS system could plot the data on the digital map.

Since web-GIS allows users to store, edit and disseminate cultural heritage data over the Internet, many cultural heritage research works and studies were implemented in the past, which prove the importance of web-GIS to cultural heritage, especially in the cultural heritage conservation and promotion [12] [13].

1.2 Web-GIS and web-based 3D visualisation

Web-based 3D visualisation has recently become commonplace in many domains including in cultural heritage. One of the main reasons for this was the introduction of Web Graphics Library (WebGL) which brought about plugin-free 3D rendering to the web. WebGL is a cross-platform, royalty-free web standard developed by the Khronos Group. These days most web browsers natively support WebGL without requiring an additional plugin installation at the user end. However, WebGL is a low-level API (application programming interface), thus it is cumbersome and relatively complicated to use raw WebGL to implement 3D graphics and scenes. Nevertheless, there are many high-level libraries such as three.js, Babylon.js, CesiumJS, which take advantage of the advanced web technologies including WebGL and offer an easy-to-use environment to developers in order to visualise and create sophisticated 3D graphics in a web browser. Furthermore, in recent years many lightweight, performance, and precision-oriented 3D formats, such as 3D Tiles and glTF, for the web have emerged. These advanced web-based 3D visualization technologies and 3D file formats along with web-GIS systems are now being widely utilized by cultural heritage researchers and professionals from all around the world. Thus, the above-mentioned applications of web-GIS to cultural heritage dissemination, conservation, analysis, etc may now include not only 2D geospatial cultural heritage content but also 3D geolocated and time-referenced cultural heritage models.

1.3 The semantic web and linked open data

The idea of the World Wide Web was first introduced by Tim-Berners-Lee in 1989. It has now become the largest information hub for reading, sharing, and exchanging information, as well as the largest platform for online interacting with people. Since its inception, there has been a great deal of progress in web technologies. One of the relatively recent technological advancements to the web is the semantic web, which is also known as the web of data, and web 3.0. The semantic web extends the core principles of the existing web rather than replacing it and adds semantics to the web content so that the machines can process the data and understand the meaning behind the data. The semantic web offers some standardized components including ontologies which could be used to give well-defined meaning to data, a resource description framework (RDF) data model to represent web resources in a machine-readable format, SPARQL Protocol and RDF Query Language (SPARQL) to query data in the semantic web, triple stores to store RDF data to name a few [14]. The structured and interlinked data in the semantic web is known as linked data. In contrast, if the linked data is open data, which is freely available for everyone to use, this data is referred to as linked open data [15].

On the other hand, the geospatial semantic web is an extension of the semantic web which is used for catering geospatial linked data. Since vector geospatial data has distinct features such as the geometry of the data (e.g., point, line, polygon), a spatial reference system, a topology of geometries among others, the geospatial linked data require special components to express, represent and query them in the geospatial semantic web. A geospatial query language named GeoSPARQL has been developed, which serves two purposes in the geospatial semantic web. It is a small geospatial vocabulary to describe geospatial information in the geospatial semantic web, and a geospatial query language for querying geospatial linked data. It is also worth mentioning that GeoSPARQL is a standard created by the Open Geospatial Consortium (OGC) – a non-profit organization that implements open standards for the global geospatial community. There are also some geospatial triple stores for storing geospatial linked open, and geospatial data models that can represent geospatial linked data. These components along with the raster data representation in the geospatial semantic web are discussed in detail in later chapters.

The cultural heritage domain, as an early adopter of various digital technologies, has already started adopting semantic web and geospatial semantic web concepts. Cultural heritage data published cultural heritage organizations such as museums, libraries, and archives has often complex and diverse data relationships. A very simple example is a cultural heritage object may be movable or immovable, may have condition states such as reconstructed, ruined, wrecked, etc, may have different condition states for certain historical time periods among others. Furthermore, a cultural heritage object may be interrelated to other cultural heritage objects. To describe and document cultural heritage data concepts and relationships in a consistent and standardized way across cultural heritage organizations, CIDOC CRM Special Interest Group has developed an ontology called CIDOC Conceptual Reference Model (CRM)[1]. This cultural heritage ontology intends to be a common one in the cultural heritage domain and provide the 'semantic glue' for cultural heritage data published by cultural heritage organizations. Furthermore, in recent years many purpose-built cultural heritage platforms have been developed, such as WissKI, Arches, ResearchSpace, that may facilitate publishing cultural heritage data as linked open data. These purpose-built cultural heritage platforms along with other geospatial semantic web concepts pertinent to the cultural heritage field are discussed in detail in later chapters.

1.4 Research aims

The main aims of this PhD project were to:

- Equip cultural heritage researchers and professionals with an overview of the geospatial semantic web concepts pertinent to cultural heritage including tools, frameworks, and FOSS purpose-built semantic web and the geospatial semantic web platforms that can be exploited to implement geospatial semantic web cultural heritage repositories.
- Develop guidelines and methodology, which do not involve programming or scripting, for generating geospatial linked open data including RDF

1 http://www.cidoc-crm.org/ (last accessed on 11 October 2021)

data generation using CIDOC-CRM ontology and interlinking the resulting RDF data to knowledge bases such as DBpedia.

- Develop a reusable, completely extendable, novel web repository as well as methodology based on FOSS, easy-to-implement frameworks for long-term archiving and dissemination of 3D cultural heritage models.
- Enhance digital cultural heritage exploration by integrating cloud computing concepts, digital interactive map, linked open data from DBpedia and GeoNames, and 3D digital cultural heritage models into a novel web-GIS-based cloud architecture.

1.5 Research objectives and questions

1.5.1 Research objective one

The geospatial semantic web is a domain-independent which could be applied to data from any domain including cultural heritage. However, currently, almost all disciplines such as cultural heritage, archaeology, and medicine have domain-specific ontologies. These domain ontologies are developed due to the reason that every discipline has its own set of concepts and jargon which could not be expressed by general, domain-independent ontologies. In recent years, cultural heritage developed a few ontologies such as CIDOC-CRM, CRMgeo, and implemented some purpose-built platforms for publishing cultural heritage data as linked open data. There have also been some large-scale cultural heritage projects that successfully employed geospatial semantic web concepts to cultural heritage data. Nevertheless, there are a few technical challenges and limitations of the geospatial semantic web such as representing raster data as linked open data, 3D geospatial semantic web, etc, which are limiting the use of the geospatial semantic web in cultural heritage. Thus, this PhD project started by exploring, analysing, and identifying key benefits and major limitations in the current state-of-the-art in geospatial semantic web concepts pertinent to cultural heritage.

The specific research questions derived from this research objective were:

What is the current state-of-the-art in geospatial semantic web concepts pertinent to the cultural heritage domain?

What are the current major limitations of the geospatial semantic web that are limiting its use in cultural heritage?

1.5.2 Research objective two

Recently, many RDF generation tools have been developed that can facilitate converting geospatial and non-geospatial data into RDF data. Moreover, there are some interlinking frameworks that can find semantically equivalent geospatial resources in related RDF data sources. However, most of the RDF generation tools and interlinking frameworks require a solid knowledge of semantic web and geospatial semantic web concepts to successfully deploy them. Thus, generating geospatial linked open data and interlinking it with the related RDF data is a challenging task for many domain professionals including cultural heritage. Furthermore, many existing linked open data sources are sparsely interlinked which was demonstrated by the research findings of Schmachtenberg*, et al.* [16]. Schmachtenberg, Bizer and Paulheim [16] analysed the adoption of linked open data best practices, including interlinking, in different disciplines such as media and life sciences and concluded that 44% of published linked open data datasets are not linked to any other datasets. Thus, the next research objective in the PhD project was to provide a comparative evaluation of features and functionality of the geospatial RDF generation tools and interlinking frameworks which do not require expertise in computer programming, as many cultural heritage researchers and professionals may have limited expertise in computer programming. Furthermore, another sub-objective within this objective was to implement a complete methodology in order to demonstrate RDF data generation using CIDOC-CRM ontology, and to interlink the resulting RDF data with a leading knowledge base of DBpedia.

The specific research question corresponding to this objective was:

How could cultural heritage people, as a non-programmer, transform cultural heritage data into linked open data and interlink it with other knowledge bases such as DBpedia without computer programming knowledge?

1.5.3 Research objective three

In recent years, web-based cultural heritage repositories (Omeka, 3D-COFORM) have started emerging for long-term archiving and dissemination of 3D cultural heritage models. Long-term archiving of 3D cultural heritage models is vitally important to ensure digital conservation of cultural heritage, whereas dissemination of cultural heritage on the web is essential to reach a wider audience. There are also some successful commercial web repositories such as Sketchfab and TurboSquid in which users can store and visualize 3D cultural heritage models on the web. Nevertheless, it is still challenging to find, use and re-use 3D cultural heritage models, especially the ones generated in research projects. In fact, Champion and Rahaman [17] surveyed 14 proceedings of major digital heritage events and conferences to examine the accessibility of 3D models in published research articles from 2012 to 2017. According to their findings, out of 1483 examined conference articles 17.9% or 264 articles incorporated 3D models or images of 3D models, and only 9 articles had accessible 3D content. Thus, cultural heritage professionals, cultural heritage organizations, and the cultural heritage domain in general need more projects, and easy-to-implement, re-usable and extendable methodologies to build a web repository for long-term archiving and dissemination of 3D cultural heritage models. Therefore, the next objective in the PhD project was to develop a reusable, completely extendable, novel web repository as well as methodology based on free and open-source, easy-to-implement frameworks for long-term archiving and dissemination of 3D cultural heritage models.

The specific research questions were:

Could digital interactive map, 3D digital cultural heritage models, and geospatial data such as geolocation be integrated to achieve long-term archiving and dissemination of 3D digital cultural heritage models?

Could this be achieved using free and open-source frameworks, at no cost?

1.5.4 Research objective four

Cloud computing refers to the on-demand delivery of software applications, compute power, database, storage, and other IT resources through an internet-based cloud services platform such as Amazon Web Services, Google Cloud, and Microsoft Azure to name just a few [18]. Cloud computing infrastructures offer many benefits over traditional IT infrastructures including but not limited to:

- Access as much or as little computing resources as you need. This benefit allows scaling up and down the computing resources as needed with only a few minutes' notice.
- Pay only when you consume computing resources and pay only for how much you consume. This means computing resources that are not being used can be automatically switched off and can be switched on when there is a demand for computing resources. Furthermore, payment is calculated based on how much computing resources you consume.
- Focus on your services, clients, and software application developments rather than on physical servers.

Considering cloud computing provides almost unlimited cloud computing power, storage, and other benefits such as listed above, the next objective in the PhD project was to integrate the web repository and methodology from the previous research objective with cloud computing concepts.

The specific research question relating to this research objective was:

How could cloud computing, digital interactive map, and geolocated 3D digital cultural heritage models be integrated for long-term archiving and dissemination of 3D digital cultural heritage models?

1.5.5 Research objective five

As mentioned previously, the cultural heritage domain has already started implementing linked open data concepts to cultural heritage data.

Nevertheless, there are not many research projects in which the above-mentioned concepts (cloud computing, digital maps, and 3D cultural heritage models) and linked open data from leading knowledge bases such as DBpedia and geographical databases such as GeoNames are integrated to enhance digital cultural heritage exploration. Moreover, the integration of these distinct technologies may offer new dimensions of interaction and exploration of digital cultural heritage. To demonstrate the high potential of these technologies, the last objective in the PhD project was to further extend the previous methodology and web repository by integrating with linked open data from DBpedia and GeoNames as well as raster data.

The research question derived from this research objective was:

Could we enhance digital cultural heritage exploration by integrating the above-mentioned concepts (cloud computing, digital maps, and geolocated 3D cultural heritage models) with linked open data from DBpedia and GeoNames platforms?

1.6 Significance and contributions

The following paragraphs describe the significance and contributions of this PhD project to the cultural heritage and GIS fields.

1.6.1 Significance and contributions of the research objective one

As mentioned previously, many cultural heritage institutions have started implementing semantic web and geospatial semantic web concepts to cultural heritage data. A conceptual survey of the geospatial semantic web concepts pertinent to the cultural heritage field was needed, as it may equip cultural heritage professionals and practitioners with an overview of all the necessary tools, and FOSS semantic web and geospatial semantic web platforms which could be used to facilitate geospatial semantic web-based cultural heritage platforms. Furthermore, at the time of publishing my research article entitled ‘A Survey of Geospatial Semantic Web for Cultural Heritage’ which addresses the research objective one, no one had conducted a conceptual survey of the geospatial semantic web concepts pertinent to the cultural heritage domain.

The contributions to cultural heritage and GIS fields resulting from this research objective are as follows:

- A review on the state-of-the-art geospatial semantic web concepts pertinent to the cultural heritage field.
- A complete methodology to convert geospatial cultural heritage data in a vector data model into machine-readable and processable RDF data, with a case study to demonstrate its applicability.
- Discussion on the major FOSS semantic web and geospatial semantic web-based purpose-built platforms for cultural heritage institutions.
- A summary of leading cultural heritage projects that successfully employed geospatial semantic web concepts.
- Discussion on the limitations of the geospatial semantic web concepts which may also result in the generation of new ideas and research questions for both GIS and cultural heritage fields.

1.6.2 Significance and contributions of the research objective two

Currently for many non-technical domain professionals and researchers, including for cultural heritage, generating linked open data, and interlinking it with related RDF data is a challenging task that may require substantial expertise in computer programming. Furthermore, many linked open data sources are currently sparsely interlinked [16]. Thus, the following contributions were produced from research objective two:

- A comparative evaluation of features and functionality of the current state-of-the-art geospatial RDF data generation tools and interlinking frameworks. This evaluation was specifically performed for cultural heritage professionals and researchers with limited expertise in computer programming. Thus, the tools and frameworks were selected based on pre-defined criteria which are discussed in detail in the relevant chapter.

- A complete methodology that demonstrates how these evaluated tools and frameworks can be applied to generate geospatial linked open data. The methodology employs CIDOC-CRM ontology, which is then interlinked with the related RDF data from DBpedia.
- Discussion on key challenges and limitations of the evaluated tools, frameworks, and the presented methodology.

1.6.3 Significance and contributions of the research objective three

As discussed previously, it is still challenging to find, use and re-use 3D cultural heritage models, especially the ones generated in research projects. Furthermore, in most cases, 3D cultural heritage models have associated geolocation that allows them to be mapped on the digital interactive map. Thus, cultural heritage professionals, researchers, and the cultural heritage domain in general need more projects and easy-to-implement, re-usable and extendable methodologies to build a web repository for long-term archiving and dissemination of 3D cultural heritage models. The contributions to cultural heritage and GIS fields resulted from the research objective three are as follows:

- An easy-to-implement, completely extendable, novel methodology to build a web repository for geolocated 3D cultural heritage based on FOSS frameworks and tools. The web repository integrates digital maps, 3D digital cultural heritage models, and geospatial data such as geolocation.
- The presented web repository can be implemented by cultural heritage institutions at no cost as all frameworks and tools are based on FOSS frameworks and tools.
- Unlike many previous projects and methodologies which developed the web repositories from the ground up, this web repository is based on FOSS, easy-to-implement, database-driven content management system namely KeystoneJS and other FOSS frameworks and tools, which significantly accelerate and ease the development process.

1.6.4 Significance and contributions of the research objective four

As mentioned previously, cloud computing infrastructures offer many technical benefits over traditional infrastructures. Some of the key benefits of cloud computing are scalability, almost unlimited computing power and storage capacity, data loss prevention using data backups to name just a few. All these benefits of cloud computing could be applied in digital cultural heritage platforms to solve various technical bottlenecks such as storage capacity, uneven workload where the user requests to the infrastructure fluctuate up and down, and data loss due to hardware failure. The contributions to cultural heritage and GIS fields resulting from the research objective four are as follows:

- A novel AWS-based cloud architecture that integrates cloud computing concepts, a digital interactive map, and 3D digital cultural heritage models. More specifically, this architecture extends the previous web repository and integrates it with cloud computing concepts.

1.6.5 Significance and contributions of the research objective five

Many cultural heritage organizations around the world are currently implementing semantic web concepts to cultural heritage data. Some of the common reasons for this initiative are: to achieve better cultural heritage data integration among cultural heritage institutions; to offer better cross-collection search interfaces to end-users; to better engage end-user; to better analyse and explain the links and relationships among cultural heritage sites, objects, artifacts, and others. However, there are not many research projects whereby the cloud computing concepts, 3D cultural heritage models, and linked open data DBpedia and GeoNames are integrated into a web-GIS architecture to enhance digital cultural heritage exploration. The integration of these distinct technologies may also offer new dimensions of interaction and exploration of digital cultural heritage. Furthermore, re-using existing linked open data from DBpedia and GeoNames promotes a reusable principle of the FAIR (findable, accessible, interoperable, and reusable) data principles.

The contributions to cultural heritage and GIS fields resulting from the research objective five are as follows:

- A novel cloud-based web-GIS architecture that integrates digital interactive map, linked open data from DBpedia and GeoNames platforms with the locally stored 3D digital cultural heritage models (https://unesco-chv.curtin.edu.au/mapplatform, last accessed on 30 November 2021).
- A completely re-designed digital interactive map that includes updated map controls, clustering of overlapped markers on the map, and visualisation of raster data in the form of a digital elevation model among others.
- Recommendations and suggestions on the platform, which were collected from cultural heritage researchers.

1.7 Scope of the research

As it can be seen from the descriptions of the research aims and objectives, this PhD project is interdisciplinary involving distinct fields and subfields such as cultural heritage, web-GIS, linked open data, cloud computing, and 3D web-based visualisation, as shown in However, it has some limits and coverage of the research, which are described in the next paragraphs.

In research objective two, the thesis presents a comparative evaluation of features and functionality of the current state-of-the-art RDF generation tools and interlinking frameworks. Since this evaluation was performed for cultural heritage researchers and professionals who might have limited expertise in computer programming, it did not attempt to conduct performance benchmarking of the tools and frameworks.

In the research objective three, the concepts such as 3D surveying and 3D modelling techniques, efficient search and retrieval of big cultural heritage data, and 3D web-GIS-based analysis were out of the scope of the research.

1.8 Thesis structure

The structure of the PhD thesis is presented in Figure 2. As shown in Figure 2, the thesis is subdivided into seven chapters. Chapters two to six resulted from the research objectives one to five respectively.

It is worth noting that this PhD thesis was completed in the form of the thesis by compilation in which the main chapters are peer-reviewed, published research articles.

As each of these research articles includes a literature review related to the subject topic, there is no separate literature review chapter included in the thesis.

Furthermore, since these chapters are stand-alone research articles published in different academic journals, their formats may differ based on the formatting requirements of each individual journal.

The next paragraphs discuss the content of each chapter in the PhD thesis.

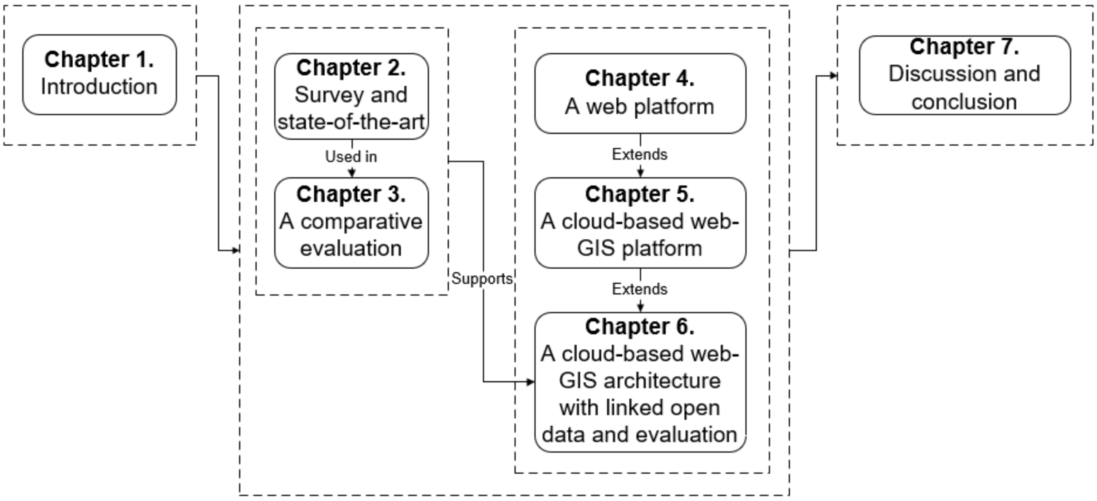


Figure 2. Thesis structure.

Chapter 1 provides an introduction to the research work including research objectives, significance, and contributions.

Chapter 2 reviews the state-of-the-art geospatial semantic web concepts pertinent to the cultural heritage field; presents a methodology to convert geospatial cultural heritage data into machine-readable and processable RDF data; discusses leading cultural heritage projects that successfully employed geospatial semantic web concepts; analyses major limitations and of geospatial semantic web concepts which are limiting its use in cultural heritage.

Chapter 3 comparatively evaluates features and functionality of the geospatial RDF generation tools and interlinking frameworks which do not require expertise in computer programming. It also presents a complete methodology to demonstrate RDF data generation based on CIDOC-CRM ontology and to interlink the resulting RDF data with one of the leading knowledge bases namely DBpedia. This methodology is based on the previous methodology in Chapter 2. The difference between them is the methodology in Chapter 2 does not deal with mapping cultural heritage data into an ontology such as CIDOC-CRM and does not demonstrate the interlinking process to knowledge bases such as DBpedia. Chapter 3 concludes with a discussion on the major limitations of the evaluated tools and frameworks which might be informative to cultural heritage researchers and professionals planning to use these tools and frameworks.

Chapter 4 presents a reusable, completely extendable, web platform as well as methodology based on FOSS, easy-to-implement frameworks for long-term archiving and dissemination of 3D cultural heritage models. This web platform integrates digital maps, 3D digital cultural heritage models, and geospatial data such as geolocation.

Chapter 5 extends the web platform and the methodology from the previous chapter and integrates cloud computing concepts into the web platform.

Chapter 6 further extends the web platform and the methodology presented in Chapter 5 and presents a novel cloud-based web-GIS architecture that integrates cloud computing, 3D digital models, and linked open data from DBpedia and GeoNames to enhance cultural heritage exploration. Furthermore, Chapter 6 presents recommendations and suggestions on the platform, which were collected from cultural heritage researchers through an online anonymous survey and direct online communication.

Chapter 7 discusses the outcomes and results of this PhD project including the limitations of the project. Furthermore, it proposes future directions which may tackle these limitations.

Chapter 2 A Survey of the Geospatial Semantic Web for Cultural Heritage[2]

This chapter provides a survey on the state-of-the-art in geospatial semantic web concepts pertinent to cultural heritage. Furthermore, it presents a) a methodology to convert geospatial cultural heritage data in a vector data model into machine-readable and processable RDF data; b) discussion on leading cultural heritage projects that employed geospatial semantic web concepts; c) and discussion on major limitations of the geospatial semantic web that are limiting the use of the geospatial semantic web in cultural heritage.

Abstract: The amount of digital cultural heritage data produced by cultural heritage institutions is growing rapidly. Digital cultural heritage repositories have therefore become an efficient and effective way to disseminate and exploit digital cultural heritage data. However, many digital cultural heritage repositories worldwide share technical challenges such as data integration and interoperability among national and regional digital cultural heritage repositories. The consequence is dispersed and poorly linked cultured heritage data, backed by non-standardized search interfaces, which thwart users' attempts to contextualize information from distributed repositories. A recently introduced geospatial semantic web is being adopted by a great many new and existing digital cultural heritage repositories to overcome these challenges. However, no one has yet conducted a conceptual survey of the geospatial semantic web concepts for a cultural heritage audience. A conceptual survey of these concepts pertinent to the cultural heritage field is therefore needed. Such a survey equips cultural heritage professionals and practitioners with an overview of all the necessary tools, and free and open-source semantic web and geospatial semantic web platforms that can be used to implement geospatial semantic web-based cultural heritage

2 This chapter is a copy of the peer-reviewed research article - Nishanbaev, I.; Champion, E.; McMeekin, D.A. A Survey of Geospatial Semantic Web for Cultural Heritage. Heritage 2019, 2, 1471–1498. https://doi.org/10.3390/heritage2020093

repositories. Hence, this article surveys the state-of-the-art geospatial semantic web concepts, which are pertinent to the cultural heritage field. It then proposes a framework to turn geospatial cultural heritage data into machine-readable and processable RDF data to use in the geospatial semantic web, with a case study to demonstrate its applicability. Furthermore, it outlines key free and open-source semantic web and geospatial semantic platforms for cultural heritage institutions. In addition, it examines leading cultural heritage projects employing the geospatial semantic web. Finally, the article discusses attributes of the geospatial semantic web that require more attention, that can result in generating new ideas and research questions for both the geospatial semantic web and cultural heritage fields.

2.1 Introduction

Recent advances in remote sensing technologies and imaging devices have opened up new possibilities for digital recording of cultural heritage (CH) sites and collections. Indeed, these technologies allow for the production of very realistic digital replica of CH sites and collections [19]. This generated CH data is often geographically referenced, thereby incorporating geographical location and time references, the resulting geospatial data often appears in a wide range of geospatial file formats. In turn, geospatial data, and location and time references can be used to discover interesting connections and relationships among cultural heritage resources. Hence, geospatial data is a major potential component in the CH field [20,21].

Furthermore, the volume of digital cultural heritage data is growing significantly, as CH sites and CH collections are being digitized at exponentially accelerating rates by CH digitization projects, government bodies and CH institutions (galleries, libraries, archives, and museums) around the globe [22,23]. Accordingly, the ever-increasing growth of digital CH data facilitated the development of cultural heritage repositories, which are playing a major role in digital preservation and dissemination, tourism, CH research among others [24,25].An apt example of this is the Europeana portal, which provides metadata information about digital cultural heritage content from European countries [26,27]. However, there are many other smaller repositories all over the world, in which data is published as raw dumps in different file formats lacking structure and semantics. Some of the major technical challenges in these repositories are data integration and interoperability among distributed repositories at the national and regional level. As a result, poorly linked CH data is fragmented in several national and regional repositories, backed by non-standardized search interfaces. All these technical challenges are limiting users' attempts to contextualize information from distributed repositories [28].

The semantic web is a new extension of the existing World Wide Web, it offers a set of best practices for publishing and interlinking structured data on the Web. In the semantic web, data is published in a machine-readable and processable format, and data has explicit meaning to machines defined by ontologies, and the resulting datasets are interlinked [14]. The geospatial semantic web is an extension of the semantic web, where geospatial data has explicit meaning to machines defined by geospatial ontologies. Geospatial data has distinct features such as geometries, a coordinate reference system (CRS), a topology of geometries, all of which require special attention when encoding into a machine-readable and processable format. In return, the geospatial semantic web offers capabilities such as GIS-based analysis, spatio-temporal queries and interlinking with other external data sources to provide geospatial context for the specific topic. These features of the geospatial semantic web are uncommon in the semantic web because the semantic web is not designed to deal with complex geospatial data [29,30]. Since the geospatial semantic web incorporates geospatial data as well as semantic web concepts, several new and existing CH repositories have adopted the geospatial semantic web concepts to deal with geospatial CH data and to overcome the above-mentioned issues of interoperability and integration [31-33]. Moreover, several web platforms based on these concepts have been developed to improve data integration, interoperability, long-term preservation, accessibility of digital CH data, such as WISSKI[3], Arches[4], OMEKA[5], etc.

As the geospatial semantic web is a new emerging concept, a limited number of surveys have been conducted. Li, *et al.* [34] reviewed the state of the geospatial semantic web, outlined challenges and opportunities of exploiting big geospatial data in the geospatial semantic web, and proposed future directions. Kuhn, *et al.* [35], Nalepa and Furmańska [36] explained major innovations in the field of Geographic Information Science brought about by semantic web concepts. Faye, *et al.* [37] and Rohloff, *et al.* [38] surveyed triple store technologies for linked data. Battle and Kolas [39] discussed the overall state of geospatial data in the semantic web including state of a query language GeoSPARQL in industry and research. Buccella, *et al.* [40] surveyed widely used approaches to integrate geospatial data using ontologies. Ballatore, *et al.* [41] reviewed geospatial data in open knowledge bases such as DBpedia, LinkedGeoData, GeoNames, OpenStreetMap, particularly paying attention to the quality of geodata, and to crowdsourced data.

However, there has not yet been conducted a conceptual survey of the geospatial semantic web concepts for a cultural heritage audience. A conceptual survey of these concepts pertinent to the cultural heritage field is therefore needed. Such a review equips CH professionals and practitioners with a useful overview of all the necessary tools, and free and open source (FOSS) semantic web and geospatial semantic web platforms that can be used to implement geospatial semantic web-based CH repositories.

Hence, the objectives of this article are as follows:

- To review the state-of-the-art geospatial semantic web concepts pertinent to the CH field.
- To propose a framework to turn geospatial CH data stored in a vector data model into machine-readable and processable RDF data in order

3 http://wiss-ki.eu/ (last accessed on 9 April 2019)
4 https://www.archesproject.org/ (last accessed on 9 April 2019)
5 https://omeka.org/ (last accessed on 9 April 2019)

to use in the geospatial semantic web with a case study to demonstrate its applicability.

- To outline key FOSS semantic web and geospatial semantic web-based purpose-built platforms for CH institutions.
- To summarize leading cultural heritage projects employing geospatial semantic web concepts.
- To discusses attributes of the geospatial semantic web concepts requiring more attention, which can result in the generation of new ideas and research questions for both the field of the geospatial semantic web and CH.

The remainder of the article is structured as follows: In Section 2, it discusses a web evolution and the state-of-art in the geospatial semantic web including cultural heritage domain ontologies. In Section 3, it proposes a framework to turn CH data into machine-readable and processable RDF data with its subsequent storage and retrieval. It demonstrates the applicability of this framework with a case study. Section 4 outlines key FOSS semantic web and geospatial semantic web-based purpose-built platforms for CH institutions. In Section 5, leading CH projects are presented that employ semantic web and geospatial semantic web concepts. Section 6 outlines challenges and technical limitations of the geospatial semantic web followed by Section 7's concluding summary.

2.2 Web evolution and the geospatial semantic web

The idea of the World Wide Web was first introduced by Tim-Berners-Lee in 1989. It has now become the largest global information space for sharing information. The web technologies witnessed a great deal of progress in the last two decades (see Figure 1). The first generation of the web, also known as a read-only web, web of documents, and Web 1.0, allows users only to search for information and read it. The websites are based on static HTML pages represented as hypertext in which elements can be linked to different resources. Hypertext Transfer Protocol (HTTP) is used as a means of communication between clients and servers [42].

The second phase in the evolution of the web is Web 2.0, also referred to as a read-write web and web of people. It has several interpretations (see for example [43]), but in general, it can be understood as an idea to describe the web as a user-centric and socially-connected platform, where people have more interaction and collaboration than in its predecessor. A set of new technologies have been introduced along with Web 2.0 such as AJAX (JavaScript and XML), Adobe Flex, Google Web Toolkit, etc. This allowed development of more dynamic, interactive websites and web applications, in which users can publish content to the web and modify the existing one. The foundations of Web 2.0 can also be demonstrated by the development of well-known web applications including social networking platforms (e.g. Facebook), web-blogs (e.g. WordPress), Wikipedia, media-sharing platforms (e.g. YouTube) among others [44].

In recent years, a new extension to the Web has started to emerge named Web 3.0, which is the third phase in web evolution. Web 3.0 is also known as the semantic web and the web of data. The previously discussed Web 1.0 is a read-only web to share information, and Web 2.0 is a read-write web which offers increased collaboration and social interaction among users. In the second phase, users have the capability to publish content to the web, with hypertext links to other resources, and machines can display it. However, it is only for human consumption and machines themselves do not have an ability to process and “understand” the content. The semantic web is being developed to overcome these problems. The semantic web extends the core principles of web 2.0 rather than replacing them, and adds semantics to the web by structuring web content and facilitating software applications to process and ‘understand’ the structured content. The semantic web incorporates several standardized components by the World Wide Web Consortium (W3C) such as Resource Description Framework (RDF) data model to represent web resources in a machine-readable format, ontologies to give well-defined meaning to data, SPARQL Protocol and RDF Query Language (SPARQL) to query data in the semantic web, triple stores to store

RDF data to name a few (to be discussed in later sections) [14]. This standardization in the semantic web provides a shared concept allowing cultural heritage data to be published, interlinked, searched and consumed more effectively than with web 2.0.

However, as mentioned in the introduction section, a problem with the semantic web arises when CH data is supplied as geospatial data in various geospatial file formats along with coordinate reference system (CRS), and topology. Geospatial data management in the semantic web is therefore tackled by its extension termed the geospatial semantic web. The geospatial semantic web brings together concepts of the semantic web and geospatial data, and adds a spatio-temporal dimension to the semantic web. CH data often includes geospatial data, for instance, current and historical places, CH objects and artifacts, and historical maps are typically associated with some type of locational and time-based references. To encode geospatial CH data into a machine-readable and processable format, the geospatial semantic web offers distinct technical specifications, including:

- A RDF data model
- Geospatial ontologies of the geospatial semantic web
- Geospatial query languages of the geospatial semantic web
- Spatio-temporal triple stores to store geospatial RDF data

In the next sub-sections, these specifications are discussed in detail.

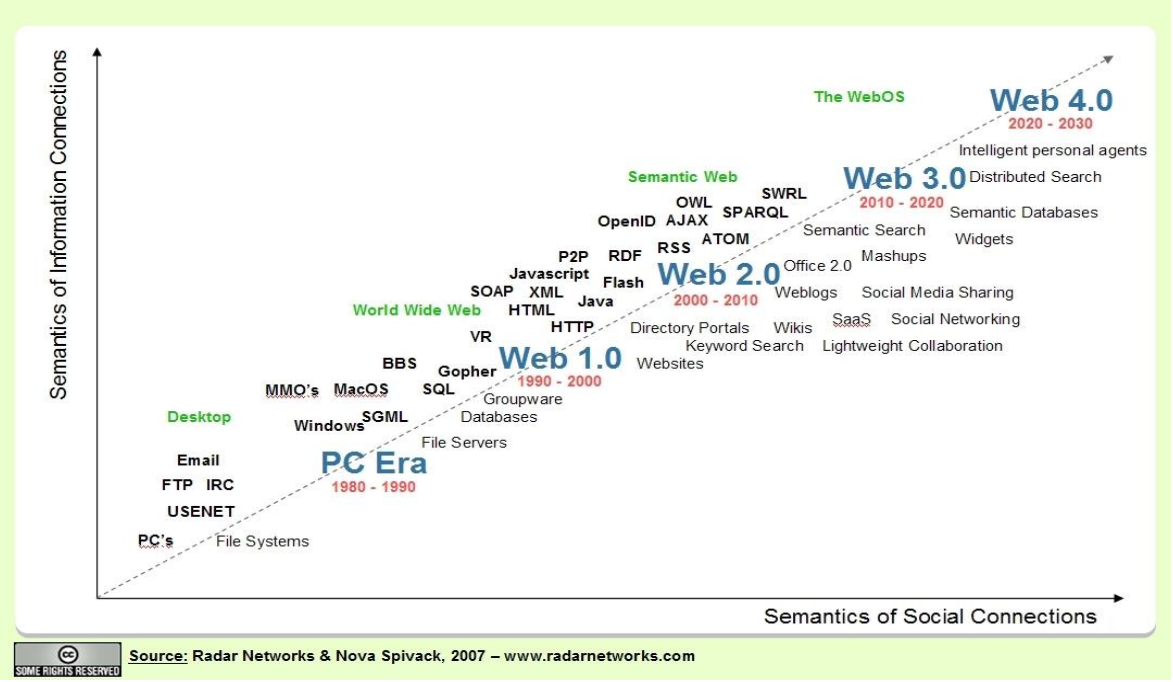


Figure 3. Web evolution from PC Era to Web 4.0.

Source: http://www.novaspivack.com/technology/web-3-0-the-best-official-definition-imaginable. Copyright © 2007 Radar Networks & Nova Spivack. CC BY 2.0.

2.3 Data models

RDF is an abstract data model to represent web resources in the semantic web. It is designed to be read and most importantly understood by machines to provide better interoperability amongst computer applications. RDF structures data in the form of subject, predicate, and object, known as triples. The subject defines resource being described, the predicate is the property being described with respect to the resource, and the object is the value for the property. The subject and predicate are both represented by Uniform Resource Identifier (URI), while the object can be a URI or a string literal. An important feature of RDF is flexibility for resources to be a subject in one triple and an object in another one [14]. For instance, “Curtin University - Bentley campus is based in Perth city” and “Perth city is the largest city in Western Australia”. From a semantic web perspective, in this example, Perth city is an object in the first example and a subject in the second. This allows RDF to find connections between triples, which is referred to as linked data. RDF is used in combination with ontologies to provide semantic information about resources being described. Ontologies and vocabularies are discussed in the next section. The RDF data model is not only stored in an XML file format, there are some alternative existing RDF formats. Despite the difference in formats, resulting triples are logically equivalent. Some of the most-used RDF serialization formats are RDF/XML, N-Triples, JSON-LD, Turtle, and Notation3 (N3). For an explanation of RDF formats we recommend an article by Gandon, *et al.* [45].

2.4 Geospatial ontologies and geospatial semantic web query languages

The term "ontology" is originated from the field of philosophy, in which it refers to the subject of existence. In computer science and information science fields, an ontology is a data model, which is used to specify domain based concepts and relationships between these concepts [46]. There also exists a term called "vocabulary" in the semantic web, which is often referred to be equal to the term "ontology". According to W3C[6], "There is no clear division between what is referred to as 'vocabularies' and 'ontologies'. The trend is to use the word 'ontology' for more complex, and possibly quite formal collection of terms, whereas 'vocabulary' is used when such strict formalism is not necessarily used or only in a very loose sense." Ontologies and vocabularies in the semantic web provide a structure for organizing implicit and explicit concepts and relationships. There are several reasons to use an ontology in the semantic web and the geospatial semantic web such as to analyze domain knowledge, or to separate domain knowledge from operational knowledge [47]. However, the main reason to use ontologies in the semantic web and the geospatial semantic web is to improve data integration by tackling an issue of semantic heterogeneity[7].

Semantic heterogeneity is a term to describe a disagreement about the meaning and interpretation of data [48]. Ontologies can be one of three types, namely, domain ontology, and upper ontology (also known as top-level

6 https://www.w3.org/standards/semanticweb/ontology (last accessed on 9 April 2019)
7 https://www.w3.org/standards/semanticweb/ontology (last accessed on 9 April 2019)

ontology) and hybrid ontology. Domain ontology specifies concepts, terminologies, and thesauri that are associated within a certain domain. For instance, a geospatial ontology defines concepts and relationships for geospatial data. To be more precise, spatio-temporal concepts and geospatial relations such as equal, disjoint, touch, within, contain can be mapped to geospatial ontologies, which then can facilitate shareable and reusable geospatial information. Upper ontology specifies concepts which can be used across a wide range of domains. It represents concepts for general things such as an object, properties, space, roles, functions and relation, which can be found in many domains. Hybrid ontology is a combination of domain ontology and upper ontology [49]. In the next sub-sections, the article discusses geospatial ontologies and a cultural heritage domain ontology CIDOC-CRM in more detail. Furthermore, there are query languages termed SPARQL and GeoSPARQL, which are query languages for data in RDF in semantic web and the geospatial semantic web respectively. SPARQL is a standard by W3C, while GeoSPARQL is a standard by Open Geospatial Consortium (OGC). The next sub-sections cover those languages in detail as well.

2.4.1 GeoSPARQL vocabulary and query language

In recent years, geospatial organizations, research groups, triple store vendors have attempted to implement their own geospatial ontologies and strategies in order to deploy geospatial data in the semantic web. Unfortunately, this led to inconsistency. Hence, properly represented spatial RDF data in one organization became incompatible with spatial RDF data from another organization [50]. To resolve this, a standardized geospatial ontology and query language termed GeoSPARQL has been proposed by OGC. OGC is an international not for profit organization, which implements quality open standards for the global geospatial community. GeoSPARQL is a small vocabulary to describe geospatial information to supply for the geospatial semantic web as well as a geospatial extension of the semantic web query language of SPARQL. It can be combined with other domain-based ontologies to fulfill the needs for geospatial RDF data. GeoSPARQL implements a basic core for representing and accessing geospatial data published in a machine-readable and processable format in the geospatial semantic web [51]. Published RDF data can be queried with spatial functions provided by GeoSPARQL to discover interesting connections and relationships among cultural heritage resources.

An example of GeoSPARQL query is illustrated in Figure 4. This GeoSPARQL query finds all performing arts centre's — in this case, denoted as PAC — which are located within 30000 meters from the place "The Perth Mint" in the city of Perth, Western Australia. It uses the GeoSPARQL "geof:distance" function to calculate the distance between each performing arts centre and "The Perth Mint". In this example, it is assumed that URI for the data is located at http://example.com/.

```
PREFIX example: <http://example.com/>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX geof: <http://www.opengis.net/def/geosparql/function/>
SELECT ?p
WHERE {
      ?p a example: PAC ;
     geo:hasGeometry ?pgeo .
     ?pgeo geo:asWKT ?pWKT

     example:ThePerthMint .
     geo:hasGeometry ?lgeo .
     ?lgeo geo:asWKT ?lwkt .
FILTER (geof:distance (?pWKT, ?lwkt,
     units:m)< 20000)
}
```

Figure 4. Example of a GeoSPARQL query to calculate a distance between objects

2.4.2 CIDOC-CRM and CRMgeo

CIDOC-Conceptual Reference Model (CRM) is an ontology developed to define concepts and relationships in CH data. It has been developed by two working groups of CIDOC namely CIDOC Documentation and Standards Working Group and CIDOC CRM SIG. The ontology is not developed for CH institutions in any one country or a region, but any CH institution in the world can benefit from using it. In addition, it is an ISO standard since 2006. The main aim of the ontology is to provide semantic interoperability and integration among heterogenous CH resources. It offers more than 80 classes and 130 properties, which CH organizations can organize their datasets with and define relationships among them. It follows an object-oriented data model and offers the possibility for extensions [52].

As mentioned earlier, CH data is often accompanied by a spatio-temporal reference. Hence, there have been many projects which attempted to integrate geoinformation with CIDOC-CRM. For instance, AnnoMAD System integrated Geographic Markup Language (GML) with CIDOC-CRM to annotate spatial descriptions in free-text archeological data [53], CLAROS[8] project used basic spatial coordinates together with CIDOC-CRM. Recently a formal geospatial extension to CIDOC-CRM termed CRMgeo has been developed. CRMgeo has been integrated with the GeoSPARQL vocabulary to add spatio-temporal classes and properties to the ontology. It allows encoding location and time-related concepts and relationships of CH data under CIDOC-CRM. It adheres to formal definitions, encoding standards and topological relations used in the GeoSPARQL vocabulary, thus geospatial functionalities of GeoSPARQL can be fully utilized on CH data published under CRMgeo. The integration method of CIDOC-CRM with the GeoSPARQL vocabulary is given in Figure 5. The letters 'SP' denotes classes and 'E' represents properties. GeoSPARQL properties are highlighted in blue, where "Feature" is a subclass of "Spatial Object" and can have "Geometry" in one of two serialization formats, well-known text (WKT) or GML [54].

8 http://www.clarosnet.org/XDB/ASP/clarosHome/ (last accessed on 9 April 2019)

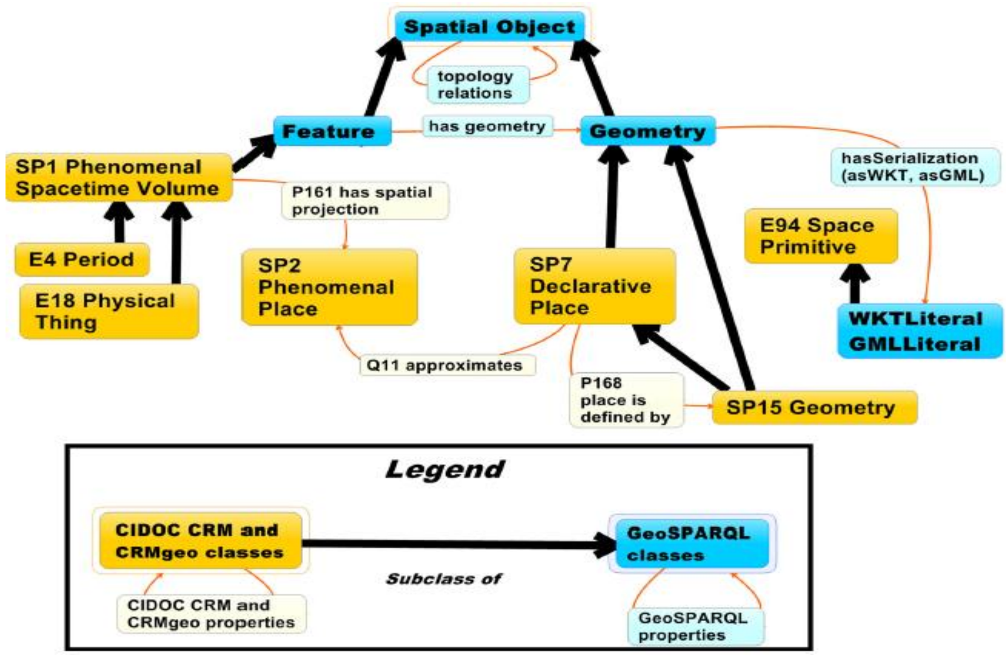


Figure 5. CRMgeo classes and their relationship to GeoSPARQL [54].
Copyright © 2016, The Authors. CC BY 4.0.

2.4.3 Geospatial functions in SPARQL query language

SPARQL Protocol and RDF Query Language is a query language for the semantic web. It is one of the key pieces of technology stack of the semantic web. SPARQL is not designed to query geospatial data and does not provide a comprehensive set of geospatial query capabilities. However, it supports several simple geospatial functionalities such as “st_intersects” (intersection between two geometries), “st_within” (check if geometry A is within geometry B) and others. If the case is to use simple geospatial queries on the semantic web, such as to check if a CH site is located within specified proximity or to check if geometries of two or more CH sites intersect, then SPARQL can handle them. The following example of a SPARQL query in Figure 6 demonstrates geospatial functionality of “st_intersects”, performed on a DBpedia semantic web endpoint9 at http://dbpedia.org/snorql/. DBpedia is a semantic web version of a Wikipedia project and allows the querying of Wikipedia content using SPARQL.

9 https://dbpedia.org/sparql (last accessed on 9 April 2019)

This query returns all places that lie within 10 kilometres of Perth city in Western Australia.

```
SELECT DISTINCT ?name ?URIResource ?coordinates
WHERE {
    <http://dbpedia.org/resource/Perth> geo:geometry ?sourcegeo.
    ?URIResource geo:geometry ?coordinates;
    rdfs:label ?name.
    FILTER( bif:st_intersects( ?coordinates, ?sourcegeo, 10 ) ).
    FILTER( lang( ?name) = "en" ) }
```

Figure 6. Example of a geospatial query in SPARQL.

The result of the query is shown in Table 1. The DBpedia semantic web endpoint provides results in JSON, XML and XML+XSLT formats.

Table 1. SPARQL geospatial query result in the DBpedia endpoint.

name	URIResource	coordinates
"Balcatta, Western Australia"@en	:Balcatta,_Western_Australia	"POINT(115.81700134277 31.875999450684)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Bassendean, Western Australia"@en	:Bassendean,_Western_Australia	"POINT(115.94899749756 31.905000686646)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Churchlands, Western Australia"@en	:Churchlands,_Western_Australia	"POINT(115.79000091553 31.92200088501)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Claremont, Western Australia"@en	:Claremont,_Western_Australia	"POINT(115.78199768066 31.979999542236)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Coolbinia, Western Australia"@en	:Coolbinia,_Western_Australia	"POINT(115.85399627686 31.913000106812)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Dianella, Western Australia"@en	:Dianella,_Western_Australia	"POINT(115.87400054932 31.888000488281)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Doubleview, Western Australia"@en	:Doubleview,_Western_Australia	"POINT(115.7799987793 31.895999908447)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Glendalough, Western Australia"@en	:Glendalough,_Western_Australia	"POINT(115.82399749756 31.917999267578)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Herdsman, Western Australia"@en	:Herdsman,_Western_Australia	"POINT(115.80400085449 31.920999526978)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Inglewood, Western Australia"@en	:Inglewood,_Western_Australia	"POINT(115.88300323486 31.916999816895)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Innaloo, Western Australia"@en	:Innaloo,_Western_Australia	"POINT(115.79100036621 31.893999099731)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Joondanna, Western Australia"@en	:Joondanna,_Western_Australia	"POINT(115.83599853516 31.909999847412)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Menora, Western Australia"@en	:Menora,_Western_Australia	"POINT(115.86100006104 31.916000366211)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Mirrabooka, Western Australia"@en	:Mirrabooka,_Western_Australia	"POINT(115.87000274658 31.863000869751)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Mount Lawley, Western Australia"@en	:Mount_Lawley,_Western_Australia	"POINT(115.87220001221 31.934499740601)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Nollamara"@en	:Nollamara	"POINT(115.84600067139 31.881999969482)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Osborne Park, Western Australia"@en	:Osborne_Park,_Western_Australia	"POINT(115.81199645996 31.898000717163)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Hazelmere, Western Australia"@en	:Hazelmere,_Western_Australia	"POINT(115.92199707031 31.900999069214)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Nedlands, Western Australia"@en	:Nedlands,_Western_Australia	"POINT(115.80699920654 31.982000350952)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"West Perth, Western Australia"@en	:West_Perth,_Western_Australia	"POINT(115.8430557251 31.950000762939)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Bedford, Western Australia"@en	:Bedford,_Western_Australia	"POINT(115.89099884033 31.909000396729)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Belmont, Western Australia"@en	:Belmont,_Western_Australia	"POINT(115.92700195312 31.944999694824)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Carlisle, Western Australia"@en	:Carlisle,_Western_Australia	"POINT(115.91600036621 31.976999282837)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Morley, Western Australia"@en	:Morley,_Western_Australia	"POINT(115.90699768066 31.886999130249)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Mount Pleasant, Western Australia"@en	:Mount_Pleasant,_Western_Australia	"POINT(115.84799957275 32.027000427246)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Perth (suburb)"@en	:Perth_(suburb)	"POINT(115.84999847412 31.950000762939)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Waterford, Western Australia"@en	:Waterford,_Western_Australia	"POINT(115.88400268555 32.016998291016)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Wilson, Western Australia"@en	:Wilson,_Western_Australia	"POINT(115.91100311279 32.026000976562)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Ashfield, Western Australia"@en	:Ashfield,_Western_Australia	"POINT(115.93800354004 31.916999816895)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Embleton, Western Australia"@en	:Embleton,_Western_Australia	"POINT(115.91000366211 31.903999328613)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"Jolimont, Western Australia"@en	:Jolimont,_Western_Australia	"POINT(115.80899810791 31.94700050354)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>
"South Perth, Western Australia"@en	:South_Perth,_Western_Australia	"POINT(115.8629989624 31.982000350952)"^^<http://www.openlinksw.com/schemas/virtrdf#Geometry>

2.5 Spatio-temporal triple stores

Currently there are two existing approaches to make RDF data available for use by the geospatial semantic web. The first approach is to use direct mapping languages that can construct RDF data from relational database management systems (RDBMS). It allows the accessing of data in RDBMS as if it were RDF data in a triple store. The RDF data in this case becomes virtual or read-only RDF data. The main advantage of using this approach is there is no need to duplicate data and store them in an actual triple store. There are some RDBMS which are shipped together with this RDBMS to RDF capability. For instance, databases such as Jena SDB, IBM DB2-RDF have this RDBMS to RDF capability. If a database does not support this capability, then additional systems can be used such as D2RQ, R2RML and Ultrawrap. The second approach is to use a triple store to store RDF data. A triple store is a database which allows storing and accessing RDF data. This article discusses the second approach in detail.

The semantic web stores RDF data in a different data model to SQL-based RDBMS or non-relational database systems (NoSQL) databases. In the semantic web an RDF database, a graph database, and a triple store are used interchangeably to refer to storage systems for RDF data. Currently, there are various existing triple stores that can store RDF data, such as Blazegraph[10], Apache Jena,[11] and others. However, for storing RDF data with spatio-temporal information, dedicated spatio-temporal triple stores should be used, such as Parliament, Oracle Spatial and Graph 12c, and Strabon. Hence, a feature matrix for a few of the well-known triple stores (see Table 2) is provided. It can be seen from Table 2 that all five spatio-temporal triple stores support GeoSPARQL query language capabilities, and GeoSPARQL geometry and topology extensions except OpenLink Virtuosa. It supports the GeoSPARQL query language. However, regards geometry it supports only point data and thus no support is available for lines and polygons. All of them can handle multiple CRS (Coordinate Reference Systems) but uSeekM manages only WGS_84 (World Geodetic System 1984).

Parliament is a triple store and rule engine compliant with RDF and OWL semantic web concepts, however, it does not have an integrated query processor to accept semantic web queries. Hence, it is usually paired with

---

10 https://www.blazegraph.com/ (last accessed on 9 April 2019)
11 https://jena.apache.org/ (last accessed on 9 April 2019)

Jena or Sesame to process queries and thanks to third-party integrations it has extra possibilities such as spatio-temporal support and numerical indexing [55]. Regards Parliament's limitations, it is difficult to configure for novice developers, and — at the time of writing — it does not support Sesame version 2 or later versions.

Oracle Spatial and Graph 12c is the most advanced triple store for enterprise-class deployments at the time of writing. It provides management features including graph database access and analysis operations, which can involve geospatial analyses as well through a Groovy-based shell console. Furthermore, it offers a considerable number of advanced performance accelerators and other enhancements for the semantic web, in particular, optimized schema to store and index RDF data, fast retrieval of RDF data based on Oracle-text and Apache Lucene technologies and RDB-to-RDF conversion among others [56]. A limitation of this triple store is its proprietary license, in addition to its complexity, which requires technical professionals to set-up and maintain.

uSeekM is not a triple store but an extension library for triple stores based on Sesame. This extension, when paired with Sesame based triple stores, supports OpenGIS Simple Features including Point, Line, Polygon and GIS relations such as Intersects, Overlaps, and Crosses.[12] One of the advantages of using this extension is it provides new modern types of indexing such as GeoHash and Quadtree, PostgisIndexer, which should result in better performance, particularly in searching and retrieval of spatio-temporal data in the geospatial semantic web. On the other hand, the limitation of uSeekM is that it does not support multiple CRS but only WGS_84.

OpenLink Virtuosa is an engine which incorporates a web server, RDF quad-store (i.e. graph, subject, object, and tuples), SPARQL processor, and OpenLink Data Spaces for a Linked Open Data based Collaboration Platform (i.e. application suite for a wiki, webmail, bookmarks, etc.) [57].

12 https://www.openhub.net/p/useekm (last accessed on 9 April 2019)

Strabon is an open-source RDF triple store based on the well-known Sesame RDF store. It extends the Sesame with a query processor engine and storage of their custom build stRDF (data model for the representation of spatio-temporal RDF data) and stSPARQL (an extension to SPARQL) query language similar to OGC GeoSPARQL to query stRDF data [58]. A major limitation of Strabon is that most OGC GeoSPARQL functionalities are not supported at the time this article was written.

Table 2. Feature matrix of spatio-temporal triple stores.

	Parliament (in the case paired with Jena)	**Oracle Spatial and Graph 12c Release 2**	**uSeekM**	**OpenLink Virtuosa 7.2.5.1**	**Strabon 3.0**
Core GeoSPARQL Support	Yes	Yes	Yes	Yes	Yes
GeoSPARQL Geometry Extension Support	Yes	Yes	Yes	Points only	Yes
GeoSPARQL Topology Extension Support	Yes	Yes	Yes	No	Yes
Multiple CRS	Yes	Yes	No (Only WGS_84)	Yes	Yes
Origin	Native built with C++ language	Native built	RDF4J (Sesame)	Native built	RDF4J (Sesame)
DBMS as a back-end	Berkeley DB	Oracle Database 12.2	PostgreSQL/PostGIS	Virtuoso DBMS	PostGIS, MonetDB
R-tree index	Yes	Yes	Yes	Yes	Yes
License	BSD Licence	Proprietary	Open Source, Apache Licence 2.0	GPL v2 or Commercial	Mozilla Public Licence
Other	Support for integration with Jena and Sesame. Supports different servlet engines Jetty, Tomcat, Glassfish or others.	Oracle Exadata supports fast computations on triples. Support for integration with Apache Lucene and	Support for integration with Sesame, also supports Quadtree, Geohash	The free version does not include all the functionalities.	Most Powerful Spatio-Temporal RDF Store. Has a custom built stRDF data model and stSPARQL query language

Developed by	BBN Technologies	Oracle Corporation	Open Sahara	OpenLink Software	European FP7 projects 'SemsorGrid4 Env' and 'TELEIOUS'
Official Website	http://parliament.semwebcentral.org/ [last accessed on 9 April 2019]	https://www.oracle.com/technetwork/database/options/spatialandgraph/overview/index.html [last accessed on 9 April 2019]	https://www.openhub.net/p/useekm [last accessed on 9 April 2019]	https://virtuoso.openlinksw.com/ [last accessed on 9 April 2019]	http://www.strabon.di.uoa.gr/ [last accessed on 9 April 2019]

2.6 Proposed framework for the geospatial semantic web

As mentioned before, CH data often accompanied by spatio-temporal references, which comes in various geospatial file formats. This section demonstrates a workflow to turn geospatial CH data in a vector file format into RDF data, with its subsequent storage and retrieval features, illustrated in Figure 7.

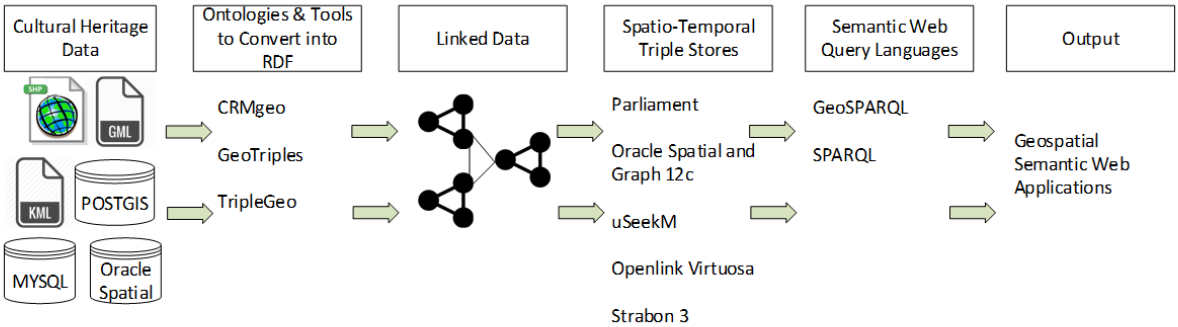


Figure 7. Workflow: from cultural heritage data to the geospatial semantic web

Geospatial CH data in a vector data model that is stored in spatio-temporal databases such as PostGIS and Oracle Spatial or in vector file formats such as ESRI Shapefiles, Keyhole Markup Language (KML), GML can be converted into RDF data using tools, as for example GeoTriples[13] and TripleGeo[14]. These tools extract geospatial features in vector files such as points, lines, polygons along with thematic attributes such as identifiers, names from the geospatial CH files and transform them into machine-readable and processable RDF data. Once the data is converted into RDF

13 http://geotriples.di.uoa.gr/ (last accessed on 9 April 2019)
14 https://github.com/GeoKnow/TripleGeo (last accessed on 9 April 2019)

data using these tools, it can be populated into a spatio-temporal triple store for storage and retrieval. A triple store provides an endpoint, which is designed to publish and access RDF data in addition to alter existing RDF data. Semantic web query languages such as GeoSPARQL and SPARQL allow querying RDF data from the endpoint and the resulting query can be used by geospatial semantic web applications.

The following paragraphs present this framework with a case study. The tool TripleGeo can encode data represented only in a vector data model, hence, it is the main criteria for data suitable for this framework. Also, the geometry, and thematic attributes such as ID, name are other important data characteristics to take into account, but temporality has not been considered in this framework. For this case study, data about statutorily protected places in Perth city in Western Australia is used, available under a Creative Commons license at the webpage of the Heritage Council of Western Australia[15].The geometry of the data is represented in a vector data model and stored in an ESRI Shapefile format. It contains 2017 polygons, each polygon hold thematic attributes such as FID (identifying number for each polygon), shape name (polygon), unique place number, place name, place location, length and area of for the shape, as shown in Figure 8. This data is then converted into RDF Turtle syntax using the above-mentioned tool TripleGeo. The RDF file contains all the polygons included in the original file, however, a sample from the RDF file that is the place named Peter Pan Statue, Queen's Garden, is illustrated in Figure 9. The prefixes in the RDF data are URI namespaces for GeoSPARQL ontology, for XML schema, for URI resource of the file (http://examplePerthWesternAustralia/URI#). A property "Georesource" in the RDF data represents an identity of the polygon, which is "Geom_polygons_2172". The geometry of the polygon is in RDF based on WKT, which is linked to WKT via the property "geo:asWKT".

Once the data is converted into RDF data, it is populated into a spatio-temporal triple store "Strabon", which offers an endpoint as shown in Figure 10. This endpoint then allows the accessing of RDF data stored in the triple

---

15 https://catalogue.data.wa.gov.au/dataset/heritage-council-wa-state-register (last accessed on 9 April 2019)

store. Figure 11 illustrates the query respond that contains the previously discussed Peter Pan Statue, Queen’s Garden, as a polygon in SPARQL/JSON format. This endpoint can accept queries and the result of the query can be used to build geospatial semantic web applications. Variables *s*, *p* and *o* denotes subject, predicate and object respectively.

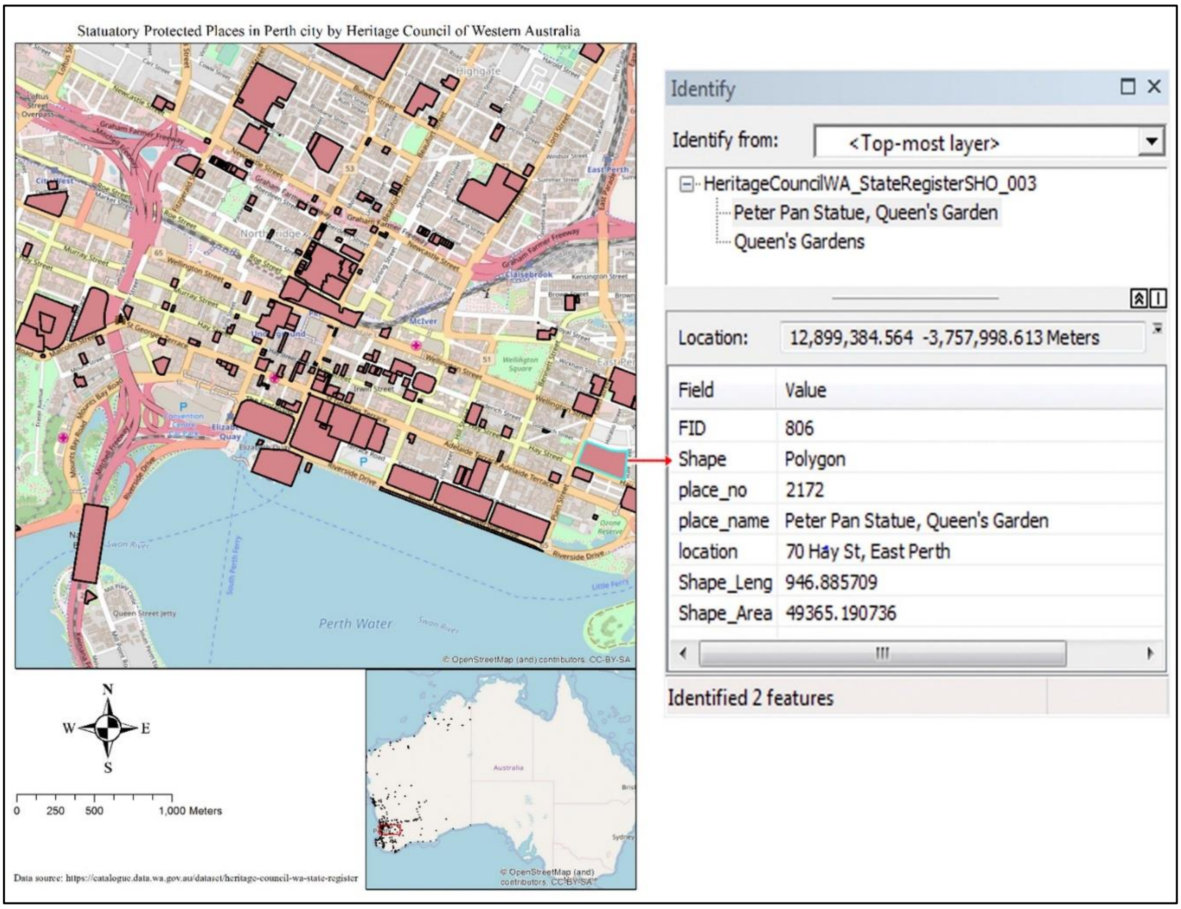


Figure 8. Statutory protected places in Perth city, Western Australia.
Source: https://catalogue.data.wa.gov.au/dataset/heritage-council-wa-state-register. Copyright © Bernhard Klingseisen, Landgate. CC BY 4.0.

```
@prefix geontology:  <http://www.opengis.net/ont/geosparql#>.
@prefix geo:         <http://www.opengis.net/ont/geosparql#>.
@prefix sf:          <http://www.opengis.net/ont/sf#>.
@prefix xsd:         <http://www.w3.org/2001/XMLSchema#>.
@prefix georesource: <http://examplePerthWesternAustralia/URI#>.
@prefix dc:          <http://purl.org/dc/terms/>.
georesource:Geom_polygons_2172
        a           sf:Polygon ;
        geo:asWKT "POLYGON ((12899179.4855 -3757986.4516000003, 12899179.689199999
-3757985.9180999994, 12899183.1587 -3757976.8027999997, 12899235.0636 -3757840.4525999986,
12899368.2546 -3757891.428199999, 12899384.6736 -3757896.8847999983, 12899401.395799998
-3757901.375500001, 12899418.327799998 -3757904.8726999983, 12899435.409200002
-3757907.3691000007, 12899452.582400002 -3757908.855700001, 12899469.7881 -3757909.3304000013,
12899486.967500001 -3757908.7989999987, 12899504.061299998 -3757907.2611000016,
12899486.690699998 -3758087.1761000007, 12899486.0904 -3758087.8982000016, 12899485.454300001
-3758088.583799999, 12899484.782499999 -3758089.2344999984, 12899484.0797 -3758089.844900001,
12899483.345600002 -3758090.4168, 12899482.5821 -3758090.944600001, 12899481.795200001
-3758091.430399999, 12899480.983399998 -3758091.8720000014, 12899480.156100001
-3758092.2661999986, 12899478.327300001 -3758092.9571, 12899476.467300002 -3758093.540100001,
12899474.5841 -3758094.0111000016, 12899472.6838 -3758094.3704999983, 12899470.7725
-3758094.6180000007, 12899468.856600001 -3758094.7536000013, 12899466.942200001
-3758094.7771999985, 12899465.0372 -3758094.6893000007, 12899463.175900001 -3758094.495000001,
12899461.460499998 -3758094.238499999, 12899459.760699999 -3758093.9048000015, 12899458.0812
-3758093.4943000004, 12899179.4855 -3757986.4516000003))"^^geo:wktLiteral.
georesource:polygons_2172
        a           georesource:polygons , georesource:Perth , geo:Feature ;
        <http://www.w3.org/2000/01/rdf-schema#label>
                    "2172"@en ;
        georesource:name "Peter Pan Statue, Queen's Garden"@en ;
        geo:hasGeometry georesource:Geom_polygons_2172.
```

Figure 9. Example of polygon representation of RDF in TURTLE syntax in GeoSPARQL

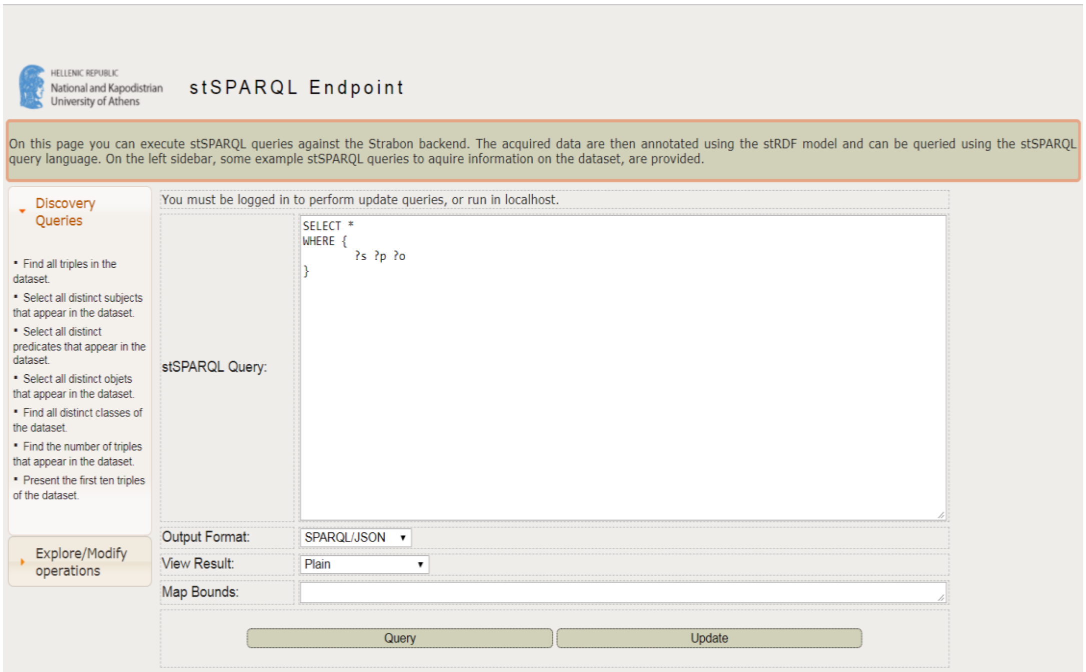
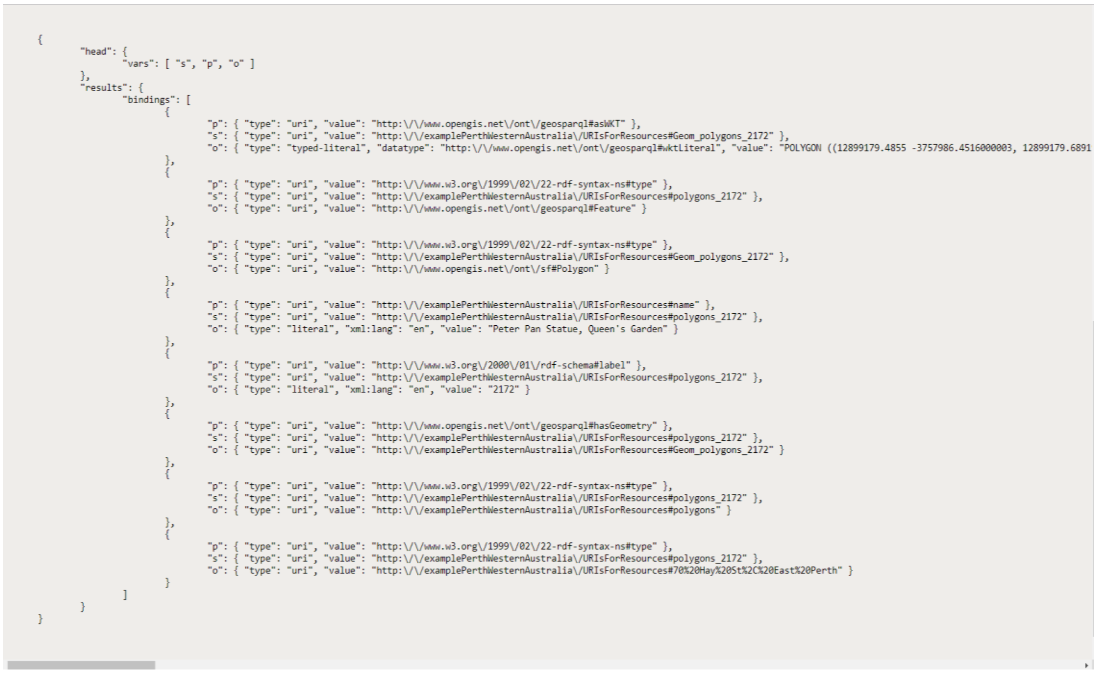


Figure 10. Endpoint of the Strabon triple store.
Copyright © 2017 KRR&A Team. Mozilla Public License Version 2.0.




```
{
    "head": {
        "vars": [ "s", "p", "o" ]
    },
    "results": {
        "bindings": [
            {
                "p": { "type": "uri", "value": "http:\/\/www.opengis.net\/ont\/geosparql#asWKT" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#Geom_polygons_2172" },
                "o": { "type": "typed-literal", "datatype": "http:\/\/www.opengis.net\/ont\/geosparql#wktLiteral", "value": "POLYGON ((12899179.4855 -3757986.4516000003, 12899179.6891
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.w3.org\/1999\/02\/22-rdf-syntax-ns#type" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "uri", "value": "http:\/\/www.opengis.net\/ont\/geosparql#Feature" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.w3.org\/1999\/02\/22-rdf-syntax-ns#type" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#Geom_polygons_2172" },
                "o": { "type": "uri", "value": "http:\/\/www.opengis.net\/ont\/sf#Polygon" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#name" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "literal", "xml:lang": "en", "value": "Peter Pan Statue, Queen's Garden" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.w3.org\/2000\/01\/rdf-schema#label" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "literal", "xml:lang": "en", "value": "2172" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.opengis.net\/ont\/geosparql#hasGeometry" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#Geom_polygons_2172" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.w3.org\/1999\/02\/22-rdf-syntax-ns#type" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons" }
            },
            {
                "p": { "type": "uri", "value": "http:\/\/www.w3.org\/1999\/02\/22-rdf-syntax-ns#type" },
                "s": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#polygons_2172" },
                "o": { "type": "uri", "value": "http:\/\/examplePerthWesternAustralia\/URIsForResources#70%20Hay%20St%2C%20East%20Perth" }
            }
        ]
    }
}
```


Figure 11. A query response of the endpoint.

2.7 Free and open-source semantic web and geospatial semantic web-based purpose-built platforms for cultural heritage institutions

In the past years, some FOSS purpose-built platforms for CH instructions have been developed, which incorporate semantic web or geospatial semantic web concepts. As these platforms are FOSS, they are completely free to use, and the source code is also available for free, which can be configured and extended without any restrictions to meet the needs of CH projects. Not only can this type of platforms offer unlimited flexibility, but also decrease software costs for CH institutions. Table 3 provides some information about these platforms such as supported semantic web and geospatial semantic web concepts, supported content types and visualizations. The following paragraphs discuss each one of them in more detail.

WissKI is a web-based virtual research environment (VRE) and content management system (CMS), which aims to become a Swiss army knife for scholars from any discipline who work with object-centric documentation. It has been developed by three scientific institutions namely "the Digital

Humanities Research Group of the Department of Computer Science at the Friedrich-Alexander-University of Erlangen-Nuremberg (FAU)”, “the Department of Museum Informatics at the Germanisches Nationalmuseum (GNM)” and “the Biodiversity Informatics Group at the Zoologisches Forschungsmuseum Alexander Koenig (ZFMK)”. It implements semantic web concepts to support scientific projects in CH institutions that collect, store, manage and communicate knowledge, and it provides long-term availability and interoperability of research outputs, as well as identity of authorship, and authenticity of information. People with a limited technical background can benefit from this easy-to-use platform to create enriched content for the World Wide Web, including for the semantic web and geospatial semantic web, presented in a Wikipedia-like style incorporating textual, visual and structured information. It supports two ways of creating content on the platform. The first way is to enter data using traditional form-based interfaces and the second is to perform an aggregation of data in free texts by utilizing web-based knowledge management approaches. In the latter case, a parser in the system recognizes named entities such as date and time phrases, names, and place names, and enters them to the system. Users of this platform can perform semantic text annotations by submitting free text, which is then analyzed to connect entities (e.g., names, places, dates) to the system’s knowledge base. WissKI is a modular extension of the well-known Drupal CMS. It is shipped with an ontology Erlangen CRM, which is an OWL-DL 1.0 (Web Ontology Language) implementation of CIDOC-CRM, however, any other ontology can also be imported to the system. A triple store called ARC[16] is used as a storage for RDF, which offers a SPARQL endpoint to query available data in the knowledge base [59].

Arches is a heritage inventory and management system developed through a collaboration between the Getty Conservation Institute and the World Monuments Fund. It is a geospatially-enabled platform for the CH field to document immovable heritage such as CH buildings, CH landscapes, and archeological sites. The goals of the platform are to offer long-term

16 https://github.com/semsol/arc2/wiki (last accessed on 9 April 2019)

preservation and interoperability for CH datasets. It employs CIDOC-CRM as an ontology to define concepts and relationships within datasets. This relationships in the system can be visualized using an interactive and exploration software called Gephi[17]. International Core Data Standards for Archeological and Architectural Heritage (CDS) are used to define data fields in the platform. These fields can then be mapped to entity classes in CIDOC-CRM. Furthermore, it follows OGC standards which means the platform is compatible with desktop GIS applications. GIS features of the platform include a map-based visualization with comprehensive spatial queries, as well as drawing, importing and editing CH resource geometries. Results derived from the spatial query and editing of CH resource geometries can be exported into GIS formats (e.g., shapefile), and vocabulary concepts can be exported into the SKOS format. Another interesting feature of the Arches platform is its mobile app called "Arches Collector[18]", which allows the collecting of heritage data in the field. The collected data is then synchronized with the installed Arches platform. Through use of this feature, administrators of the platform can design projects by setting up participants, type of data to be collected, place of the data collection, and collection time. Approved users can then connect to the Arches platform, download the project, collect the data and save it to the Arches platform. Arches is developed using a Python-based Django[19] web framework together with other web frameworks (Require.js, Backbone.js, jQuery, Bootstrap). Arches platform is fully customizable, which means other controlled vocabularies including the semantic web, and related geospatial semantic web standards can be implemented on top of the existing platform [60].

ResearchSpace is a web-based platform that offers a collaborative environment for humanities and CH research projects. It has been developed by the British Museum and utilizes semantic web concepts. It includes work to integrate heterogeneous datasets without losing meaning or perspective. This platform uses CIDOC-CRM to define concepts and relationships in the

17 https://gephi.org/ (last accessed on 9 April 2019)
18 https://arches.readthedocs.io/en/stable/using-arches-collector/ (last accessed on 9 April 2019)
19 https://www.djangoproject.com/ (last accessed on 9 April 2019)

datasets. The project also aims to implement a contextual search system and other tools to enhance collaboration in projects. It offers several types of instruments to explore and visualize complex datasets. For instance, “Narratives” – this tool allows creating documents incorporating semantically defined entities and interactive visualizations. “Semantic Diagram” is another tool to explore connections between objects, people, events, and places in the system. An image annotation tool is another interesting feature of the platform, which allows the attaching of text information to a selected area of the image. ResearchSpace is based on “metaphacts”[20], which is a knowledge graph platform.[21]

Omeka is a content management system developed jointly by Corporation for Digital Scholarship, the Roy Rosenzweig Center for History and New Media, and George Mason University. According to Cohen [61] who introduced Omeka, “Omeka, [derives]from the Swahili word meaning ‘to display or layout goods or wares; to speak out; to spread out; to unpack.’”. It provides a web platform for CH institutions to publish and manage digital CH collections. Omeka offers three options for using this platform. The first option is “Omeka S”, which is a web-based platform for publishing CH data, and for interlinking datasets with other online resources using semantic web concepts. The second option “Omeka Classic” offers a web publishing platform but without the use of semantic web concepts. “Omeka.net” is the third option, which provides a web publishing platform without involving semantic web concepts in their hosted service. In this article, “Omeka S” is discussed as it includes publishing data for the use in the semantic web and the geospatial semantic web. A feature distinguishing “Omeka S” from other platforms that otherwise similar, is it allows the management of multiple sites from one single installation. This could be helpful for organizations needing to run multiple sites and manage them from a single platform. Not only can “Omeka S” provide an easy-to-use platform, but it also provides a possibility to extend the functionality with modules. It is shipped with four pre-built ontologies namely Bibliographic Ontology, Dublin Core, Dublin Core Type, and Friend of

20 https://metaphacts.com/ (last accessed on 9 April 2019)
21 https://www.researchspace.org/ (last accessed on 9 April 2019)

a Friend. In addition, other ontologies can also be imported to the platform. In regard to visualization methods, it supports map, image, and 3D model visualizations. Omeka is written in a PHP programming language and requires a LAMP web service stack to run, which are developed via Linux, Apache HTTP Server, MYSQL and PHP.[22]

Fedora (Flexible Extensible Digital Object Repository Architecture) is a robust and modular repository system, which includes integration with semantic web concepts. A number of universities, research institutions, government agencies and CH organizations around the globe have contributed to the development of the framework. It is currently led by the Fedora Leadership Group. Fedora is fully customizable and is used by CH institutions, universities, research institutions to store and manage digital content. Nevertheless, it is not a complete application which can be used from uploading data to the presentation of digital content. It is a framework upon which institutions can build their platforms. Middleware solutions such as Samvera[23], Islandora[24] can also be used on top of it. Fedora supports defining relationships using ontologies and expressing them in RDF. It supports integration with external triple stores so resulting RDF triples can be stored in a plugged-in triple store. Fedora is written in a Java programming language and supports MYSQL and PostgreSQL databases. Files in Fedora can be stored in a local storage or in cloud storage solutions such as Amazon S3 [62].

[22] https://omeka.org/s/docs/user-manual/ (last accessed on 9 April 2019)
[23] https://samvera.org/ (last accessed on 9 April 2019)
[24] https://islandora.ca/ (last accessed on 9 April 2019)

Table 3. FOSS Semantic Web and the Geospatial Semantic web-based purpose-built platforms for cultural heritage institutions.

Name	Developed by	Originating Software	Built-in CIDOC-CRM Support	Supported Semantic Web and Geospatial Semantic Web Standards	Supported Content Types and Visualizations	Software Dependencies	Official Website
Wisski	The Digital Humanities Research Group of the Department of Computer Science at the Friedrich-Alexander-University of Erlangen-Nuremberg (FAU), the Department of Museum Informatics at the Germanisches Nationalmuseum (GNM) and the Biodiversity Informatics Group at the Zoologisches Forschungsmuseum Alexander Koenig (ZFMK)	Drupal – Open Source Content Management System	Yes / Erlangen CRM	Semantic Text Annotations Can load nearly any ontology Uses a triple store – ARC (https://github.com/semsol/arc2/wiki) SPARQL	Text, Images, Time bars, maps, Graph-based visualizations are under development	Apache 2.X + Mod Rewrite (for clean URL support) MySQL 5.5.3 or higher PHP Version 5.5.9 or higher Drupal 8.X	http://wiss-ki.eu/
Arches v4	Getty Conservation Institute and World Monuments Fund	Python Django Framework Require.js Backbone.js jQuery Bootstrap	Yes	Extendable other standards can be implemented	Text, Images, Maps Time wheel – a searchable graphical distribution of time data Graph – network visualization based on Gephi Compatible with satellite imagery and online map services Arches Collector – a mobile app to collect heritage data	Operating System: Linux Ubuntu or Windows PostgreSQL 9.6 with PostGIS 2.3 GDAL and GEOS Python 2.7 Yarn Mapnik 2.2 JDK ElasticSearch 5.2.1 CouchDB 2.x	https://www.archesproject.org

Table 3. Continued.

Name	Developed by	Originating Software	Built-in CIDOC-CRM Support	Supported Semantic Web and Geospatial Semantic Web Standards	Supported Content Types and Visualizations	Software Dependencies	Official Website
ResearchSpace	The British Museum	Metaphacts (https://metaphacts.com/)	Yes	Extendable other standards can be implemented	Text, Images, Narratives Image annotations Image comparisons – overlay and side-by-side Network Visualizations with diagrams	Operating System: Linux Ubuntu 16.04 or MacOS Sierra Git Docker Java 8 Scala Interactive Build Tool Node.js Yarn	https://www.researchspace.org/ https://github.com/researchspace/researchspace
Omeka S	Roy Rosenzweig Center for History and New Media at George Mason University	PHP Zend Framework 3 Doctrine 2 EasyRDF PHPUnit jQuery	No - CIDOC-CRM can be imported separately	Any ontology / vocabulary can be imported	Text, Images, 3D files, Maps 3D Model Visualization	Operating System(s): Linux Apache HTTP Server MYSQL PHP	https://omeka.org/
Fedora 5.0.2	DuraSpace	Java	No – CIDOC-CRM and any other ontologies/vocabularies can be imported separately	Triple Stores for RDF SPARQL	Any type of file format	Operating System(s): Window, Linux, Mac OSX JAVA 8, Servlet 3.0 MYSQL and PostgreSQL are also supported if needed Files can be stored in a local file storage or cloud-based solutions such as Amazon S3	https://duraspace.org/fedora/

2.8 Applications of the geospatial semantic web in cultural heritage domain

Linked CH data can be accessed from various sources. These sources can be divided into three categories, namely data aggregating platforms (e.g., Europeana, ARIADNE), metadata connectivity platforms (e.g., Pelagios), Gazetteers or APIs (e.g., GeoNames). In the following sections, we discuss key projects related to those categories that employed geospatial semantic web concepts.

2.8.1 Pelagios commons

One of the first steps in applying the geospatial semantic web in the CH domain or a project in a closely aligned direction is Pelagios, run by a consortium of Pelagios Commons, funded by The Andrew W. Mellon Foundation. The aims of the project are as follows:

- Create links among ancient places and help users to reference their ancient or historical geo-related data.
- Make links more discoverable and visualize them in a meaningful way.

They achieve the above aims by providing a web-based infrastructure which allows the Recogito tool to annotate text against images (i.e., ancient maps, images in digital books). Pleiades Gazetteer of the Ancient World can retrieve referenced places of the ancient world as URIs (Uniform Resource Identifier). The storage of the data in the system is in RDF format and the vocabulary used to describe places is the Open Annotation Data Model[25]. Furthermore, VoID (Vocabulary of Interlinked Datasets)[26] is used to describe general metadata about places including a general description, details of a publisher and license information. The infrastructure does not store data being referenced as it is not a data repository nor is it a data aggregator platform, rather it stores metadata about places and refers to URIs of places. As for visualization of referenced ancient places, Pelagios offers map-based visualization where users can search for ancient places that are of interest to them and examine the network of visualization of ancient places, and extract relevant data. Pelagios offers an HTTP API that allows other third-party applications to send a request and consume raw data from the API in different RDF formats. This provides an opportunity to take Linked Open Data from Pelagios and deploy it in other projects or develop mashups and applications [63].

---

25 http://www.openannotation.org/ (last accessed on 9 April 2019)
26 https://www.w3.org/TR/void/ (last accessed on 9 April 2019)

2.8.2 ARIADNE (now ARIADNE PLUS)

ARIADNE (Advanced Research Infrastructure for Archeological Dataset Networking in Europe) is a project of the European Union (EU) funded by the Seventh Framework Programme of the European Commission. It was built upon a consortium of 24 partners in EU as well as 15 associate partners. The main objective of the project was to develop an aggregator infrastructure that facilitates data connection, sharing and searching among European archeological institutions, hence achieve better interoperability across disperse archeological data collections. To achieve it they developed ARIADNE Catalogue Data Model as archeological institutions in Europe store collections of data in different formats, languages, and metadata schemas.

This model stores metadata information about archeological data and follows a “who, what, where, when” paradigm. To describe and encode information (e.g., monuments, pottery, and excavations) related to the “what” pattern, they used Art and Architecture Thesaurus (AAT)[27] of the Getty Research Institute. For the “where” pattern the spatial coordinates (longitude, latitude) in WGS_84 CRS have been used, whereas the information and a schema for encoding the information related to a “when” pattern have been facilitated in collaboration with a PeroidO[28] temporal gazetteer of historical periods. The portal itself has been developed using a PHP-based Laravel MVC (Model-Viewer-Controller) frameworks and conceptual classes for archeological data have been implemented by CIDOC-CRM, and the above mentioned “who, where, what, when” paradigm information mapped into DCAT (Data Catalog Vocabulary)[29] vocabulary [64].

2.8.3 PARTHENOS

PARTHENOS (Pooling Activities, Resources, and Tools for Heritage E-research Networking, Optimization, and Synergies) is an EU project funded by the EU Horizon 2020 program. The main aim of the project is to build a framework, which would work as a bridge between existing EU digital

27 http://www.getty.edu/research/tools/vocabularies/aat (last accessed on 9 April 2019)
28 http://perio.do/en/ (last accessed on 9 April 2019)
29 https://www.w3.org/TR/vocab-dcat/ (last accessed on 9 April 2019)

humanities aggregator infrastructures namely ARIADNE, CENDARI, CLARIN, CulturalItalia, DARIAH, EHRI, and TGIR. The partners argue that this framework eventually will facilitate an environment where humanists can find and use available humanities data resources including: download and process the data, share digital tools to re-use data, build dynamic virtual environments to find specific resources relevant to their research area, run different computational services on the data, and share the results between collaborators. To accomplish this, they have been developing a framework called the Content Cloud Framework, comprising a common data model that describes available data, services, and tools of all involving infrastructures but in a standardized manner. The data and tools are made available via user interfaces and semantic endpoints of SPARQL [65].

2.8.4 Geospatially linked data in digital gazetteers

Digital gazetteers can be loosely defined as a geodatabase for place names or toponyms. They usually include attributes and features for places such as a name of the place, location details (coordinates representing point, line or polygon), type of place (region, country, etc.) and others [66]. Digital gazetteers have paramount importance in providing access to location information that is used in research projects [67], geographic context retrieval systems [68], and location-based services, to name but a few. With the advancement of the geospatial semantic web technologies, new digital geospatial semantic gazetteers have been evolving and existing non-semantic ones have been transforming into geospatial semantic ones. Geospatial semantic web gazetteers are rich sources of geo-linked open data that are readily available for consumption in the geospatial semantic web and related projects. Examples of gazetteers available as geo-linked open data include the Getty Thesaurus of Geographic Names (TGN), GeoNames, the Pleaidas Gazetteer of Ancient World and so on.

TGN is based on the Getty vocabulary and includes geo-information (location, relationships, bibliography) related to historical places and, most importantly, has a focus on places relevant to CH, art, and humanities. Despite the fact that it does not provide information on the map interface, the results returned from the queries can be visualized on top of maps using other open source mapping libraries [69].

GeoNames gazetteer is a world geographical database containing location information of all countries in different languages, and users can add new place names and edit existing ones. GeoNames have developed their own ontology to represent geospatial data in RDF to the geospatial semantic web, and there is a semantic query endpoint available that can return queries in RDF format.[30]

In fact, all geospatial semantic digital gazetteers provide access to their linked data via a semantic query endpoint, which accepts geospatial semantic web queries, process and returns a response. The notable advantage of the semantic and the geospatial semantic web regarding querying is their ability to send federated queries. This can be explained as dividing a query into subqueries and sending to semantic query endpoints. This allows the easy integration of linked data from different semantic query endpoints, in a standardized data model (RDF) [70]. This is one of the powerful features of the semantic web and geospatial semantic web, and it facilitates an easier way to build applications with a suitable combination of heterogeneous data from various resources.

2.9 Technical limitations of the geospatial semantic web

As discussed in the previous sections, in recent years the geospatial semantic web has matured a great deal and has benefited from many advancements developed by large-scale research projects, as well as benefitting from smaller individual explorations. However, since it is still a new technology, there are many technical challenges that need to be tackled to take full advantage of it and successfully implement it across different disciplines, in a similar fashion to current GIS systems. Some of the global geospatial semantic web challenges include raster data representation and query, 3D data representation and query, as well as big geospatial data.

30 https://www.geonames.org/ (last accessed on 9 April 2019)

There are also many other technical challenges exist such as disambiguation, a plethora of data models to choose from for encoding geospatial data, entity matching, and others. Nevertheless, this article only concentrates on major obstacles hindering the geospatial semantic web from being applied to as many domains as GIS systems are applied to.

2.9.1 Raster data

Raster data is a data model for representing spatial phenomena, that is to say, it is used for representing various types of geospatial data such as continuous data (e.g. elevation, precipitation, temperature), thematic data with classifications (e.g. land use with different classifications such as vegetation, soil) [71]. The sources of raster data include images coming from satellites, UAV (unmanned aerial vehicles), scanned maps, and many others. In the CH domain, raster data can be employed in many case-studies, such as analyzing, visualizing, monitoring CH sites [72,73]. However, at the time of writing this research article, raster data representation is not supported in the geospatial semantic web, more specifically RDF data models GeoSPARQL, stRDF and others do not support raster data. Hence, there is no consistent way of encoding raster data for the geospatial semantic web. The main reason for that is raster data consists of a matrix of cells each containing a value or values that represents information (e.g., spectral value, category, magnitude, and height). Therefore, converting raster files into RDF would increase the size of the triples substantially or it might even require a few triples to encode a raster file, which in the end could create performance issues [74,75]. Nevertheless, there have already been several attempts research projects attempting to solve this issue. For example, there is a methodology that builds vector objects using image segmentation and then creates RDF for vector objects [76]. This workaround is arguably not a complete solution to represent raster files in RDF as it still does not represent raster geometry in RDF.

2.9.2 3D geospatial semantic web

In recent years there has been an ever-increasing interest in advancing 3D modelling technologies that can create a virtual replica of CH objects, display

3D CH models in web browsers, and in the enrichment of 3D CH models with related contextual information using annotation concepts and the semantic web [77,78]. In the CH domain, there have been numerous large-scale projects aligned with the above-indicated points such as INCEPTION (Inclusive CH in Europe through 3D Semantic Modelling) [79], 3D-COFORM [80] and others. Furthermore, many CH institutions now have a repository of 3D models that are available to online users for exploration [24,81].

However, the 3D geospatial semantic web remains an active research area, in particular, representation and 3D model encoding in RDF, and querying 3D geospatial semantic web are challenges still to be tackled. Consequently, a 3D geospatial semantic web, yet to emerge, should allow the geo-interlinking of 3D CH models with relevant valuable information, provide new possibilities and insights, and answer research questions in the CH domain and beyond. An example of initial steps in the 3D geospatial semantic web is a project by Hor*, et al.* [82], that has the objective to develop an RDF data model that semantically integrates CityGML-3D data encoding standard in GIS and BIM (Building Information Model), and query the data model using SPARQL query language. They claim that the new model has all the classes and properties of both CityGML and BIM, although they have not dealt with spatio-temporal aspects of the geospatial semantic web (which is the main characteristic of web-based geospatial data). Therefore, more research is needed to define consistent ways to encode, interlink, query, process and visualize structured 3D data with the 3D geospatial semantic web.

2.9.3 Big geospatial data and the geospatial semantic web

In recent years, the term “big data” has become ubiquitous in many domains including in CH. To this end, an array of platforms has been developed that deals with big CH data [83,84]. However, the term big data is cumbersome to contextualize and to describe how much data exactly it refers to as it is a relative term and can change quickly over time [85]. Nevertheless, five main characteristics of big data can be defined: high volume; high variety; high velocity; high veracity (i.e. varying quality); and value [86], in the case of geospatial big data, it is a large volume of geospatial data, in addition to the other four factors.

Even though big data concepts have been applied to many CH projects [83,84], big geospatial data is still a new topic to explore and apply in the CH domain. Geospatial big data has many worthwhile characteristics similar to big data, however, heterogeneity and complexity of data are the primary challenges here. Thus, processing and analyzing various types of geospatial data and an enormous volume of distributed data should be further researched as present semantic reasoners and triple stores are not yet able to handle big geospatial data. On the other hand, the geospatial semantic web may also be a solution for dealing with the challenge of heterogeneity in big spatial data. For instance, Koubarakis*, et al.* [87] hold a view that searching, integration and using big spatial data in other applications could be significantly improved if they are published using geospatial semantic web concepts. As a proof of concept, they generated a methodology showing clear benefits in a case-study of wild-fire-monitoring service. Although big geospatial linked data is still very much in its infancy, in the future it may provide new opportunities and may be applied in a relevant and effective way in the CH domain to solve many issues of interoperability, discovery, and processing of heterogeneous big CH data.

2.10 Conclusion

Currently in many CH repositories data is published as raw dumps in different file formats lacking structure and semantics. Some of the biggest technical challenges in these repositories are data integration and interoperability among distributed repositories at the national and regional level. Poorly linked CH data is fragmented in several national and regional repositories, backed by non-standardized search interfaces. All these technical challenges are limiting the capabilities of users to contextualize information from distributed repositories. The geospatial semantic web is a new paradigm of publishing, interconnecting, and consuming data on the Web, and it is being adopted by many CH research projects and CH repositories as a solution. Furthermore, in recent years a few FOSS semantic web and geospatial semantic web purpose-built platforms have been developed for CH institutions in order to ease the adoption of this technological progress.

This article provided a conceptual survey of the geospatial semantic web for a CH audience. Firstly, it discussed the state-of-the-art geospatial semantic web concepts pertinent to the cultural heritage field including standardized ontologies by OGC – GeoSPARQL, and ISO – CIDOC-CRM. Also, it discussed a geospatial extension to CIDOC-CRM ontology CRMgeo. Further, it compared the technical features of widely used geospatial triple stores to identify their advantages, benefits, disadvantages, and pitfalls. Secondly, it proposed a framework to turn geospatial cultural heritage data stored in a vector data model into machine-readable and processable RDF data to use in the geospatial semantic web (with a case study to demonstrate its applicability). Thirdly, it outlined key FOSS semantic web and geospatial semantic web-based purpose-built platforms for CH institutions. Next, it summarized leading CH projects that employed the geospatial semantic web concepts. Finally, it provided attributes of the geospatial semantic web requiring more attention, which may generate new ideas and research questions for both the field of the geospatial semantic web and CH.

We strongly suggest that employing standardized geospatial semantic web concepts will allow the creation of heterogeneous, multi-format, disperse cultural heritage data that will prove to be more accessible, interoperable and reusable than ever before. However, raster data representation and query, 3D data representation and query, and big geospatial data, are some of the challenges of the geospatial semantic web, requiring both more thoughtful attention and more results-oriented research projects.

Supplementary Materials: The following are available online at http://www.mdpi.com/2571-9408/2/2/93/s1 - Statutory protected places in RDF Turtle syntax based on GeoSPARQL ontology.

Chapter 3 A Comparative Evaluation of Geospatial Semantic Web Frameworks for Cultural Heritage[31]

This chapter provides a comparative evaluation of RDF generation tools and interlinking frameworks. Furthermore, a complete methodology for converting geospatial cultural heritage data into linked open data was proposed using CIDOC-CRM ontology, and the knowledge base of DBpedia to which the resulting RDF data was interlinked. This methodology is based on the previous methodology in Chapter 2. The difference between them is the methodology in Chapter 2 does not deal with mapping cultural heritage data into an ontology such as CIDOC-CRM and does not demonstrate the interlinking process to knowledge bases such as DBpedia. Chapter 3 concludes with the discussion on major limitations of the evaluated tools and frameworks which might be informative to cultural heritage researchers and professionals planning to use these tools and frameworks.

Abstract: Recently many RDF data generation tools have been developed to convert geospatial and non-geospatial data to 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 the tools and frameworks. In addition, the article provides a methodology to generate geospatial cultural heritage RDF data and to

31 This chapter is a copy of the peer-reviewed research article - Nishanbaev, I.; Champion, E.; McMeekin, D.A. A Comparative Evaluation of Geospatial Semantic Web Frameworks for Cultural Heritage. Heritage 2020, 3, 875–890. https://doi.org/10.3390/heritage3030048

interlink it with the related RDF data. This methodology uses a CIDOC-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.

3.1 Introduction

In recent years, Geographic Information Systems (GIS) has become a popular technology for cultural heritage (CH) researchers and professionals. One reason for this is the enormous possibility of GIS such as the 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 [88,89].

On the other hand, GIS is also undergoing marked changes in its 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 is 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 a such way that its meaning is well-defined to machines, and the data is interlinked with other related data [14]. Whereas 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 [29,30].

The CH domain has successfully implemented some Geospatial Semantic Web concepts. For instance, Hiebel, Doerr and Eide [54] developed a geospatial ontology called CRMgeo that provides a schema to integrate geospatial ontology GeoSPARQL with CIDOC-CRM, one of the most widely used cultural heritage ontologies. CRMgeo offers the necessary classes and relations to model spatio-temporal aspects of cultural heritage objects in the CIDOC-CRM ontology [54]. Another example is a WarSampo – Finnish Second World War on the Semantic Web project by Hyvönen, *et al.* [90] 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 [90]. The SKOS32 is an application of the 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 the recent research findings by Schmachtenberg, Bizer and Paulheim [16], who analyzed the adoption of LOD best practices including interlinking in different domains such as media and life sciences, 44 % of the published LOD datasets are actually 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 have the objective to conduct performance benchmarking of the tools, neither does it consider the tools that require computer programming such as RDFLib33, ontospy34, or similar. Instead, the geospatial RDF generation tools and interlinking

32 https://www.w3.org/TR/swbp-skos-core-spec/ (last accessed on 8 June 2020)
33 https://github.com/RDFLib/rdflib (last accessed on 8 June 2020)
34 https://pypi.org/project/ontospy/ (last accessed on 8 June 2020)

frameworks are selected based on the 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. In order 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. This data is 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, it 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 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.

3.2 Background literature

3.2.1 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 [88,91], GIS for analysis and visualization of CH [7,92] 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 [14]. Some fundamental concepts of the Semantic Web are an RDF data model, the RDF Vocabulary Definition Language (RDFS) and 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 relationships between these entities [14,46]. SPARQL Protocol and RDF Query Language (SPARQL)[35] 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)36 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 has distinct features such as geometry, 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

35 https://www.w3.org/TR/sparql11-query/ (last accessed on 8 June 2020)
36 https://www.w3.org/standards/semanticweb/ (last accessed on 8 June 2020)

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 [50,93]. It is a standard created by the OGC37-non-profit organization that implements open standards for the global geospatial community. Strabon[38] and Parliament[39] (when paired with Apache Jena[40]) are some of the well-known examples of geospatial triple stores that store geospatial RDF data, providing endpoints for sending GeoSPARQL queries.

3.2.2 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 [94]. 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 [54]. 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 [33,63], WarSampo [90], Omeka[41], and Getty Thesaurus of Geographic Names (TGN)[42] among others, have all successfully implemented geospatial semantic web concepts.

---

37 https://www.ogc.org/ (last accessed on 8 June 2020)
38 http://www.strabon.di.uoa.gr/ (last accessed on 8 June 2020)
39 https://github.com/SemWebCentral/parliament (last accessed on 8 June 2020)
40 https://jena.apache.org/ (last accessed on 8 June 2020)
41 https://omeka.org/s/ (last accessed on 8 June 2020)
42 https://www.getty.edu/research/tools/vocabularies/tgn/ (last accessed on 8 June 2020)

3.3 Methodology

3.3.1 Criteria for selecting RDF generation tools

In the last two decades, many RDF generation tools have been developed, 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 12 has been developed that facilitated the selection of the RDF generation tools. The first process in the diagram is to identify potential RDF generation tools, which are found from three different resources. The first resource is the official "RDF Generation Tools[43]" 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.

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

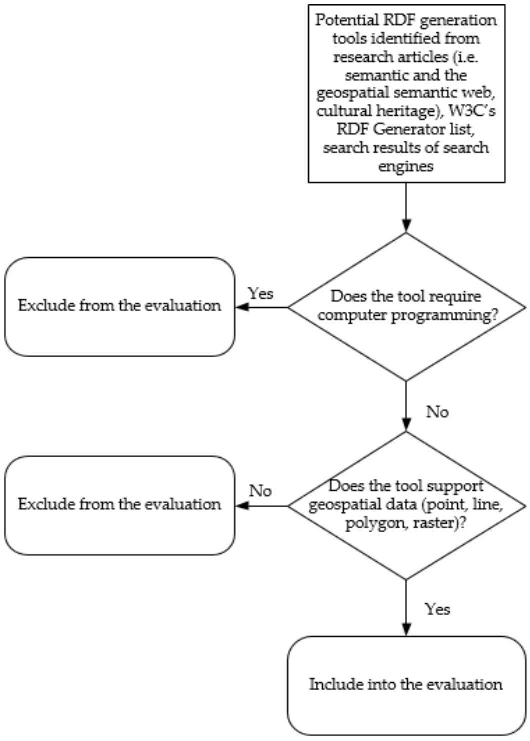


Figure 12. A flowchart diagram for selecting geospatial RDF generation tools

3.3.2 Criteria for selecting RDF linking frameworks

Generally, linking RDF data in many interlinking frameworks includes 3 steps, pre-processing, linking configuration, and postprocessing. In the first step, the data to be interlinked is cleaned and transformed as in many cases datasets may contain irregular characters. Afterwards, 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 [95]. The owl:sameAs[44] property indicates that two URI references refer to the same thing. In the article, we have compared

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

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+[45], SERIMI[46], KNOFUSS[47] 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.

3.4 Comparative evaluation

3.4.1 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 12. Firstly, a comparative evaluation of RDF data generation tools is provided. Afterwards, a detailed discussion on 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 4) 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, and ontology import option and others. It can be seen from Table 4 that KARMA Data Integration Tool can take as an input data geospatial database PostGIS and JSON file format which enable 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, 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.

45 http://ri-www.nii.ac.jp/SLINT/index.html (last accessed on 8 June 2020)
46 https://github.com/samuraraujo/SERIMI-RDF-Interlinking (last accessed on 8 June 2020)
47 https://code.google.com/archive/p/knofuss/downloads (last accessed on 8 June 2020)

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.

One remarkable tool for generating geospatial RDF data is Karma[48], which is free and open-source software (Apache Licence 2.0) and enables users to produce RDF data from several data sources such 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 [96]. 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-CRM[49]. 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 13. The mapping generator

---

48 https://usc-isi-i2.github.io/karma/ (last accessed on 8 June 2020)
49 http://www.cidoc-crm.org/Version/version-5.0.4 (last accessed on 8 June 2020)

takes, as input, the geospatial data source along with other configurations. Based on the defined input and configurations, the tool creates a so-called R2RML or RML mapping which includes rules to generate 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 the 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 that, 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 [97]. 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-spatial[50], which performs the translation from GeoSPARQL to SQL. This approach allows accessing geospatial data stored in the geospatial databases as linked geospatial data using GeoSPARQL queries [98].

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

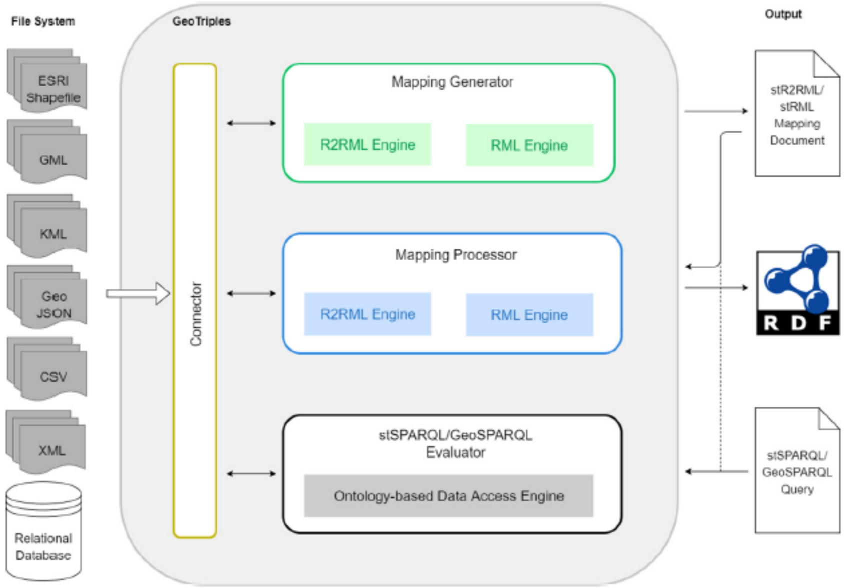


Figure 13. A workflow of generating RDF in GeoTriples[51]

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 converts them into RDF data. The tool is based on the utility called Geometry2RDF[52] and provides a command-line interface [99]. A user of the tool should provide the geospatial data source and the 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 converting datasets to 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

---

51 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.
52 https://github.com/boricles/geometry2rdf/tree/master/Geometry2RDF (last accessed on 8 June 2020)

service in three ways, which are linking based on a SPARQL endpoint, RDF file, and Apache Standbol Entity Hub.

Table 4. 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	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	ESRI Shapefiles XML GML KML JSON and GeoJSON CSV Databases: PostGIS and MonetDB	ESRI Shapefiles GML KML Databases: Oracle Spatial, PostGIS, MySQL, IBM DB2 with Spatial extender	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	D2RQ Engine	Geometry2RDF	OpenRefine is written

	a Java programming language	(http://d2rq.org/)	(https://github.com/boricles/geometry2rdf/tree/master/Geometry2RDF)	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

3.5 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 on 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, pre-processing functions. To compare supported features of the RDF linking frameworks, Table 5 was developed. It can be seen from Table 5, 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 of supported output formats, all of these frameworks support at least two different RDF syntaxes, while only LIMES support 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.

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 in 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 if two strings are identical), RatcliffObershelp (calculated by dividing the matching characters of two strings by the total number of characters in those strings), vector space measures such as Euclidean and Manhattan distance to name just a few. For 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 in 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 interlinking RDF data sources with a published RDF data source using a SPARQL endpoint. This feature is beneficial when RDF data needs 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 which can be used on geospatial RDF data. For instance, a geographical distance measure (computes the geographical distance between two points), 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 converting to an RDF data model and 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.

Table 5. 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	• N-Triples (N3) • Turtle (TTL) • Notation3 (N3) • Tab Separated Values (TAB) • Comma Separated Values (CSV)	• Alignment • Comma Separated Values (CSV) • JavaScript Object Notation (JSON) • RDF • XML • SPARQL endpoint	• JSON • TSV • CSV • XML • RDF various syntaxes • MS Excel

	• SPARQL endpoint		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 License	Apache License, 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 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 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, lowercases, etc

3.6 Methodology for producing linked geospatial cultural heritage data

A methodology for producing linked geospatial CH data consists of five steps as shown in Figure 14. 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 is cleaned by removing whitespaces and making it 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 completed, the RDF data can be exported and fed in to 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 Cloud[53] that stores the collection of RDF data accessible to people and machines.

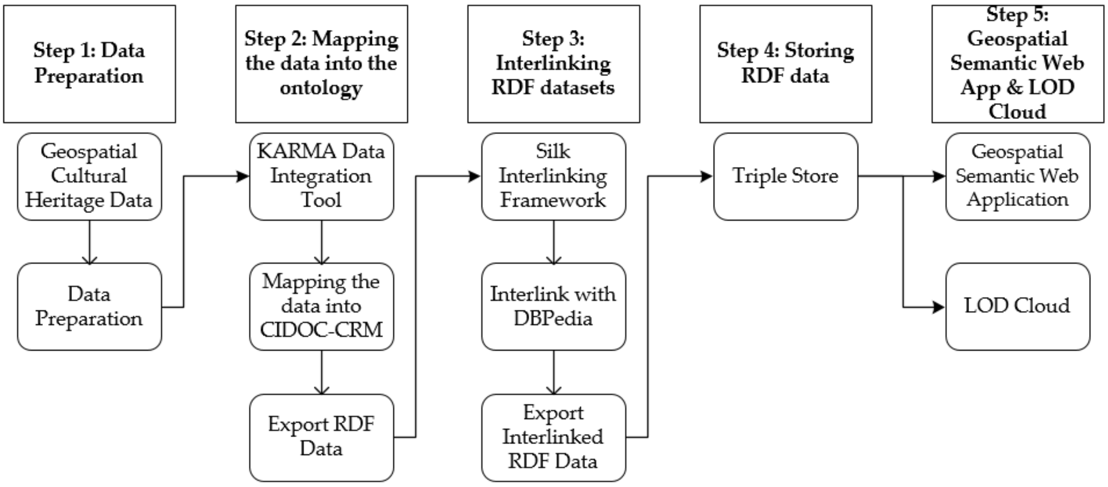


Figure 14. From geospatial cultural heritage data to linked geospatial cultural heritage data.

In order to demonstrate the applicability of this methodology, we mapped a sample geospatial CH data and interlinked it with the related RDF data from DBpedia. The sample data is about CH places located in Western Australia. This data is freely available on the website of the Government of Western

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

Australia[54]. 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. The attributes of the data and mapping are illustrated in Figure 15. As discussed previously, in the second step the data should be interlinked. As an RDF interlinking framework, the SILK Interlinking Framework was selected. The framework requires the data source 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, the 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. For the DBpedia data, it was 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 16.

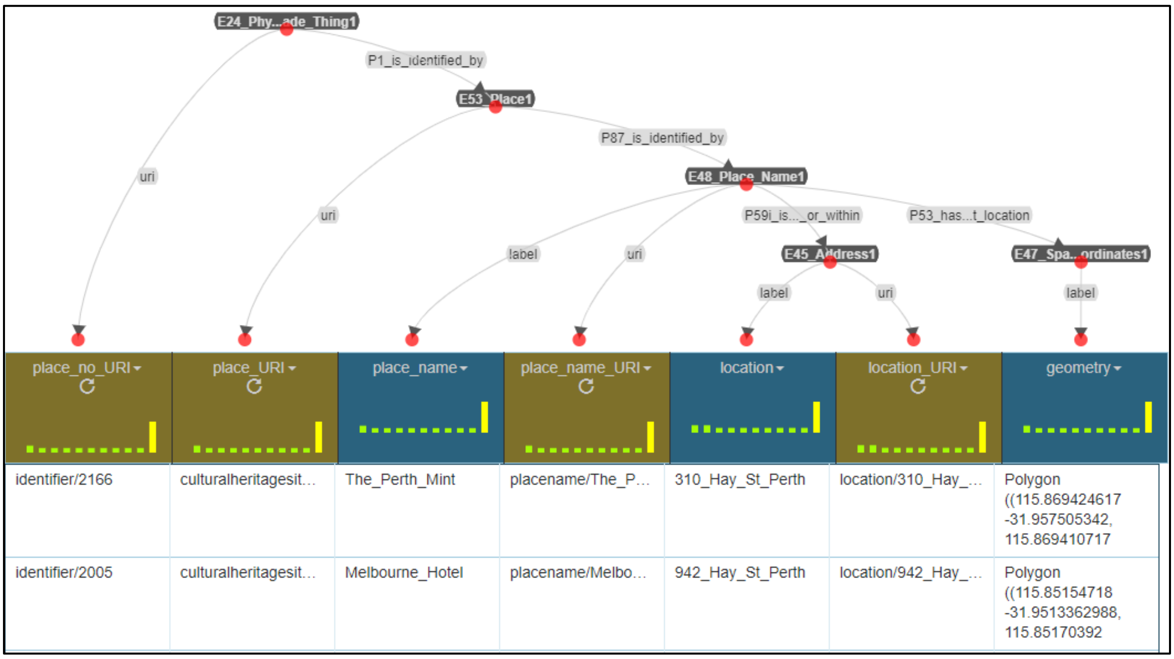


Figure 15. Mapping cultural heritage data into CIDOC-CRM ontology in Karma Data Integration Tool.

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

Figure 16. Cultural heritage places of Western Australia in DBpedia SPARQL endpoint.

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 which were replaced with spaces as shown in Figure 17. 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, a 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 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.* [100].

Then, the interlinking process was performed, illustrated in Figure 18. 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 19. As mentioned previously, in step four interlinking RDF data can be stored in a specialized RDF store called a triple store. The RDF data then can 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.

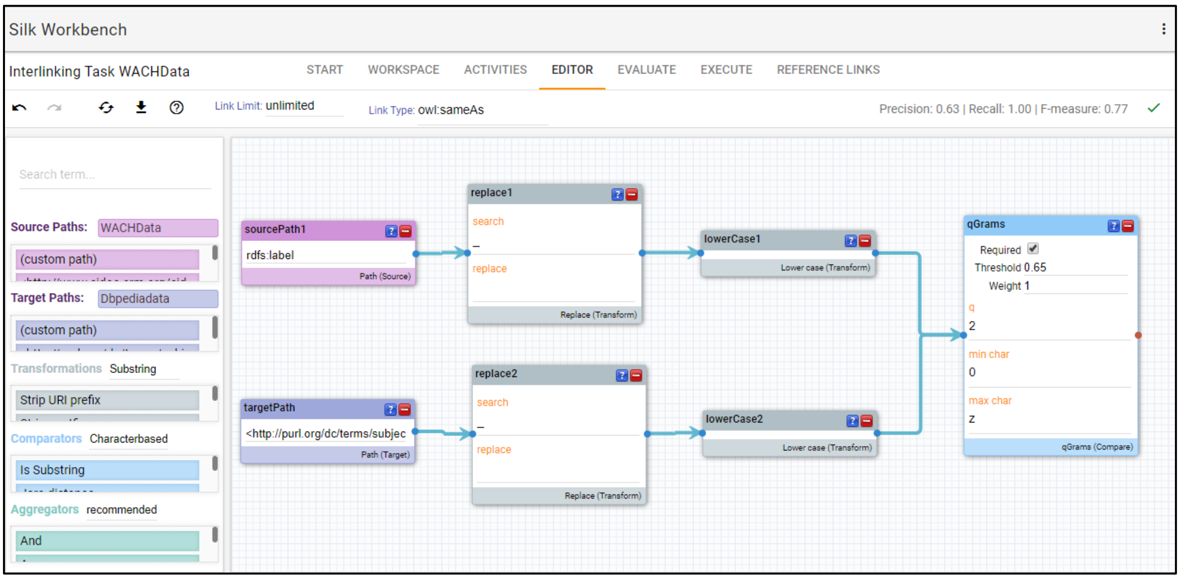


Figure 17. Configuration of the interlinking datasets in Silk Interlinking Framework.
Copyright © 2020 University of Mannheim. Apache Licence 2.0.

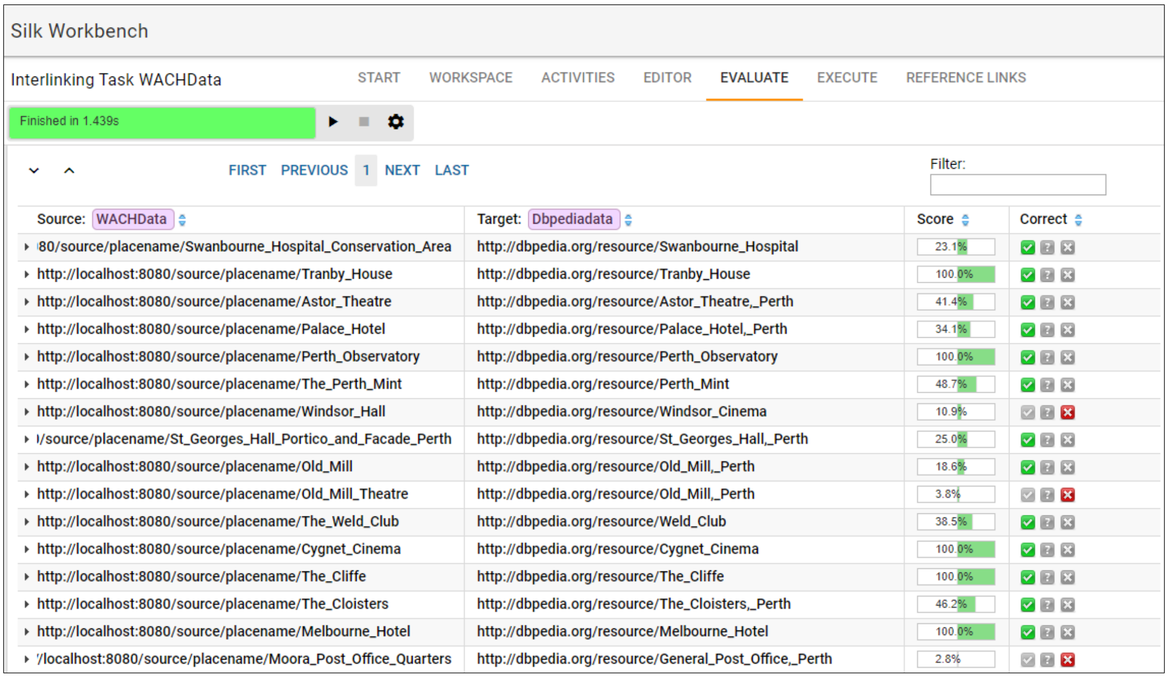


Figure 18. Interlinking process in Silk Interlinking Framework. Copyright © 2020 University of Mannheim. Apache Licence 2.0.

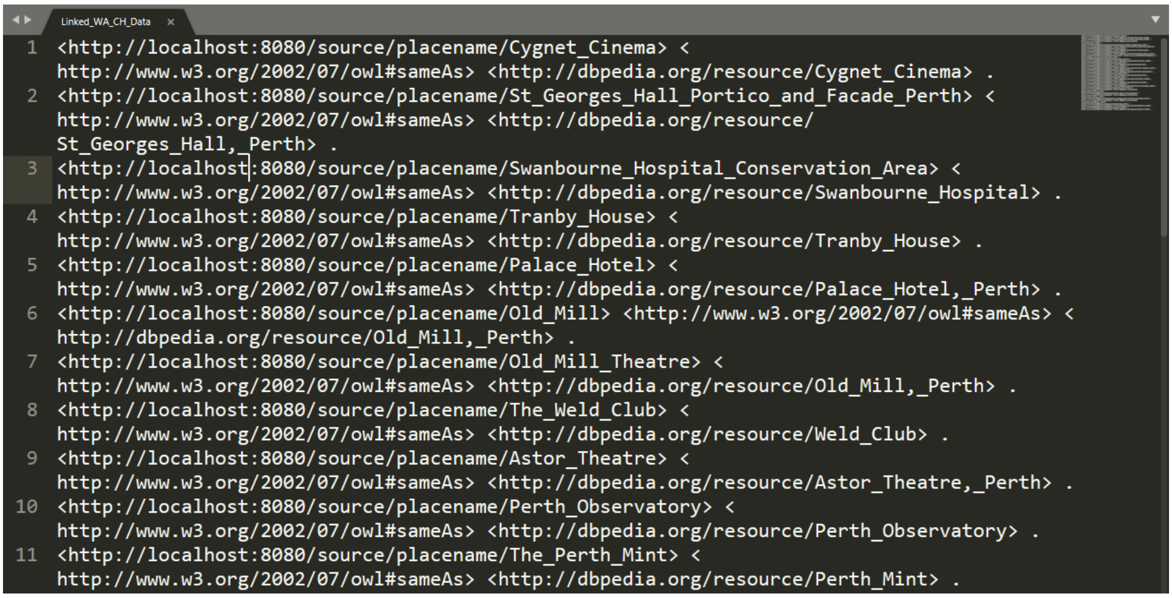


Figure 19. Results of the interlinking process

3.7 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 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 my previous research article [101]

3.8 Conclusion

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 the recent research findings by Schmachtenberg, Bizer and Paulheim [16], who analyzed the adoption of LOD best practices including interlinking in different domains such as media, life sciences, geographic, etc., 44 % of the 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. In order 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. This data was 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 on some of the key challenges and limitations that CH researchers and professionals may encounter when using the evaluated tools and frameworks including the methodology.

Chapter 4 A Web Repository for Geo-located 3D Digital Cultural Heritage Models[55]

This chapter presents a web platform that integrates digital maps, 3D digital cultural heritage models, and geospatial data such as geolocation. Since this web platform is based on easy-to-implement FOSS frameworks and tools, it may be easily replicated, by cultural heritage institutions and professionals at no cost, for long-term archiving and visualization of geo-located 3D digital cultural heritage models on the Web.

Abstract: Recent advances in 3D surveying and web technologies have made a significant contribution to the digital conservation and dissemination of cultural heritage. 3D cultural heritage models are now a critical component in the cultural heritage conservation, which are also employed for other use cases in education, research, tourism, virtual and augmented reality. The World Wide Web is used as a primary medium for dissemination of 3D cultural heritage models, while databases including databases of web repositories for long-term archiving. This article aims to report a new methodology and a web repository, which integrates maps, 3D models, and geospatial data such as geolocation. It can, therefore, be utilized for long-term archiving and visualization of geo-located 3D digital cultural heritage models on the Web. Unlike many previous related projects which developed the web repositories from the ground up, the web repository built using this methodology is based on free and open-source, easy-to-implement content management system namely KeystoneJS and other associated frameworks, which are reusable and completely extendable. It can, thereby, help cultural heritage organizations and cultural heritage professionals facilitate the rapid development of web repositories for geo-located 3D digital cultural heritage models, which can also be further extended per project requirements. While this methodology is presented for the cultural heritage domain in mind, in the

---

55 This chapter is a copy of the peer-reviewed research article - Nishanbaev, I. A web repository for geo-located 3D digital cultural heritage models. Digital Applications in Archaeology and Cultural Heritage. 2020, 16, e00139. https://doi.org/10.1016/j.daach.2020.e00139

long term it can be employed and extended for use in a wide range of domains such as archaeology, engineering, and geographic information systems (GIS) among others.

4.1 Introduction

In the last years, 3D digital documentation of tangible cultural heritage (CH) received a great deal of attention. With recent advances in remote-sensing technologies and imaging devices, CH professionals now equipped with powerful 3D surveying tools and 3D modelling software applications to create an impressive 3D replica of CH sites and objects [19,102,103]. These developments, in turn, have generated in the last decade a great deal of large-scale and individual 3D CH digitization projects across the globe. A "CyArk"[56] project is one of the highly relevant examples, which has now created 3D digital models of over 200 different CH sites on all seven continents. Another apt example is "3D-ICONS", a European project funded by the European Commission, which has massively digitized Europe's archaeological and historical buildings in a 3D form in order to contribute 3D content to the "Europeana"[57] portal [3,104].

Alongside the advances in 3D modelling technologies, in recent years web-based CH repositories have started emerging for long-term archiving and dissemination of 3D digital CH models. Long-term archiving of 3D CH models is of paramount importance to ensure digital conservation of CH, while dissemination of CH is essential to reach the wider audience possible. These concepts are also underlined by international CH agencies such as *The United Nations Educational, Scientific and Cultural Organization* (UNESCO)[58], *International Council of Monuments and Sites* (ICOMOS)[59]. There is also a CH agency namely *The International Committee for Documentation of Cultural Heritage* (CIPA), which advises on surveying methods for CH monuments and sites [105]. Furthermore, 3D digital CH

56 https://www.cyark.org/ (last accessed on 8 August 2019)
57 https://www.europeana.eu (last accessed on 8 August 2019)
58 https://whc.unesco.org/en/convention/ (last accessed on 8 August 2019)
59 https://www.icomos.org/en/charters-and-other-doctrinal-texts (last accessed on 8 August 2019)

models are now not only utilized for plain visualization but also for different use cases in education, research, tourism, virtual and augmented reality among others [106-108]. Regarding web-based repositories for 3D digital CH models, there have been some innovative projects in the CH domain in the last decade. 3D-COFORM is among those projects, which offers an integrated repository system. It can work as a web repository for distributed institutions to store, manipulate and export 3D digital CH models, which can be suitable for a CH institution with several sub-institutions or sub-divisions in different locations [80]. Another well-known platform for CH institutions is OMEKA. It is a purpose-built platform for CH institutions that can be downloaded and can work as a web repository. This platform offers a 3D storage and 3D visualization features.[60] On the other hand, there are some successful commercial repositories available on the Web such as Sketchfab, TurboSquid, which can be used to publish, to store and to visualize 3D CH models over the Web. Although there are some downloadable web repositories and commercial repositories on the Web, 3D CH models are difficult to find, use and re-use. In fact, Champion and Rahaman [17] surveyed 14 proceedings of leading digital heritage events and conferences by *The International Society for Virtual Systems and Multimedia* (VSMM), *Computer Applications and Quantitative Methods in Archaeology* (CAA), *International Committee of Architectural Photogrammetry* (CIPA), *The European Mediterranean Conferences* (EuroMed), and *The Digital Heritage International Congress*, to examine accessibility of 3D models in published scholarly research articles, from 2012 to 2017. According to their findings, out of 1483 examined conference articles only 17.9% or 264 articles incorporated 3D models or images of 3D models. Strikingly, only 9 articles had accessible 3D content. Hence, CH professionals, CH organizations and CH domain in general, need more projects and easy-to-implement methodologies to build a web repository for long-term storage and dissemination of 3D digital CH models. Moreover, in most cases, 3D digital CH models incorporate spatial information such as geolocation. This article aims to present a web repository, and a reusable and extendable methodology to encourage the

60 https://omeka.org/ (last accessed on 8 August 2019)

development of web repositories for geo-located 3D CH models. Moreover, unlike many previous related projects which developed the web repositories from the ground up, the web repository built using this methodology is based on free and open-source, easy-to-implement content management system namely KeystoneJS and other associated frameworks, which are reusable and completely extendable. This allows to accelerate and ease the development process and requires less technical and programming skills, which CH professionals and CH organizations can take advantage of. While this methodology is presented for the CH domain professionals in mind, in the long term it can be employed and extended for use in a wide range of domains such as archaeology, engineering, geographic information systems, etc.

4.2 Contributions and scope of the article

The contributions of this article to the CH domain are as follows:

- A new methodology, to build a web repository for 3D digital CH models, based on open-source, database-driven, easy-to-implement content management system namely KeystoneJS and other associated frameworks which integrates maps, 3D CH models, and geospatial data such as geolocation.
- This methodology and the web repository can be implemented in CH institutions such as museums, galleries, archives to facilitate storage and visualization of 3D CH models and relating information. Since a 3D visualization can be a very close replica of the reality, it can help to arouse interest in the CH organization, for instance into a museum. Thereby, it can help to promote national and international tourism.
- This methodology and the web repository extend the existing body of knowledge and expertise for building web repositories for 3D models and help to tackle the challenges associated with integrating geospatial data and maps with 3D models.

The concepts such as 3D surveying and 3D modelling techniques, metadata, efficient search and retrieval of big cultural heritage data, 3D web-GIS and 3D analysis are out of the scope of this article.

The remainder of the article is organized as follows: Section 2 outlines some previous related work. Section 3 provides an overview of the background literature related to 3D visualization on the Web, 3D frameworks, 3D file formats, and web content management systems. Section 4 presents a methodology used to develop the web repository followed by a results and discussion part in Section 5. Section 6 provides conclusion and planned future work.

4.3 Related work

In recent years there have been a few projects related to the development of a web repository for 3D CH models. The next paragraphs summarize some of these projects and provide an overview including the purpose of each project and employed technologies such as server-side programming language, 3D visualization framework, and database management system.

A project by Dhonju, Xiao, Mills and Sarhosis [24] developed a prototype termed "Share Our Cultural Heritage" (SOCH), which offers a web-based as well as a mobile-based geo-crowdsourcing platform to document and share CH sites. The main aim of the platform is to encourage public participation to document CH sites in remote locations that are not easily accessible. The platform collects geo-located CH data from the public via a web interface and supports image visualization, 3D visualization in addition to the web-based generation of 3D models from images among others. The platform has been implemented using a PHP server-side programming language, and 3D visualization achieved through a free and open-source library named Potree. OpenLayers mapping library was used to visualize maps and the PostgreSQL/PostGIS database was employed for storing datasets. Another related work is by Felicetti and Lorenzini [109], which implemented a semantic enabled persistent web repository for 3D digital CH models. The repository offers storage and retrieval of 3D models and aims to improve CH data interoperability. It is based on the Fedora v3.4 platform by "Fedora Commons", and the datasets in the repository are stored in the digital object storage provided by the Fedora platform. It uses the "Apache Solr" query framework to provide a searching capability in the repository and employs other plugins by QGIS, Blender and Google SketchUp. Another notable project, related to the development of repository and platform for 3D CH models, is MayaArch3D [110-112]. It has been developed through a collaboration of five institutes namely the German Archeological Institute, GIScience Research Group, Bruno Kessler Foundation, University of Nebraska and Honduran Institute of Anthropology and History. The main aim of the project is to develop a web platform, which stores spatial and non-spatial data, and which not only visualizes 3D models but also allows users to perform GIS analyses on 3D models. This platform is based on a Geomajas framework, which incorporates a collection of free and open-source GIS libraries and tools for building Web-Geographic Information Systems (GIS) applications. The datasets in the platform are stored in a PostGIS database and FileMaker, while 3D visualization is achieved through a 3D visualization library of Three.js.

Furthermore, there are other scholarly research projects in a closely related direction such as work by Kiourt, *et al.* [113], which implemented a dynamic web-based virtual museum application termed DynaMus. This application aims to provide virtual museum exhibitions by employing 3D models, game engines, and linked open data. Guarnieri, *et al.* [114] developed a web application which included a segmentation of 3D CH models to define user-selectable areas on a 3D model. However, as mentioned in the previous section, this work presented in the article differs from previous projects in a way that the proposed web repository and the methodology is based on a FOSS, easy-to-implement, database-driven content management system namely KeystoneJS and other associated frameworks. Hence, it helps CH organization and professionals facilitate the rapid development of web repositories for 3D CH models, which can be further extended per project requirements.

4.4 Background literature

4.4.1 3D visualisation on the web

The World Wide Web has now become a primary medium for sharing information and data, which can display data in different forms including in a 3D form. These days, 3D visualization on the Web is not a new topic. From the early 1990s engineers started working on 3D visualization technologies on the Web. In 1994 the VRML consortium released a “Virtual Reality Modelling Language”, also known as “Virtual Reality Markup Language” (VRML), file format, which was designed to represent 3D interactive vector models on the Web. At that point in time, rendering of 3D models required a 3D/VRML plugin of the Netscape web browser. There were three major releases of the format namely VRML, VRML 2.0 and VRML 97. In 1997 the last revised version of the format become an ISO standard.[61] At that time, VRML aroused a lot of interest among engineers, however, it was not widely used as a web standard due to some reason such as limited bandwidth and slow internet access, and technological hardware limits of computers. Furthermore, many researchers believe that it does not provide fast performance on the visualization of complex and large-scale 3D models as it is an interpreted language, not a compiled one [115]. In 1997, VRML superseded by Extensible 3D (X3D) and the VRML Consortium was renamed to the Web3D consortium. X3D is a royalty-free, ISO-standardized, XML-based file format to represent 3D models which can be integrated with HTML5 and XML.[62]

With the introduction of Web Graphics Library (WebGL), it brought about a plug-free 3D rendering to the Web. WebGL is a cross-platform, a royalty-free web standard developed by the “Khronos Group”, and is supported as a default feature by major browsers such as Google Chrome, Microsoft Edge, and Mozilla Firefox among others.

---

61 http://www.web3d.org/documents/specifications/14772/V2.0/index.html (last accessed on 8 August 2019)
62 https://www.web3d.org/standards (last accessed on 8 August 2019)

4.4.2 3D frameworks

Since WebGL[63] is a low-level 3D graphics API, plenty of libraries has been implemented to ease and to accelerate the development of 3D applications. In the next paragraphs, some of the well-known 3D frameworks are briefly discussed.

CesiumJS[64] is an open-source, JavaScript library by “Analytical Graphics Inc” for building interactive 3D globes and maps. It leverages on WebGL for hardware-accelerated graphics. It supports many types of 3D geospatial formats including new formats and open standards, which have been developed by the Cesium project such as 3D Tiles, Quantized-Mesh, and CZML. 3D tiles has now been accepted as an Open Geospatial Consortium (OGC) standard. These new formats aim to stream and to render heterogeneous and massive 3D scenes, 3D building models, and 3D time-dynamic models more efficiently than previous formats and standards. Cesium has now been implemented in many domains including in aerospace and defense, autonomous vehicles, drone data capturing, 3D cities among others.

Three.js[65] and Babylon.js[66] are arguably two most popular and most-used JavaScript-based 3D libraries for creating and rendering 3D computer graphics in a web browser. They are both actively developing projects and have strong community of developers, yet three.js has larger community than Babylon.js. This can be due to the reason that Three.js was first published to GitHub in 2010, whereas Babylon.js in 2013. Also, Three.js currently offers more features compared to the other one. On the other hand, Babylon.js has extremely well-written documentation and offers onsite playground to test the features and code. While Three.js supports its own playground, it does not support testing of many features available in the library.

63 https://www.khronos.org/webgl/ (last accessed on 8 August 2019)
64 https://cesium.com/cesiumjs/ last accessed on 18 November 2019)
65 https://threejs.org/ (last accessed on 18 November 2019)
66 https://www.babylonjs.com/ (las accessed-on 18 November)

In addition to above-mentioned 3D frameworks, there are some other alternative 3D JavaScript-based frameworks exist such as sceneJS, OSG.JS, Cannon.js among others.

4.4.3 3D file formats

A 3D file format is used to store geometry, appearance, scene and if available an animation of a 3D object in the plain text form or in the binary form. The 3D file formats can be divided into two types namely proprietary and non-proprietary. The former term is used to describe a native file format from software applications, while the latter describes open-source file formats. Currently, there are at least 140 different 3D formats available [116]. The following paragraphs discuss a few most-used, non-proprietary 3D file formats and their features.

COLLADA (Collaborative Design Activity) is a 3D file format based on XML (Extensible Markup Language) and has a '*.dae*' file extension. The format is an ISO standard and under "SCEA Shared Source Licence", which can be used for commercial and non-commercial purposes. At the time of writing, the latest version of COLLADA - 1.5.x can encode information such as geometry mesh geometry, boundary representation, material, texture, scene and animation for a 3D CH model [117]. This file format could be an appropriate format for many purposes, such as visualizing static 3D models on the web, animation, full scene graphs. However, COLLADA was built to be an interchange file format and to solve the issue of interoperability. Furthermore, it is an XML based format, hence if a 3D model is very complex, it occupies more storage and not be rendered as fast as a binary encoded format. Therefore, COLLADA could be best deployed for the purpose, where an interoperability could be an issue.

OBJ format supports both plain text form and binary encoding but only the former is open source. The format has been developed by Wavefront technologies, which is now owned by Autodesk. It supports geometry, appearance, and scene, however it does not support animation. It generates another file with an extension of *'.MTL'* (Material Template Library) to store appearance of the 3D model. This format is popular for its simplicity and ease to work with. For instance, it stores material in a separate MTL file, which gives possibility to change the material and texture of the model without editing OBJ file. Nevertheless, the limitation of the format is also linked with this, as separate storage of material and texture information of the model could be an issue in conversion of the format into another format.

gLTF[67] is a 3D format and open standard by Khronos group, which has been designed for the efficient transfer of 3D models. It has an aim to be versatile to represent various types of 3D assets, compact for efficient transmission, and easy to process on the client side. It is often referred as 'JPEG for 3D data' as many applications nowadays use this format, particularly on the Web. It supports complex 3D scenes, which can include animations, materials, etc. The core of the gLTF format is based on JSON (JavaScript Object Notation) which represents the general scene structure, cameras and animations. However, gLTF format has two additional files linked to the core JSON file, which are files with *'.bin'* and *'jpg or png'* extensions, for geometry and texture of the 3D asset respectively.

3D Tiles[68] is a 3D file format and an open specification designed for sharing, visualizing and interacting with massive heterogenous 3D geospatial data such as 3D building, photogrammetry data, and point clouds to name a few. It has been developed by 'Analytical Graphics Inc' and is based on previously discussed glTF format. The format uses spatial data structure named 'the tree' to organize 3D content.

4.4.4 Web content management systems

With a recent significant increase in demand for website maintenance and development, web content management systems (WCMS) are now being utilized frequently. WCMS are easy-to-use and deploy systems, which have been helping many institutions and individuals with limited technical skills to develop websites and blogs. They often consist of two major components namely a front-end interface to represent published content, and a back-end

67 https://www.khronos.org/gltf/ (last accessed on 18 November 2019)

68 https://github.com/AnalyticalGraphicsInc/3d-tiles/tree/master/specification#introduction (last accessed on 18 November 2019)

to publish and edit content which often includes a database as well. Drupal, WordPress, Joomla are some of the well-known examples of WCMS. Most WCMS offer a built-in hosted service, while others offer to download their software which can then be configured and hosted in the desired hosting provider [118].

4.5 Methodology

The next section describes the system architecture, which facilitated the web repository. Afterwards, employed 3D digital models are outlined, which have been used to test the web repository including some background information about each place and technical characteristics of the 3D models.

4.5.1 System architecture

To accomplish the development of a web repository for 3D digital cultural heritage models, a methodology consisting of three parts have been developed, as shown in Figure 20.

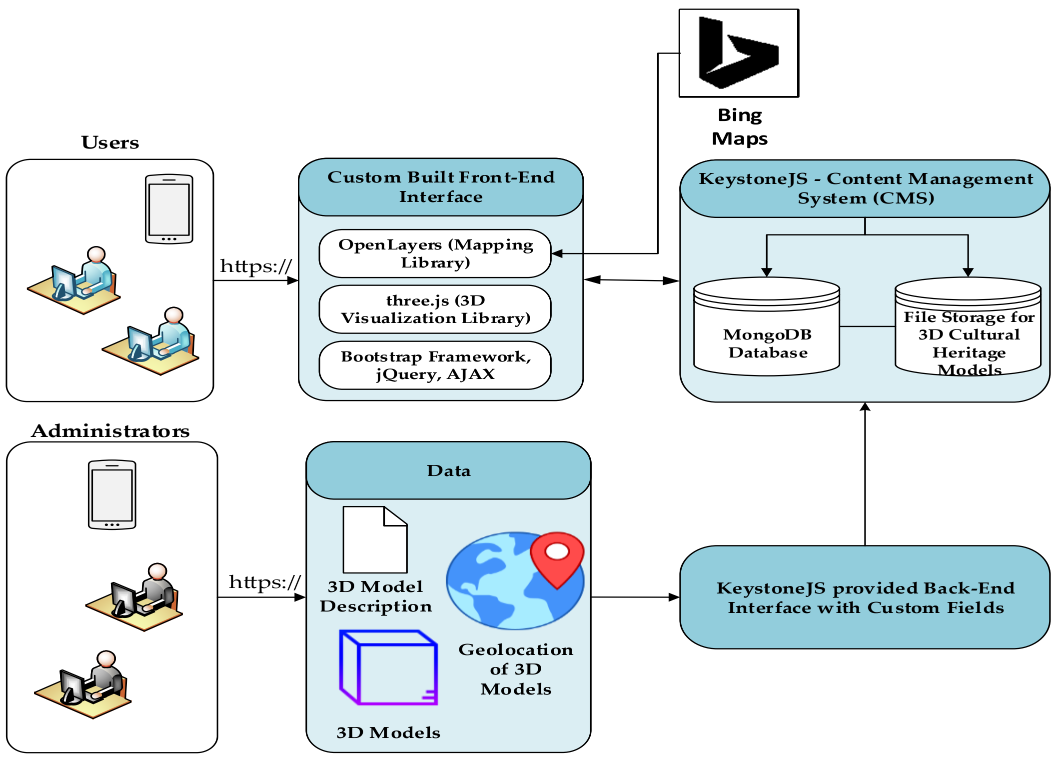


Figure 20. A methodology for the development of a web repository for 3D cultural heritage models.

The first part is the ingestion of 3D models to the content management system, which is achieved through a graphical user interface (GUI). The GUI itself has been developed using frameworks “Bootstrap” and “jQuery”. The former framework is the world’s most popular framework for building the front-end development of web applications, which is also known for its responsive and mobile-first concepts. Since Bootstrap focuses on simplifying the development of front-end web applications, it offers several ready-to-use templates and components such as navigation, forms, and typography. The latter framework is a JavaScript library for simplification of HTML DOM manipulation, event handling, and asynchronous calls. As of writing this paper, it is employed by 74% of the 10 million most popular websites[69].

The second part is the backend of the web repository, which was achieved using a WCMS named “KeystoneJS”. This WCMS is based on a “Node” runtime environment which allows developers to implement commands in a server-side in the JavaScript programming language. Hence, this web repository follows the “JavaScript everywhere” paradigm. KeystoneJS is arguably the most popular open-source content management system for Node.js. Since it offers modular architecture, it can be used for large-scale projects such as portals, e-commerce projects, RESTful web services among others. It offers a standardized set of modules, which allow rapid and easy development of web applications. Furthermore, since it is open source, it can be extended if necessary. As a standard package it offers, auto-generated admin user interface, session management for admin log-in and log out, form processing to name a few. Therefore, administrators of the WCMS can manage inputting of text information, upload of content and edit of content. KeystoneJS natively shipped together with a MongoDB database which is a cross-platform, document-oriented database system. It offers a flexible document data model and is known to be one of the best NoSQL database systems. In the web repository, all text information is stored in this database, while all 3D digital CH models are uploaded and stored in the file storage system.

69 https://w3techs.com/technologies/overview/javascript_library (last accessed on 18 November 2019)

The third part is a front-end interface for the rendering of 3D digital cultural heritage models. This has been implemented using a 3D JavaScript library termed “Three.js”, which is based on a previously discussed WebGL technology. This 3D library has been chosen because of its non-proprietary licence, integration with all web browsers that support WebGL, and relatively well-documentation among others. The library supports 3D digital models in several file formats such as gITF (.glTF), OBJ (.obj), and Point Cloud Data (.pcd) to name a few.

4.5.2 Employed 3D digital models

Figure 21. Employed 3D digital models.
In a clockwise direction, **Gosses Bluff Crater** (Image Source: https://bit.ly/2Kjrsr1, Copyright by rplzzz. CC BY-SA 2.0), **Ballarat Town Hall** (Image Source: https://bit.ly/2WS6cf1, Copyright 2019 Film Victoria*),* **Split Point Lighthouse** (Image Source: https://bit.ly/2XnIlYL, Copyright by Cafuego. CC BY-SA 2.0), **Heron Island, Great Barrier Reef** (Image source: https://bit.ly/33fcGJ6, Copyright by Jon Connell. CC BY 2.0)

In total eleven different 3D digital models have been used to test the web repository. These 3D models are 3D digital replicas of the places located in Australia. Four of these 3D models are presented in the article namely Gosses Bluff Crater, a facade of Ballarat Town Hall, Split Point Lighthouse and a part of the reef area in Heron Island, Great Barrier Reef. The images of these places are shown in Figure 21.

Gosses Bluff Crater is a registered sacred site, which is of great cultural significance to Australian Western Arrernte Aboriginal people. This place is also known as Tnorala and located in the southern Northern Territory of Australia. It is believed that an asteroid or comet with a diameter of about one kilometer crashed into the earth about 142 million years ago. Today Gosses Bluff Crater is about six kilometers wide but the original one believed to be twenty or twenty-five kilometers.[70] The second place is Ballarat Town Hall, located in the state of Victoria, which was constructed in 1870. The building is believed to be one of the fewest in the world with a peal of bells. Up to now it has been extended and altered several times. It is in the heritage list of the National Trust, which protects natural and built cultural heritage in Australia.[71] Split Point Lighthouse is a lighthouse, which was originally called Eagles Nest Point and was renamed to Split Point in 1913. It was constructed in 1891 and is located in the state of Victoria. The lighthouse helps to guide ships in the Shipwreck Coast and listed in the Victorian Heritage Register.[72] Heron Island is a coral cay in the southern Great Barrier Reef in the state of Queensland in Australia. The island is named after the herons which are part of the birdlife that inhabits the island. Furthermore, the island is home to over two hundred thousand species of birds. The entire Heron Island is a world heritage listed marine national park since September 1943.[73]

The technical characteristics of 3D digital models are presented in Table 6. All 3D models are encoded in a 3D file format of glTF, which is one of the best choices for the Web as it significantly minimizes the size and loading times of 3D models compared to other 3D file formats.

The web repository has also been tested with another 3D file format of “OBJ”. As discussed previously, it is an open format by Wavefront technologies. This file format usually incorporates three files, which are a 3D object with the “.obj” file extension, a material template library with the “.mtl” file extension, and one or more image files (with the file extension of “.jpeg”, ”.png”, etc).

---

70 https://bit.ly/2ZE5d4h (last accessed on 8 August 2019)
71 https://vhd.heritagecouncil.vic.gov.au/places/67573/download-report (last accessed on 8 August 2019)
72 http://splitpointlighthouse.com.au/about/history/ (last accessed on 8 August 2019)
73 http://www.greatbarrierreef.org/islands/heron-island/ (last accessed on 8 August 2019)

The first file with ".obj" extension stores 3D object-related information such as vertices, faces in a plain text, while the second file with the ".mtl" file extension stores a material that is the appearance of the 3D object. This file also stores the path link to the image files.

Table 6. Characteristics of employed 3D digital models.

3D Model Name	File Format	Geometry	Vertices	Size	Licence and URL
Gosses Bluff Crater	glTF 2.0	Triangles 9.1k	4.9k	59 MB	Copyright by SebastianSosnowski. CC BY 4.0 (https://bit.ly/2MjHbaW)
Ballarat Town Hall	glTF 2.0	Triangles 551.1k	290.6k	9 MB	Copyright by City of Ballarat. CC BY 4.0 (https://bit.ly/2ZzH6mY)
Split Point Lighthouse	glTF 2.0	Triangles 346.1k	174.1k	24 MB	Copyright by Stefan A Vollgger. CC BY 4.0 (https://bit.ly/2ZC4v7k)
Heron Island, Great Barrier Reef	glTF 2.0	Triangles 8.6M	4.3M	400 MB	Copyright by renataferrari. CC BY 4.0 (https://bit.ly/3162IYN)

4.6 Results and discussion

As mentioned in the previous section, the user interface of the web repository has been developed using a Bootstrap toolkit, OpenLayers mapping library, and jQuery. The user interface is illustrated in Figure 22, which has a base-map with clickable red pinpoints on top of it. These red pinpoints represent CH places and hold CH place related information such as a name, short description, and license information. In the web repository, CH places can be created by the administrator using the back-end interface, as shown in Figure 23. While creating the CH places, the geolocation of the CH places in the form of latitude and longitude can be set. The precise geolocation for the most CH places can be found in the GeoNames platform, which is a geographical database with more than 11 million place names in it. The visualization of 3D digital models is illustrated in Figure 24, which is achieved using Three.js 3D visualization library. The 3D models can be manipulated along the X, Y and Z coordinate axes, and can be zoomed in and out.

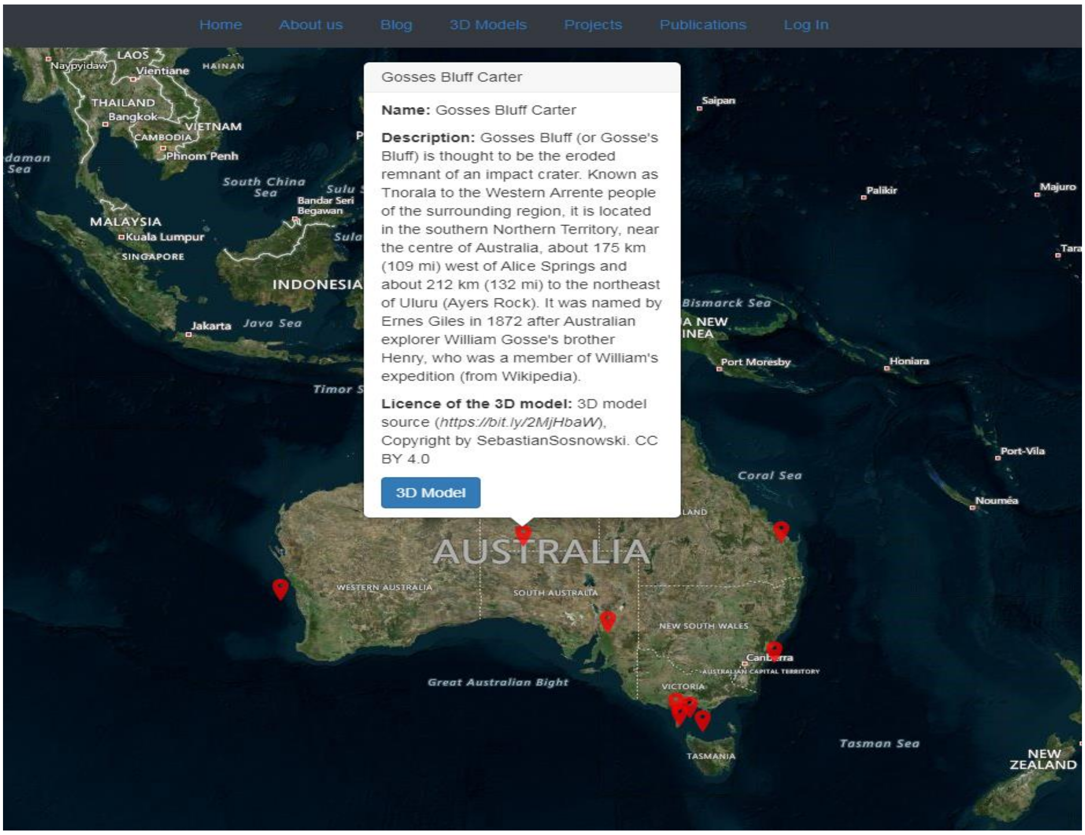


Figure 22. A front-end interface for rendering of 3D CH Models. Copyright 2019 by Microsoft Bing Maps.

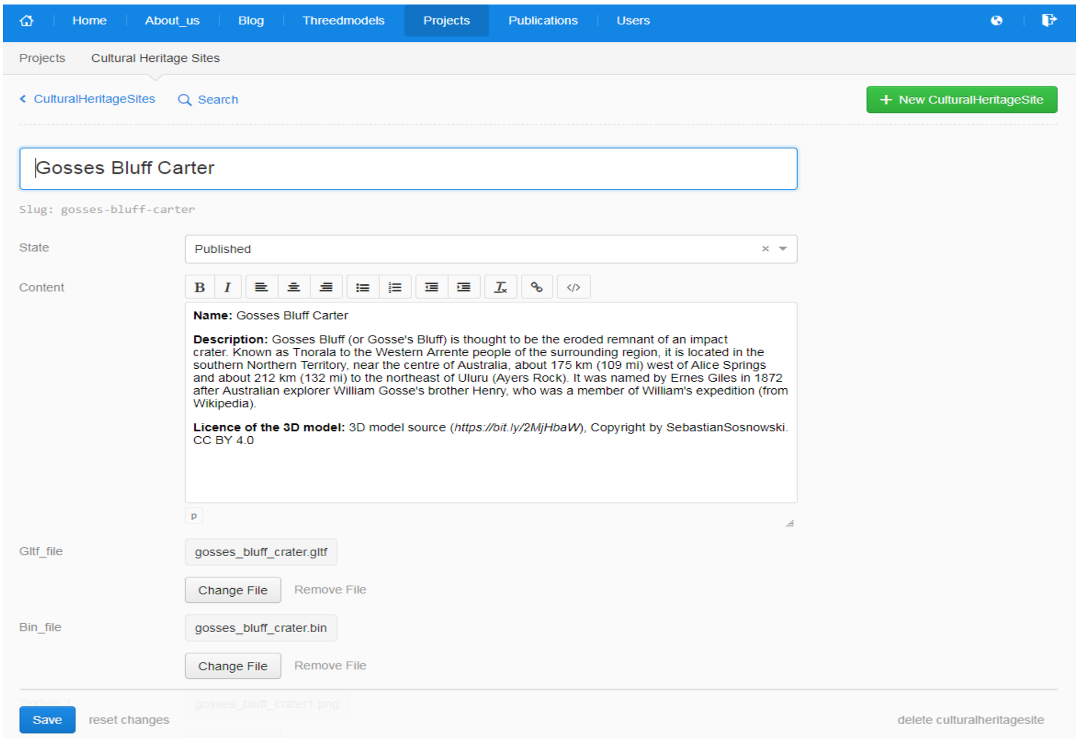


Figure 23. A back-end interface of the web repository. Copyright 2019 by KeystoneJS. MIT licence

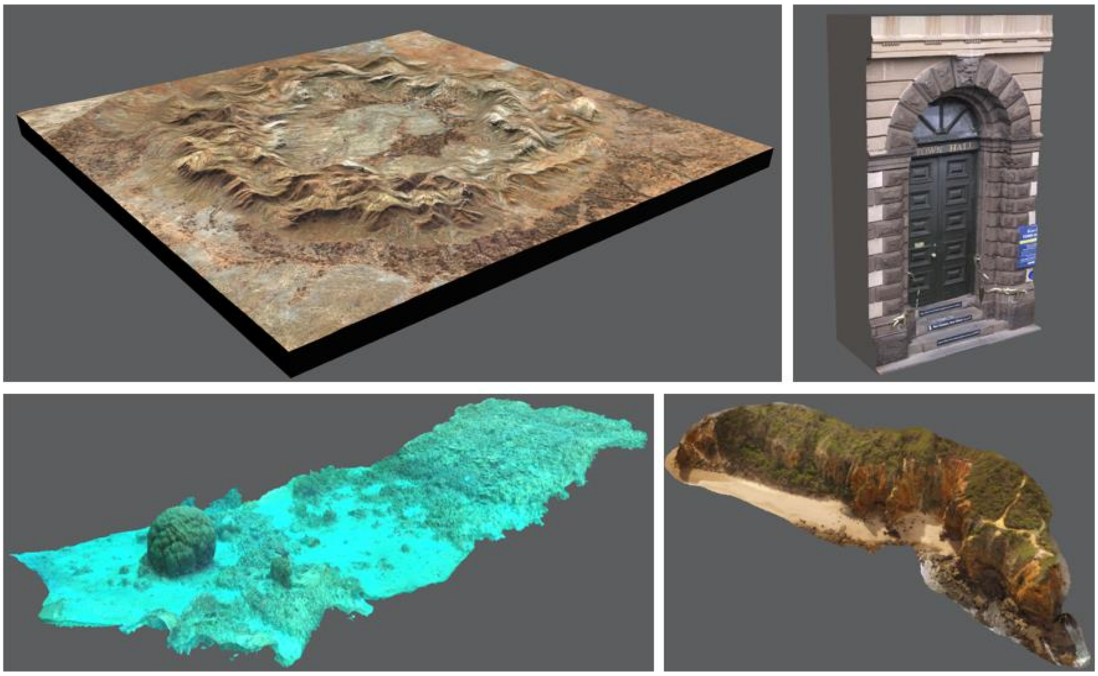


Figure 24. Visualization of 3D cultural heritage models in the web repository. In a clockwise direction, **Gosses Bluff Crater** (3D Model Source: https://bit.ly/2MjHbaW, Copyright by SebastianSosnowski. CC BY 4.0), **Ballarat Town Hall** (3D Model Source: https://bit.ly/2ZzH6mY, Copyright by City of Ballarat. CC BY 4.0), **Split Point Lighthouse** (3D Model Source: https://bit.ly/2ZC4v7k, Copyright by Stefan A Vollgger. CC BY

4.0), **Heron Island, Great Barrier Reef** (3D Model Source: https://bit.ly/3162IYN, Copyright by renataferrari. CC BY 4.0)

4.6.1 Level of customization for the front-end 3D visualization

Table 7. Level of customization for the front-end 3D visualization.

Type	Description
3D Scenes	Supports adding and removing 3D objects including dynamic adding and removing at run-time
Cameras	Supports perspective camera, orthographic camera, cube camera, and stereo camera
Animation	Supports animations such as forward and inverse kinematics, morph and keyframe.
Audio	Supports positional and non-positional audio effects in 3D scenes
Lights	Supports ambient light, directional light, hemisphere light, point light, rectangular area light, spot light
Shaders	Supports lens-flare, depth pass, post-processing and access to WebGL shader library
Built-in 3D objects	Supports different types of built-in 3D objects such as Box, Circle, Cone, 3D text among others
3D Model Loaders	gLTF, OBJ, JSON, PCD (Point Cloud Data), etc.
Level of Detail for 3D models	Supports visualizing meshes with more or less geometry based on distance from the camera
Controls	Mobile device-based orientation, drag and drop interaction, fly mode, pointer lock controls, trackball control

3D visualization in Three.js offers a great deal of customization. The Table 7 and the following paragraph describe some of the customizations, which can be implemented on top of the presented web repository.

3D models to the scene can be added and removed at run-time as Three.js capable of updating scenes if there is any change to it. Furthermore, it supports other different components such as various cameras, animation, positional and non-position audio, lights and others, which can be used to customize the front-end of the 3D visualization in the web repository. Level-of-detail (LOD) is another component supported by Three.js using which meshes get rendered with more or less resolution based on the distance from the camera. For instance, for a 3D model, three associated LODs can be attached, one for low detail, one for medium detail, and one for high detail rendered in a close distance to the 3D model. Nevertheless, if the 3D models are extremely large and the task is to offer comprehensive multi-resolution rendering with progressive loading and compression, a set of tools by Nexus[74] can be used. Nexus is a collection of tools that incorporates a tool termed ‘Nxsbuild’. This tool allows converting 3D models into a multi-resolution 3D model with ‘.nxs’ file format. Afterwards, the multi-resolution 3D model can be loaded into Three.js.

4.6.2 Scalability of the web repository and a sample use case

The proposed web repository has been implemented using a database-driven content management system (CMS) named KeystoneJS, which is based on the NodeJS runtime environment. Since NodeJS has an event-driven architecture capable of running asynchronous input and output, KeystoneJS can run several processes separately from the primary application thread. Therefore, this offers numerous benefits such as improved application performance, scalability, and enhanced responsiveness. As mentioned in the methodology section, the application has been tested with eleven different 3D models ranging the size up to 400 MB.

KeystoneJS allows creating multiple administrators without limitations, which can be necessary when working in distributed workflows. When creating a new item, for instance a 3D model, administrators can choose the states of the item among draft, published and archive states. Furthermore, all administrators can search items by the title of the item, filter items out according to the title, state, content, and location, and download the list of items in JSON and Comma Separated Values (CSV) formats.

74 http://vcg.isti.cnr.it/nexus/ (last accessed on 18 November 2019)

The web repository has been employed in a sample use case to better evaluate the benefits and limitations of the methodology as well as the web repository. This sample use case is described in the next paragraph.

Within every museum, there is a need for an online presentation of some of the available CH artifacts and assets. This action, in turn, can help to promote national and international tourism, and thereby may result in an increase in the revenue. In such cases, this methodology and web repository can be exploited to serve as a platform for online 3D demonstration of some of the CH artifacts and relating text information. Furthermore, a 3D visualization can be a very close replica of the reality, hence it can help to arouse interest in the museum.

In this use case, two administrators (administrator-A and administrator-B) have been created both with access privileges to the back-end of the application. It is assumed that they work remotely and not from the same physical location. Assuming one of them is working on the task of generating 3D models of cultural heritage objects, for instance with free and open-source software such as Regard3D[75]. The other one is preparing the text information associated with CH objects and identifying geolocation of CH objects, for instance from the "GeoNames" platform of geographical places. Once 3D models are ready, administrator-A uploads them into the platform but assigns them into the 'draft' state. When the text and geolocation information is ready, administrator-B inputs them into the platform and informs administrator-A about the progress. When they are ready to make the content available to the public, any of the administrators change the state of the content to 'published' and the content will be ready for public consumption. If they need to change any content, they can easily log in and perform necessary changes. This use case can be extended to several collaborators, and if the geolocation is not available for 3D models, the same methodology can be applied to implement a platform without a map interface.

75 http://www.regard3d.org/ (last accessed on 18 November 2019)

4.7 Conclusion and future work

3D surveying tools and techniques now allow CH professionals to 3D digitize tangible CH ranging from the smallest CH artifact to the largest CH building. This rapid evolution in technology has significantly advanced the conservation of CH in a digital form. However, the lifetime of 3D CH models is short, especially those which resulted from the individual CH research projects. According to the survey findings of Champion and Rahaman [17], who looked at 1483 digital heritage articles published in 14 recent proceedings of major CH conferences, from the examined 1483 conference articles, only 9 articles had accessible links to 3D assets. In this article, the author presented a new methodology for building a web repository, which can be employed for long-term archiving and visualization of geo-located 3D digital CH models on the Web. This article presented a web repository, and a reusable and extendable methodology to encourage the development of repositories for geo-located 3D CH models. Moreover, unlike many previous related projects which developed the web repositories from the ground up, the web repository built using this methodology is based on free and open-source, easy-to-implement content management system namely KeystoneJS and other associated frameworks, which are reusable and completely extendable. Hence, it can help CH organizations and CH professionals facilitate the rapid development of web repositories for 3D CH models, which can be then further extended per project requirements.

However, there are still many technical challenges and issues with this web repository, methodology, and generally with many web repositories for 3D digital CH models. There are many challenges exist such as digital rights management of 3D digital models, visualization of very large-sized 3D models, efficient search and retrieval of big 3D CH data to name a few. However, the next paragraphs discuss two major technical challenges, which will be further researched by the author.

Metadata and semantic interoperability: CH data integration and interoperability, and reuse of existing CH data in new applications are still some of the big technical challenges on the Web. There is a plethora of CH data in different file formats including in 3D formats available on the Web, however, data integration and interoperability of CH data across web platforms is still a daunting task. This is mainly because many CH organizations have their generic non-standardized software designs such as non-standardized metadata schema, data model and database structures. Due to these diverse software designs, web CH platforms struggle to communicate with each other coherently. This results in dispersed and poorly-linked CH data on the Web. Recently introduced semantic web concepts have been proposed a solution for these data integration and interoperability issues. The semantic web is a new extension of the existing web, which offers a set of best practices for publishing and interlinking structured data on the Web. In the semantic web, data is published in a machine-readable and processable format, and data has explicit meaning to machines defined by ontologies, and the resulting datasets are interlinked [14]. In the CH domain, there are a few ontologies, which intends to facilitate the integration and interoperability of CH data. CIDOC-CRM is one of the well-known ontologies in the CH domain, which provides definitions and structures to describe concepts and relationships in CH data. This ontology is an ISO standard since 2006. Many large-scale and individual research projects have been applying CIDOC-CRM to facilitate semantic interoperability and integration among heterogeneous CH resources [119-121]. Nevertheless, applying CIDOC-CRM ontology into the CH data still requires solid technical expertise in CIDOC-CRM classes and properties in addition to semantic web concepts. Hence, mapping CH data into CIDOC-CRM ontology and publishing CH data as LOD is still a technical challenge for many CH organizations.

Interlinking parts of 3D CH models with other relevant CH resources: While CH professionals are now well-equipped with 3D acquisition tools and 3D modelling software, interlinking parts of 3D CH models with relevant CH resources is still a challenging task. 3D digital CH models constructed using surveying methods are usually represented as meshes. Although they look like a single 3D model, they can be divided into subparts. For instance, a 3D digital CH model of an ancient building might consist of parts such as a roof, walls, windows, and doors, etc. These subparts can, in turn, be interlinked with other semantic knowledge bases such as DBpedia, GeoNames among others. This possibility is extremely important as it allows users to get information about each part of the 3D model. In the past there have been several research projects related to interlinking parts of 3D models with relevant information [78,122,123], however, semantic interlinking of parts of 3D models with knowledge bases is remaining as a daunting task.

In the next research articles, the author will explore technical solutions to encode CH data into CIDOC-CRM ontology in addition to interlinking mapped CH data with other knowledge bases. Furthermore, the author also aims to provide a solution for part-based interlinking of 3D digital CH models with knowledge bases such as DBpedia, and GeoNames.

Chapter 5 A Cloud Architecture for Processing and Visualisation of Geo-located 3D Digital Cultural Heritage Models[76]

This chapter presents a novel AWS-based cloud architecture that integrates cloud computing concepts, a digital interactive map, and 3D digital cultural heritage models. More specifically, this cloud architecture extends the web platform presented in Chapter 4 and integrates it with cloud computing concepts.

Abstract: The increasing affordability of surveying methods such as laser scanning and photogrammetry has aroused broad and current interest in 3D modelling among cultural heritage preservation specialists. This generated, in recent years, many digital cultural heritage preservation projects across the globe that aimed at documenting cultural heritage sites and objects in a 3D form. Once 3D cultural heritage models have been created, the next step is generally to assure their long-term digital storage, dissemination, and visualization. To this end, this article presents a new cloud architecture for processing and visualization of geo-located 3D cultural heritage models over the web, which has been accomplished by integrating maps, 3D cultural heritage models, and the geospatial data associated with the location of 3D cultural heritage models. The cloud architecture is based on Amazon Web Services, while the core framework for handling the content is managed by free and open-source, database-driven, easy-to-implement KeystoneJS Content Management System. All other frameworks used in the architecture such as for web mapping, 3D visualization, etc. are also based on free and open-source paradigm, which allows flexibility on extensions and re-use. The proposed architecture has been validated through a use-case applied to Australian 3D cultural heritage models.

76 This chapter is a copy of the peer-reviewed conference article - Nishanbaev, I. A Cloud Architecture for Processing and Visualization of Geo-located 3D Digital Cultural Heritage Models. In Proceedings of the 6th International Conference on Geographical Information Systems Theory, Applications and Management, 7-9 May 2020; pp. 51–61. https://doi.org/10.5220/0009341500510061

5.1 Introduction

In recent years, 3D modelling has become a widespread tool among CH preservation specialists. One of the reasons for this is the increasing affordability of surveying methods such as laser scanning and photogrammetry, which can efficiently and accurately survey tiny cultural heritage (CH) objects as well as complex CH sites [102]. This generated, in recent years, many digital CH preservation projects across the globe to document CH objects and sites in a 3D form [104]. The resulting 3D CH models are often employed not only for preservation purposes but also for other purposes such as 3D Geographic Information Systems (GIS), augmented reality, and virtual reality among others [88]. Furthermore, there are cloud computing-based web repositories, web-GIS platforms, and archives are emerging for long-term storage, visualization, and analysis of 3D digital CH models. Cloud computing offers many benefits over traditional approaches (e.g., on-premises infrastructure) such as scalability, flexibility, and potential to reduce IT (information technologies) costs among others. Scalability of the cloud computing refers to the ability of the system in which every application or infrastructure can be scaled up and scaled down based on the workload [124-126]. While flexibility in cloud computing allows employees to be more flexible in terms of accessing files through web browsers, collaboration, access from different devices and location among others [127,128]. Finally, IT costs can be reduced through a reduction in spending on hardware, software, infrastructure, and IT staff among others [129,130].

Although there are some web platforms and archives for 3D content such as OMEKA, 3D-COFORM, Sketchfab, TurboSquid, 3D CH models are still problematic to find, use and re-use. A recent survey of Champion and Rahaman [17], who surveyed 14 proceedings of leading digital heritage events and conferences, revealed that out of 1483 examined conference articles only 17.9% or 264 articles incorporated 3D models or images of 3D models. Even more dramatic is only 9 articles had accessible 3D content. Furthermore, in most cases, 3D digital CH models incorporate geospatial information relating to the location of the CH site and object.

To this end, this article presents a new cloud-based architecture for integrating maps, 3D digital CH models and geospatial data such as geolocation. The cloud architecture is based on Amazon Web Services (AWS), while the core framework for handling the content is managed by free and open-source, database-driven, easy-to-implement KeystoneJS Content Management System (CMS). All other frameworks used in the architecture such as for web mapping, 3D visualization, etc. are also based on free and open-source paradigm, which allows flexibility on extensions and re-use. The proposed architecture has been validated through a use-case applied to Australian 3D CH models. This architecture can be used, or extended for use, in a wide range of domains including GIS.

The contributions of this article to the CH and GIS domains are as follows:

- A new AWS-based cloud architecture to integrate maps, 3D CH models and geospatial data such as geolocation. All employed frameworks in the architecture are free and open-source, and completely extendable and reusable. The proposed architecture can also be used in any other cloud platform, which supports Node.js environment.

- Since content management including uploading and editing text data, 3D CH models, and geo-location within the architecture is based on a database-driven, easy-to-implement KeystoneJS CMS, it can help CH and GIS professionals with limited technical knowledge to implement the architecture. Furthermore, many time-consuming web developments such as admin user interface, session management among others are handled by the KeystoneJS CMS.

- CH organizations and institutions such as museums, galleries, and archives can implement this architecture to present available CH objects and assets in a 3D digital form on the web. This can, in turn, help to promote national and international tourism.

- GIS and CH professionals can extend this architecture according to their needs. For instance, raster and vector data can be integrated with the free and open-source framework of GeoServer, which supports geospatial data in many formats and standards.

- Finally, the article extends the existing body of knowledge and expertise in implementing a cloud-based architecture for processing and visualization of geo-located 3D digital models.

5.2 Related work

Europeana is a CH platform and aggregator with more than 30 million digitized objects from more than 2300 European CH institutions and organizations, while Europeana cloud is one of the largest projects regarding cloud-based infrastructures for CH research and data. Europeana Cloud is an initiative by Europeana Foundation to provide shared cloud infrastructure for aggregating and exchanging CH data among European institutions and organizations. To this end, it offers many services such as unique identifiers to each CH records, storage and access for heterogeneous CH data including metadata, annotation services to add additional data to CH records, tracking of changes made to CH records, flexible, scalable and customizable CH data processing capabilities among others. Europeana cloud, in terms of technical architecture, has been implemented on a hybrid cloud architecture by combining public cloud and private cloud architectures. This allows taking all the advantages offered by these two cloud architectures. For instance, public cloud offers flexibility for scaling up and down the hardware capabilities based on the workload, while the private cloud offers more resistance to failures and less dependency on third-party cloud providers to name a few [131,132].

The SACHER [133] (The Smart Architecture for Cultural Heritage in Emilia Romagna) is a project financed by Regione Emilia-Romagna within the European Regional Development Fund. It has developed a cloud-based, open-source and federated platform to manage various aspects of tangible CH such as 3D life cycle management for CH, the multi-dimensional search engine to find CH data from heterogeneous sources among others. The platform offers services both to CH professionals and the public. The cloud platform used in the project is IaaS (Infrastructure-as-a-service) OpenStack, which is often used to build private and public clouds. Regarding the server-side web technology, the project uses the Django web framework, which is written in Python programming language. Swift Object Storage was employed as a storage container for 3D models, while 3D Heritage Online Presenter (3DHOP) was chosen as a visualization framework for 3D models. The multi-dimensional search engine in the platform is based on a NoSQL database of MongoDB. The platform also incorporates a Google map through which CH places can be searched by name, address and building type [134,135].

CNR ITAB (Institute for technologies applied to cultural heritage), while collaborating with ARIADNE [136] and E-RIHS [137] infrastructure projects, developed a cloud-based modular architecture to enable archeologists to build and visualize 3D landscapes. This architecture was built on a cloud platform of ownCloud [138], which comes in three versions namely ownCloud Community, ownCloud Online, and ownCloud Enterprise. In all versions, it provides a free desktop client and iOS and Android app to upload and manage the data. The developed cloud architecture offers several services such as terrain service and gallery service. The former leverages on WebGL based Virtual Planet Builder [139] tool to load, process and visualize GIS elevation data such as Digital Elevation Models (DEM). The latter service allows users to create, delete, and update projects within the system. It displays all created terrain databases, status, and if available, a short description for each entry among others [140].

Furthermore, there are some other smaller-scale projects in this regard such as a collaborative project by the National Library of Scotland, Edinburgh Parallel Computing Centre, the National Galleries of Scotland and the Digital Preservation of Coalition. The name of the project is Cloudy Culture, which aimed to investigate the potential of EUDAT [141] cloud services in CH conservation, particularly to improve preservation and accessibility of European CH data hosted in the National Library of Scotland, and the National Galleries of Scotland. These two CH institutions collect and host an extremely large amount of digital CH data such as maps, books, articles, images among others, which require safe digital preservation [142].

Another relating project is by Pisu and Casu [143], which proposed a cloud-based web-GIS framework for documentation and dissemination of architectural heritage. The web-GIS framework resulted in a multi-scale and multi-layer information system applied to Sardinian late gothic architecture.

5.3 Background literature

5.3.1 Cloud computing

Cloud computing is the hosting and the delivery of various services over the Internet. It can include tools and resources such as servers, databases, networking, various business applications and software, which can be leased in an on-demand fashion. Cloud computing has been given many definitions, however, many researchers and professionals in this domain have adopted a definition by the National Institute of Standards and Technology of the United States (NIST), which states it as follows:

'Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.'

Cloud computing services can be grouped into three categories namely software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS).

SaaS refers to cloud computing that offers on-demand applications over the Internet. An apt example of this category of cloud computing is Salesforce [144], which offers a wide range of services to business owners (customer relationship management) to better connect with customers, potential customers, and partners.

PaaS refers to cloud computing that offers platform resources such as software development frameworks and tools. An example of this category of cloud computing is Heroku [145]. It provides software developers with virtual containers called ‘Dynos’ that can execute software applications written in various programming languages.

IaaS refers to cloud computing that offers on-demand infrastructural resources over the Internet. Examples of this category of cloud computing include Amazon Elastic Compute Cloud - Amazon EC2 [146], Google Compute Engine [147] and Digital Ocean [148] among others. This cloud computing type allows launching on-demand virtual machines (VMs), which can then be accessed to perform computing tasks.

Cloud computing also offers several cloud deployment models, which define how cloud services are made available to users. Public cloud, private cloud, hybrid cloud, and community cloud are 4 common deployment models associated with cloud computing.

Public cloud is the type of cloud computing model that supports all users who want to use cloud computing resources such as hardware and software on a subscription basis, in other words, everyone who subscribed can use the services. It is often used for application development and testing, file sharing and other purposes, in which privacy and security of data are not a high priority.

On the other hand, the private cloud is typically used by a single organization. This type of cloud computing deployment can be hosted internally or externally. This type of cloud deployment offers greater control over cloud infrastructure and a higher level of security compared to public cloud deployment.

Hybrid cloud as the name suggests combines the best of the private and the public cloud. Many organizations use this type of cloud deployment as it allows to take advantage of scalability and cost-effectiveness of public cloud as well as to execute and to store mission-critical applications and data in the private cloud.

In the community cloud deployment model, multiple organizations share computing resources. Hence, the hardware and software capabilities of the cloud are managed and secured commonly by all the participating organizations. Since access to the cloud is restricted to the members of the community, many research organizations and universities often deploy this type of cloud model to conduct collaborative research projects [149-151].

5.3.2 Web content management systems, 3D visualisation frameworks and 3D file formats

In recent years, the amount of digital content available on the Internet has increased significantly. Many organizations and individuals are, therefore, deploying web content management systems (WCMS) to provision and manage their information on the web. In essence, WCMS is a software application that can help people with limited technical knowledge to create and manage websites, and web content. WCMS often consists of two parts namely front-end and back-end. The former represents the web user-interface that users see when they visit the website. While the latter represents the server-side, in other words how the website works, changes, and updates. This usually involves databases and servers. WCMS may offer organizations and individuals many benefits such as user-friendly customization of the front-end, user-friendly web content management and editing, and easy-to-follow workflows for search engine optimizations (SEO) among others. Furthermore, there are many free and open source WCMSs available on the Web such as WordPress, Drupal, Joomla, KeystoneJS to name a few [118,152].

Currently all modern browsers such as Google Chrome, Opera, Safari and Firefox support WebGL technology that allows creating and visualization of 3D graphical applications on the Web. This technology enables web users to experience interactive 3D content on webpages without downloading and installing any plug-ins. WebGL was originally developed by Mozilla, however, currently it is maintained by the non-profit organization called Khronos Group. Since WebGL leverages two hardware components of a computer namely central processing unit (CPU) and graphics processing unit, it offers GPU acceleration to execute large-scale 3D web applications. Thus, it provides improved performance and faster running of applications. The part of the technology that runs on CPU is written on JavaScript programming language, while the GPU part is based on OpenGL ES. Despite the fast execution of 3D web applications, WebGL is a low-level 3D graphics application programming interface (API) [153]. This means developing 3D web applications using this technology can be time-consuming, and in some instances extremely complicated. For this reason, many WebGL-based JavaScript libraries have been developed to ease and accelerate the development of 3D web applications. Three.js, Babylon.js, sceneJS are among those libraries, which is widely used to develop interactive 3D web applications. These libraries offer many technical features such as effects, lights, various shaders, and virtual reality among others. They run on all web browsers that support previously discussed WebGL technology. CesiumJS is another 3D library used for developing interactive 3d maps. This library is now an open-standard of Open Geospatial Consortium (OGC).

3D file formats are used to store information relating to a 3D object such as geometry, appearance, scene and animation. 3D file formats are divided into two types in terms of licencing, which are proprietary and non-proprietary. Proprietary formats are native file formats from software applications, whereas non- proprietary formats are open-source file formats. gLTF, Collada, 3D Tiles, and OBJ are some of the well-known examples of 3D file formats.

5.4 Methodology

To accomplish the proposed cloud architecture, a methodology as shown in Figure 25 has been developed. The web application server KeystoneJS CMS and front-end application in the methodology are based on the previous research work of the author. The methodology presented in the article extends those parts with cloud computing concepts such as Amazon EC2 instance, Amazon EBS storage for 3D CH models, Elastic Load Balancing, PM2 process manager, web server of NGINX, cloud installed MongoDB database among others. This extended methodology consists of the followings:

- Cloud computing infrastructure from AWS including Amazon Elastic Compute Cloud (AWS EC2), Red Hat Enterprise Linux as an operating system, Web Application Server – Keystone JS CMS, MongoDB database, Amazon Elastic Block Store (Amazon EBS), PM2 - Process Manager, Web Server – NGINX, and Elastic Load Balancing.

- Front-end application for administrators which includes the KeystoneJS authentication system and KeystoneJS provided back-end interface with custom fields.

- Front-end application for users which includes Three.js – WebGL-based 3D JavaScript library, OpenLayers 6 – web mapping library, Bootstrap 4 for the graphical user interface, and jQuery

- Datasets including maps, Australian 3D CH models, 3D model descriptions and geospatial data in the form of geolocation such as longitude and latitude.

- Finally, users and administrators, who consume the data and manage the data respectively.

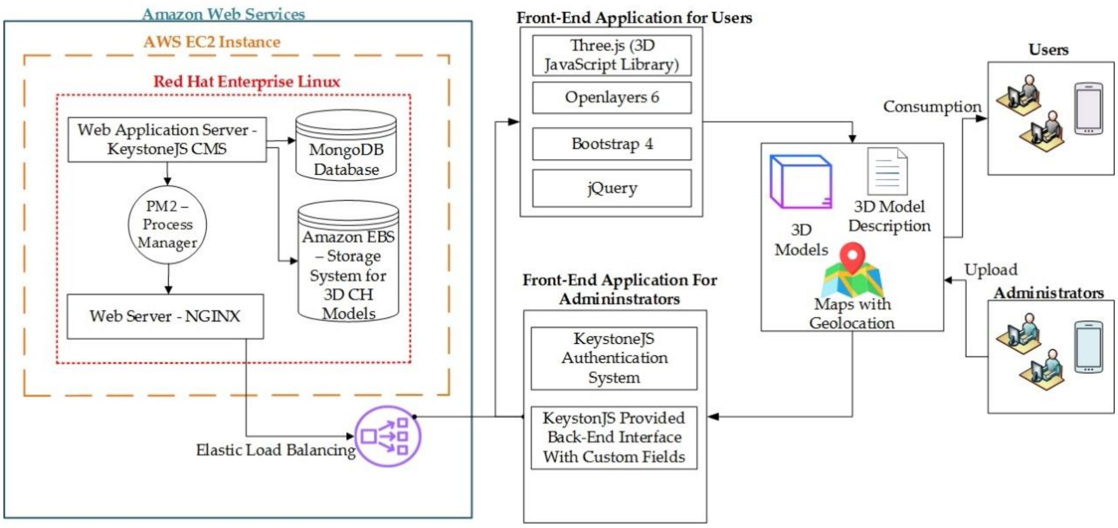


Figure 25. A methodology for the proposed cloud architecture for processing and visualization of 3D digital cultural heritage models.

5.4.1 Cloud computing

AWS has been selected as a cloud computing solutions provider as it is dominant and the largest provider in the cloud computing market. Furthermore, it offers a variety of services and virtual machines for various purposes such as from hosting a small website to Big Data analytics. In the architecture, EC2 has been deployed to get the cloud computing capacity and launch the virtual server. As an operating system of the EC2, Red Hat Enterprise Linux has been deployed which is a Linux distribution developed by Red Hat for the commercial market. This operating system was chosen because it is one of the leading Linux kernels with high security, reliability, and good community support. The web application server in this architecture is KeystoneJS CMS, which is based on Express.js framework and Node.js runtime environment. Since KeystoneJS is a database-driven CMS, it comes with a MongoDB database by default. This database has been utilized to store all the data within the architecture except for 3D CH models. The storage of 3D models has been facilitated by Amazon Elastic Block Store (EBS), which is a high-performance storage service to use with Amazon EC2. Another important feature of this storage is the possibility to enable the automated capability to back up the data into Amazon S3. NGINX has been used as a web server in the architecture, whereas PM2 Process Manger used to manage and monitor the web application of KeystoneJS. Finally, Elastic Load Balancing is used to automatically distribute incoming traffic to the web application, which helps to improve the responsiveness of the web application.

5.4.2 Front-end application for administrators

Front-end application for administrators has been accomplished with the auto-generated admin user interface by the KeystoneJS. It also handles the authentication system for administrators out of the box. KeystoneJS allows creating as many administrators as needed. However, one limitation of KeystoneJS in this regard is it does not provide a feature to assign roles to the administrators, which can grant or restrict access to certain fields and documents within the admin interface. Text data in the administrator interface is inputted via WYSIWYG (what you see is what you get) editor, while files such as 3D CH models are uploaded via the built-in file upload interface. For the uploaded content, there are three states available namely draft, published, and archived.

5.4.3 Front-end application for users

Front-end application for users was achieved using Three.js, OpenLayers 6, Bootstrap 4 and jQuery. Three.js is a cross-browser 3D JavaScript library based on previously mentioned WebGL technology, which facilitated 3D visualization in the architecture. It offers numerous features to build simple and complex 3D scenes and worlds. This 3D library has been chosen because of its non-proprietary license, integration with all web browsers that support WebGL, and relatively well-documentation among others. OpenLayers 6 is used for web mapping in the architecture, which can visualize maps including geospatial data such as vector data and raster data in different geospatial files from various sources and geospatial web mapping standards of Open Geospatial Consortium (OGC). Since OpenLayers is an open-source, mature, web mapping library with a strong community of developers, it has been selected for map and web geospatial data visualization. Bootstrap 4 framework is used as a helper framework for graphical user interface implementation. It is the world's most popular framework for building the front-end development of web applications, which is also known for its responsive and mobile-first concepts. Since Bootstrap focuses on simplifying the development of front-end web applications, it offers several ready-to-use templates and components such as navigation, forms, and typography. Finally, the jQuery framework is a JavaScript library for simplification of HTML DOM manipulation, event handling, and asynchronous calls. As of writing this article, it is employed by 74% of the 10 million most popular websites [154].

5.4.4 Datasets

As a validation of the architecture, a use case with Australian 3D CH models has been deployed. In total eleven different 3D digital models have been used in this use case. These 3D models are 3D digital replicas of the places located in Australia. Four of these 3D models are presented as a sample in the article namely Time Ball Tower-Williamstown, Cape Liptrap, Magnetic Termite and a part of the Abrolhos Islands. The images of these places are shown in Figure 26.

Figure 26. Sample of the employed 3D digital models in the use case.
In a clockwise direction, Time Ball Tower, Williamstown (Image Source: https://bit.ly/2FfKuuT, Copyright by Nick Morieson. CC BY-SA 2.0), Cape Liptrap (Image Source: https://bit.ly/2MTdNrr, Copyright by Jorge Lascar. CC BY 2.0), Magnetic Termite (Image Source: https://bit.ly/2ZO5an6, Copyright by Geoff Whalan. CC BY-NC-ND 2.0), Abrolhos Islands (Image source: https://bit.ly/37qOPaG, Copyright by ernie_greatoutdoors. CC BY-SA 2.0)

All 3D models are encoded in a 3D file format of glTF, which is one of the best choices for the web as it significantly minimizes the size and loading times of 3D models in comparison to other 3D file formats. The architecture has also been tested with another 3D file format of "OBJ, which is an open format by Wavefront technologies. This file format usually incorporates three files, which are a 3D object with the ".obj" file extension, a material template library with the “.mtl” file extension, and one or more image files (with the file extension of “.jpeg”, ”.png”, etc). The first file with “.obj” extension stores 3D object-related information such as vertices, faces in a plain text, while the second file with the “.mtl” file extension stores a material that is appearance of the 3D object. This file also stores the path link to the image files.

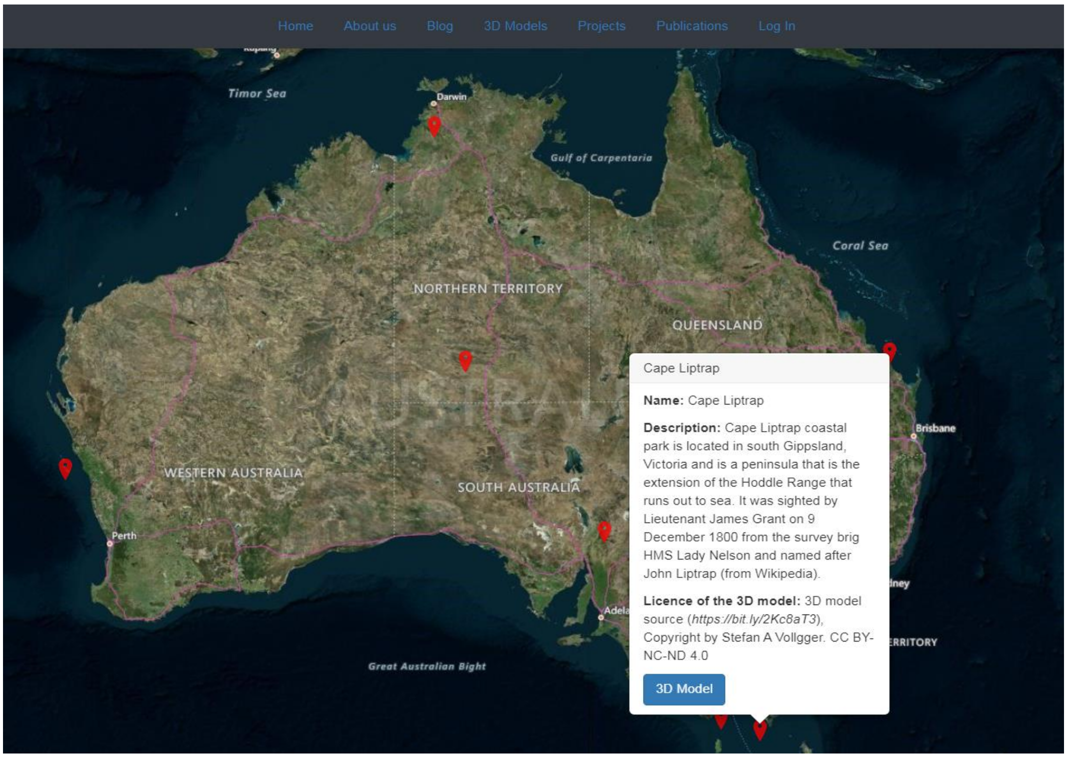


Figure 27. The web interface for users.
The map copyright by Microsoft Bing Maps 2019.

5.4.5 Users and administrators

Users of this architecture platform can view a map, interact with vector data in the form of point features which represent CH places, read descriptions of CH places, and view 3D CH models. Whereas, administrators can create and edit CH places, upload 3D content to the platform, input geospatial data such as geolocation associated with CH places, select and change the state of the content among others.

5.5 Results and discussions

The cloud architecture integrates web maps, 3D digital CH models and geospatial information associated with the 3D CH models in the form of geolocation. The overall web interface for users consists of three main components namely base map, clickable red markers which represent CH places with relevant CH information, and viewer for visualization of 3D CH models. This user interface is shown in Figure 27. Once a user clicks on a red marker, the pop-up box appears with the information relating to the CH place such as name, short description, and licence information of the 3D model. To view the 3D model, they should click onto the 3D model button in the pop-up box.

In KeystoneJS CMS there is no limitation on the number of administrators. Hence, it allows creating as many administrators as needed. A page relating to creating a new CH place and uploading 3D content to the architecture is shown in Figure 28. When creating a new CH place, administrators should input the name of the CH place, description including short information and licence information of the CH place, and 3D digital CH model to upload. 3D digital CH models are automatically uploaded into Amazon EBS storage and path to the 3D content is stored in the MongoDB database. This path is then used in the front-end application to retrieve the 3D CH model. The administrators can also select the state of the CH place from 3 options namely draft, published, archived. These states can help to manage the process of publication. For instance, they may upload the content but wish to publish it later in time. Finally, for each and every CH place administrators should input geolocation in the form of longitude and latitude. The precise geolocation for the most CH places can be found in the GeoNames platform, which is a geographical database with more than 11 million place names.

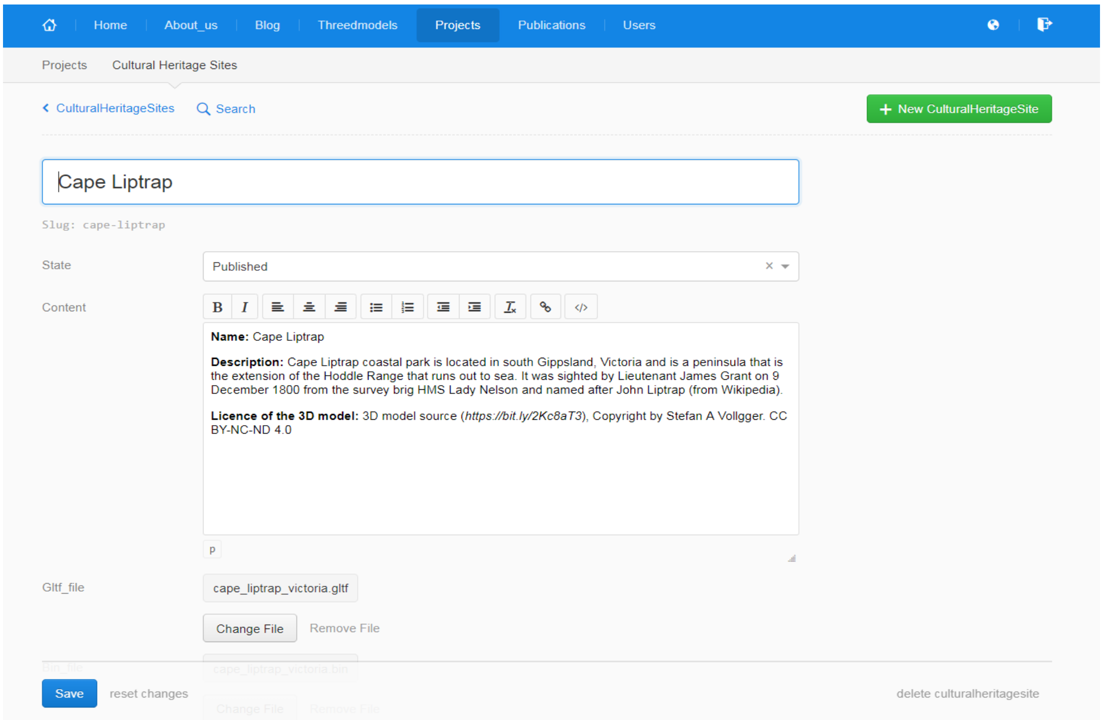


Figure 28. The web interface for administrators.
Copyright by KeystoneJS, MIT licence 2019.

In a clockwise direction, Time Ball Tower, Williamstown (3D Model Source: https://bit.ly/2ZJR3PZ, Copyright by Digital Heritage Australia. CC BY 4.0), Cape Liptrap (3D Model Source: https://bit.ly/2QFMwtl, Copyright by Stefan A. Vollgger. CC BY-NC-ND 4.0), Magnetic Termite (3D Model Source: https://bit.ly/2syNqjO, Copyright by Rupert Rawnsley. CC BY 4.0), Abrolhos Islands (3D Model Source: https://bit.ly/2syY9e3, Copyright by Maritime Archeological Association of Western Australia. CC BY 4.0)

5.5.1 Use case of Australian 3D digital cultural heritage models

The architecture has been employed in a sample use case to better evaluate its benefits and limitations. As mentioned in the datasets section, eleven different 3D digital CH models have been used in this use case. The visualization of four of these 3D CH models is shown in Figure 29, which is achieved using the Three.js 3D visualization library. The 3D models can be manipulated along the X, Y and Z coordinate axes, and can be zoomed in and out.

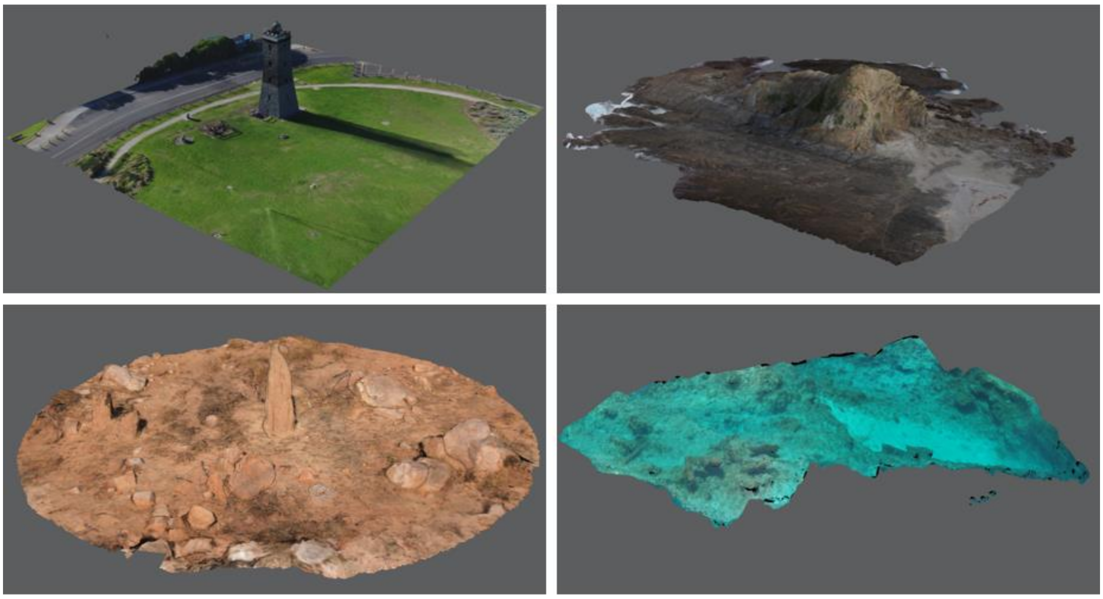


Figure 29. Use case of Australian 3D cultural heritage models.

5.5.2 Level of customization for the front-end 3D visualization

As mentioned previously, in this architecture 3D visualization is achieved using Three.js library, which offers plenty of customizations. These customizations include importing and removing complementary 3D models at run-time, various cameras such as orthographic, perspective, and stereo cameras, various audio types such as positional and non-positional, various lights such as point, directional, and spotlights, and three levels of details (LOD) among others. These customizations can take effect at run-time as Three.js is capable of updating scenes continuously.

5.5.3 Scalability of the architecture

The proposed architecture is based on AWS, which offers many features for scalability, load balancing, etc. The architecture particularly uses Elastic Load Balancing to automatically distribute incoming traffic. Whereas KeystoneJS CMS handles internal tasks within the web application. Since KeystoneJS CMS is based on the Node.js runtime environment, it supports an event-driven system capable of running asynchronous input and output. In other words, it can run several processes separately from the primary application thread. Therefore, this offers numerous benefits such as improved application performance, scalability, and enhanced responsiveness.

5.6 Conclusion

3D surveying tools and techniques now allow CH professionals to 3D digitize tangible CH ranging from the smallest CH artefact to the largest CH building. This rapid evolution in technology has significantly advanced the conservation of CH in a digital form. However, the lifetime of 3D CH models is short, especially those which resulted from the individual CH research projects. According to the survey findings of Champion and Rahaman [17], who looked at 1483 digital heritage articles published in 14 recent proceedings of major CH conferences, from the examined 1483 conference articles, only 9 articles had accessible links to 3D assets.

This article presented a new cloud-based architecture for integrating maps, 3D digital CH models and geospatial data, which can be used for processing, visualization, dissemination and digital preservation of 3D digital models.

All frameworks used in the architecture such as for content management, web mapping, 3D visualization, etc. are based on free and open-source paradigm, which allows flexibility on extensions and re-use. For instance, raster and vector data can be integrated into the architecture by implementing a free and open-source geospatial framework of GeoServer, which allows publishing geospatial data to the web using open standards such as from OGC. Furthermore, among others, on top of this architecture geospatial analyses on the web can be performed using Turf.js, web virtual reality can be implemented using React 360.

The proposed architecture has been validated through a use-case applied to Australian 3D CH models.

Chapter 6 A Web GIS-based Integration of 3D Digital Models with Linked Open Data for Cultural Heritage Exploration[77]

This chapter further extends the web platform and the methodology presented in Chapter 5 and presents a novel cloud-based web-GIS architecture that integrates cloud computing, 3D digital models, and linked open data from DBpedia and GeoNames to enhance cultural heritage exploration. Furthermore, it presents recommendations and suggestions on the platform, which were collected from cultural heritage researchers.

Abstract: In recent years, considerable efforts have been made by cultural heritage institutions across the globe to digitise cultural heritage sites, artifacts, historical maps, etc. for digital preservation and online representation. On the other hand, ample research projects and studies have been published that demonstrate the great capabilities of web-geographic information systems (web-GIS) for the dissemination and online representation of cultural heritage data. However, cultural heritage data and the associated metadata produced by many cultural heritage institutions are heterogeneous. To make this heterogeneous data more interoperable and structured, an ever-growing number of cultural heritage institutions are adopting linked data principles. Although the cultural heritage domain has already started implementing linked open data concepts to the cultural heritage data, there are not many research articles that present an easy-to-implement, free, and open-source-based web-GIS architecture that integrates 3D digital cultural heritage models with cloud computing and linked open data. Furthermore, the integration of web-GIS technologies with 3D web-based visualisation and linked open data may offer new dimensions of interaction and exploration of digital cultural heritage. To demonstrate the high potential of integration of these technologies, this study presents a novel

---

77 This chapter is a copy of the peer-reviewed research article - Nishanbaev, I.; Champion, E.; McMeekin, D. A. (2021). A Web GIS-Based Integration of 3D Digital Models with Linked Open Data for Cultural Heritage Exploration. ISPRS International Journal of Geo-Information, 10(10), 684. https://doi.org/10.3390/ijgi10100684

cloud architecture that attempts to enhance digital cultural heritage exploration by integrating 3D digital cultural heritage models with linked open data from DBpedia and GeoNames platforms using web-GIS technologies. More specifically, a digital interactive map, 3D digital cultural heritage models, and linked open data from DBpedia and GeoNames platforms were integrated into a cloud-based web-GIS architecture. Thus, the users of the architecture can easily interact with the digital map, visualise 3D digital cultural heritage models, and explore linked open data from GeoNames and DBpedia platforms, which offer additional information and context related to the selected cultural heritage site as well as external web resources. The architecture was validated by applying it to specific case studies of Australian cultural heritage and seeking expert feedback on the system, its benefits, and scope for improvement in the near future.

6.1 Introduction

Tangible cultural heritage (e.g., monuments, sites, artifacts, historical maps) is often exposed to a variety of natural disasters such as earthquakes and flooding, as well as human-made hazards such as fires or human negligence. Digital preservation of tangible cultural heritage is, therefore vitally, important to ensure that the shape and texture of the cultural heritage are not lost in case of damage by natural or human-made disasters [1]. Hence, in the last 20 years, considerable efforts have been made by cultural heritage institutions across the globe to digitise cultural heritage sites, artifacts, and historical maps, etc. for digital preservation and online representation [2,3].

On the other hand, ample research projects and studies were published that demonstrate the high potential of web-geographic information systems (web-GIS) for the dissemination and online representation of cultural heritage data [4,5]. This high potential of web-GIS in the dissemination and online representation of cultural heritage data is mainly due to the fact that the data derived from monuments, sites, and artifacts often contain spatiotemporal information such as geolocation on the Earth and time references, two critical data qualities for any web-GIS system. However, cultural heritage data and the associated metadata produced by cultural heritage institutions are heterogeneous in different file formats and languages, typically interconnected in content, and are distributed in different web services across cultural heritage institutions and even countries [155].

To make these heterogeneous data more interoperable, an ever-growing number of cultural heritage institutions are adopting linked data principles [156,157]. Linked data refer to data that are published in a way that is machine understandable; the data have explicit meaning to machines defined by ontologies, and the data are linked to external data sources and can also be linked from external data sources [14]. As an early adopter of various digital technologies and tools for digitisation, management, dissemination, and analysis, the cultural heritage domain has already implemented domain-specific ontologies such as CIDOC-CRM[78] and many cultural heritage institutions have already successfully adopted linked data publishing principles [158-160].

There have also been some relatively large-scale projects that offer a purpose-built platform for publishing cultural heritage data as linked data such as Wisski[79] Arches[80], ResearchSpace[81], Omeka S[82], etc. Furthermore, in a previous study, the authors compared and evaluated the features and functionality of the current state-of-the-art resource description framework (RDF) generation tools and interlinking frameworks, provided a comprehensive methodology for generating geospatial linked open data, and demonstrated how the geospatial linked open data could be interlinked with other external RDF data sources such as DBpedia [161]. The above-mentioned platforms, including my previous study, essentially deal with publishing data according to linked data principles, as well as an online representation of linked data. Although a cultural heritage domain has already started implementing linked open data concepts for the display of cultural heritage data, there are not many research studies presenting an

78 http://www.cidoc-crm.org/ (last accessed on 29 July 2021)
79 http://wiss-ki.eu/ (last accessed on 29 July 2021)
80 https://www.archesproject.org/ (last accessed on 29 July 2021),
81 https://www.researchspace.org/ (last accessed on 29 July 2021),
82 https://omeka.org/ (last accessed on 29 July 2021)

easy-to-implement, free, and open-source (FOSS)-based web-GIS architecture that integrates 3D digital cultural heritage models with cloud computing and linked open data.

This study presents a novel cloud architecture (https://unesco-chv.curtin.edu.au/mapplatform, last accessed on 5 October 2021) that attempts to enhance digital cultural heritage exploration by integrating cloud computing concepts, digital interactive map, linked open data from the well-known RDF knowledge base of DBpedia, and the geographical database of GeoNames, along with locally stored 3D digital cultural heritage models and the associated metadata.

Thus, the architecture promotes the re-use of the existing data, which is one of the four principles of FAIR[83] data (findable, accessible, interoperable, and reusable). The conceptual design of the architectures is shown in Figure 30. As can be seen from Figure 30, through the web-GIS architecture, the users can interact with the digital map, visualise 3D digital cultural heritage models, and explore linked open data extracted from GeoNames and DBpedia platforms, offering additional information and context related to the selected cultural heritage site or artefact, including information from external web resources.

[83] https://www.ands.org.au/working-with-data/fairdata (last accessed on 29 July 2021)

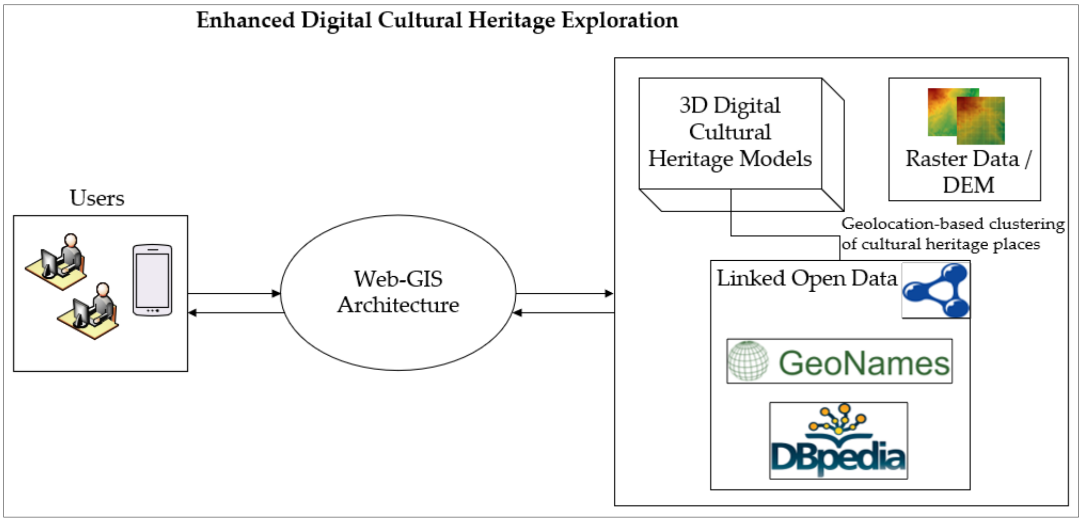


Figure 30. The conceptual design of the architecture.

Regarding the technical aspects of the architecture, first, the architecture extracts the geolocation of cultural heritage places from the DBpedia RDF data (Australian World Heritage Sites and Australia's National Heritage List), the GeoNames RDF data (any Australian cultural heritage site searched by the user), and the locally stored 3D digital cultural heritage models.

Second, if the cultural heritage places retrieved from DBpedia and the locally stored 3D digital cultural heritage models are overlapped or within 3 metres from each other, it clusters them into a single clickable marker on the digital interactive map. If the retrieved cultural heritage places and the locally stored 3D digital cultural heritage models are not overlapped and more than 3 metres away from each other, they are displayed separately. This 3-metre distancing for clustering was intentionally implemented as the geolocations might contain minor inaccuracies.

Additionally, it is worth mentioning that this 3-metre distance is flexible and can easily be increased or decreased. As for the cultural heritage places retrieved from the GeoNames platform, they are placed on top of the other markers. This was also purposely designed in this way to make cultural heritage places retrieved from the GeoNames search result easily noticeable on the interactive map. The architecture utilises web-GIS technologies to offer a digital interactive map to users and a webGL-based 3D visualisation library (three.js) for visualisation of 3D digital cultural heritage models. For storing 3D models, the architecture uses Amazon Elastic Block Store (EBS) storage of the AWS, while metadata about the 3D digital cultural heritage models such as cultural heritage place name, geolocation in the form of latitude and longitude, basic additional text information about the cultural heritage place, and the Uniform Resource Locator (URL) path to the 3D model on the EBS are stored in the NoSQL database of MongoDB.

The architecture was validated by applying it to typical use cases of Australian cultural heritage. Moreover, the study presents some expert suggestions and recommendations that may also be helpful to other cultural heritage researchers who are working on similar research topics related to digital cultural heritage exploration and dissemination.

The architecture presented in this article extends previous research of the authors on the cloud-based platform for visualisation and dissemination of digital 3D cultural heritage models [162]. Thus, the 3D visualisation part in the presented architecture was derived from the previous work of the authors. More specifically, this study extends the previous work of the authors by providing the following features:

- A novel cloud architecture that integrates digital interactive map and linked open data from DBpedia and GeoNames platforms with the locally stored 3D digital cultural heritage models. Thus, the users of the architecture can easily interact with the digital map, visualise 3D digital cultural heritage models, and explore linked open data from GeoNames and DBpedia platforms, which offers additional information related to the selected cultural heritage place, including external web resources. The architecture is facilitated by utilising FOSS frameworks except for cloud hosting. Since the architecture consists of FOSS frameworks, it allows an easy and cost-free replication of the architecture for CH institutions such as museums, galleries, and libraries.
- A completely re-designed digital interactive map that includes updated map controls, clustering of overlapped markers on the map, and

visualisation of raster data in the form of a digital elevation model, among others.

- Recommendations and suggestions on the platform, which were collected from cultural heritage researchers. In total, seven cultural heritage researchers provided recommendations and suggestions, which may also be helpful to other cultural heritage researchers working on similar research topics related to digital cultural heritage exploration and dissemination.

6.2 Related work

In recent years, there have been a few related projects that aimed to enhance the digital exploration of cultural heritage by utilising linked open data and other different technologies such as natural language processing, 3D visualisation, and Heritage Building Information Modelling (H-BIM), to name just a few. Cultural Heritage Conversational Agent (CulturalERICA) [163] is a great example that aimed to enhance digital exploration of cultural heritage by deploying natural language interaction technologies. It offers a natural language-based interaction with the Europeana semantic web knowledge base. The authors of the CulturalERICA project believe this project assists users in exploring digital cultural heritage content from Europeana.

Europeana[84] is a platform created by the European Union that contains digital cultural heritage data from more than 3,000 cultural heritage institutions across Europe. CulturalERICA project achieved a natural language interaction interface by deploying a Dialogflow natural language platform provided by Google, while Europeana REST API was used to retrieve the data from the Europeana platform. A custom node.js web service was implemented which works as a middleware application between the Dialogflow agent and the Europeana platform.

84 https://www.europeana.eu/en (last accessed on 29 July 2021

Another related large-scale project is INCEPTION—Inclusive Cultural Heritage in Europe through 3D Semantic Modelling (2015 - 2019) [79], which was a project funded by the European Union within the Horizon 2020 program. One of the aims of this project was to semantically enrich 3D digital cultural heritage models using H-BIM and semantic web technologies.

Another relatively recent project is 3D-IMP-ACT[85], which implemented an architecture that offers a web-GIS platform. The web-GIS platform integrates 360 panoramic images and 3D models for digital cultural heritage exploration and tourism [164]. According to the authors of the 3D-IMP-ACT, the project creates a "virtual network" of ancient architectures and sites. This project was funded under the program IPA CBC Interreg-Albania-Montenegro. The web-GIS platform in the project employed QGIS Server and PostGIS database at the server-side of the system. QGIS is an open-source framework that offers GIS logic and map rendering, while PostGIS is an open-source plugin for the PostgreSQL database. The front-end rendering of the map was achieved using the Lizmap[86] application, which is a front-end open-source application for the QGIS server.

There were also other projects, such as the study by Candela*, et al.* [165] or Quattrini*, et al.* [166] that demonstrated how linked open data principles, together with other technologies such as H-BIM, could be applied to enhance cultural heritage exploration. However, this work differs from previous works insofar as it provides a novel architecture utilising cloud computing concepts, web-GIS, linked open data from DBpedia and GeoNames platforms, and 3D digital cultural heritage models to enhance cultural heritage exploration, as explained in the previous section.

6.3 Literature review

In the following paragraphs, we present background literature relating to DBpedia and GeoNames platforms. DBpedia is one of the preeminent examples of linked open data, while GeoNames is a geographical database

85 https://3dimpact.italy-albania-montenegro.eu/ (last accessed on 29 July 2021)
86 https://www.lizmap.com/, last accessed on 29 July 2021

that provides geolocation-related information about places. However, at the time this article was written, the GeoNames platform also provided information about places in the linked-open-data-compatible format of RDF. Interested readers who would like to know more about the semantic web, ontologies, and the geospatial semantic web concepts pertinent to the cultural heritage field may refer to the previous work of the authors [101].

6.3.1 DBpedia

DBpedia is a crowd-sourced community project that extracts structured content from Wikipedia and serves them as linked data. It extracts Wikipedia articles and converts them into a single shared DBpedia ontology complaint format, which has more than 320 classes and 1650 properties. The project provides downloadable DBpedia data also known as DBpedia knowledge bases. It also offers a SPARQL endpoint that can be used to query the DBpedia data using a standardised SPARQL query language.

One of the greatest advantages of DBpedia over Wikipedia is Wikipedia only offers free text-based search against Wikipedia articles, whereas DBpedia offers a SPARQL endpoint that can answer complex queries [167,168]. For instance, using a Wikipedia search, it is very difficult to find all the cultural heritage places in Australia that are located within 200 km from the coast, whereas this could be easily achieved by sending a custom query to the SPARQL endpoint of the DBpedia.

Additionally, a SPARQL query language allows sending queries to multiple SPARQL endpoints within a single query statement, which is also known as a federated query. This allows merging content from different SPARQL endpoints to provide more meaningful and contextual information to users. A basic technical diagram of the DBpedia architecture is shown in Figure 31. As can be seen from Figure 31, DBpedia is deployed on top of the Virtuoso Universal Server which stores RDF data in the RDF Quad Store and offers a SPARQL endpoint alongside HTML browser and RDF browser support.

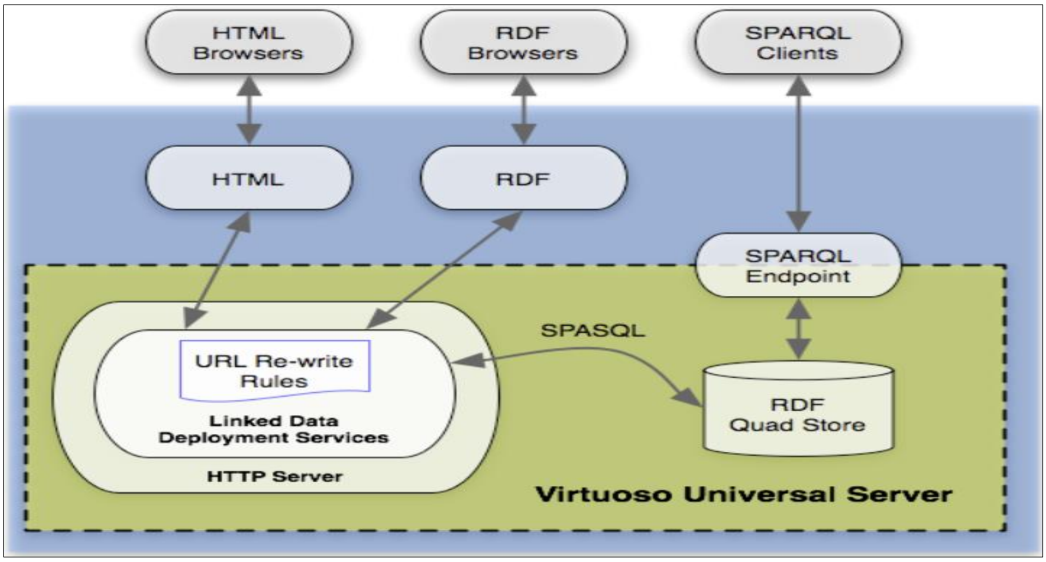


Figure 31. DBpedia architecture.
Source: https://www.dbpedia.org/about/, accessed on 29 July 2021.

6.3.2 GeoNames

GeoNames[87] is a geographical database that covers all counties in the world and contains over 11 million placenames. These placenames can be downloaded and used for free under a creative commons licence. It also offers a Representational State Transfer Application Programming Interface (REST API) which can accept a request over the Internet and respond in different file formats including in the RDF data model. Nevertheless, GeoNames REST API has a daily and hourly free usage limit per application. At the time of writing this article, GeoNames does not support a SPARQL endpoint; however, as mentioned previously, it can return the requested data in the RDF data model. Since GeoNames data sources have been interlinked with the DBpedia, most of the data in GeoNames contain RDF links that point to the relevant DBpedia URI. However, there are some placenames that are available in GeoNames but are missing in DBpedia. In these placenames, there is no RDF link to the DBpedia URI.

---

[87] https://www.geonames.org/ (last accessed on 29 May 2021)

6.4 Materials and methods

To accomplish the cloud-based integration of linked open data from DBpedia and GeoNames platforms with the locally stored 3D digital cultural heritage models, the architecture shown in Figure 32 was implemented. The next paragraphs elaborate on this architecture in more detail.

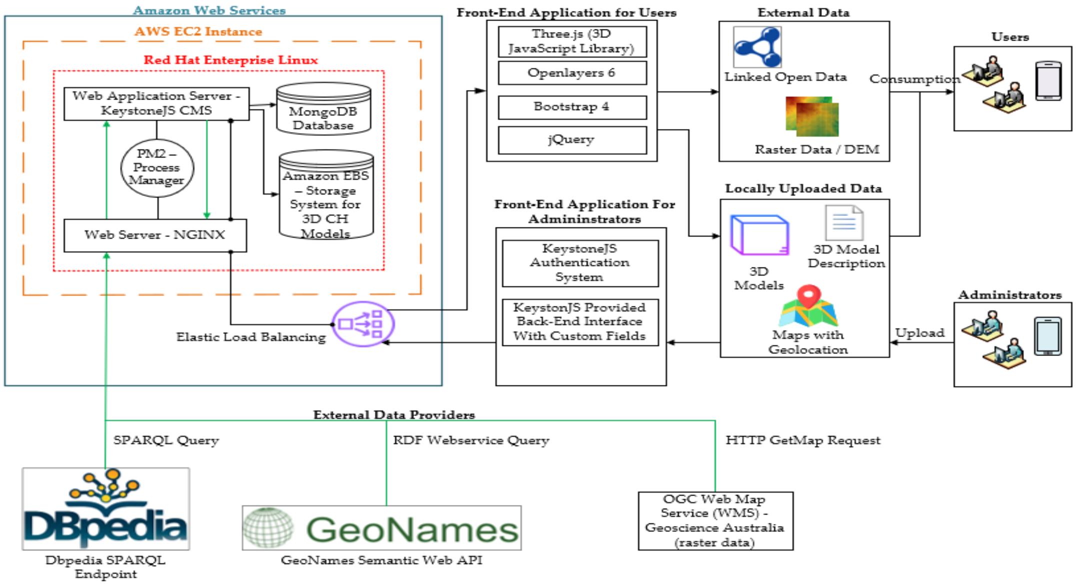


Figure 32. Proposed system architecture for integrating 3D digital models, linked open data, and raster data.

6.5 Server-side part of the architecture

As shown in the top left of Figure 32, the architecture deploys the KeystoneJS framework into the Amazon Web Services cloud platform. More specifically, the architecture utilises Amazon Elastic Compute Cloud (EC2) instance which is a web service that provides secure and resizable compute capacity on the AWS cloud. Amazon EC2 runs the Red Hat Enterprise Linux operating system, which was the preferred operating system, as it is the world’s leading enterprise Linux platform. KeystoneJS was selected as a content management system that handles the content management including uploading, editing, and publishing of 3D digital cultural heritage models and the related metadata to the architecture.

KeystoneJS is a completely FOSS database-driven content management system that makes it very easy to implement web applications and application programming interfaces (APIs). KeystoneJS is based on Node.js server-side framework and comes with the default database of MongoDB. MongoDB is one of the most popular cross-platform, document-oriented databases for web applications.

In the architecture, uploaded 3D digital cultural heritage models were stored in the AWS service—namely, Amazon EBS, while the MongoDB database stored metadata and the path to the 3D model sources in the Amazon EBS. The architecture also used a web server called NGINX which is a web server software application that acts as an intermediary server between the KeystoneJS and the clients. In my case, the clients for NGINX are front-end applications. Finally, the architecture had a PM2 process manager, and an Elastic Load Balancing service provided by AWS. The former is an advanced production process manager for Node.js-based web applications. The main aim of the PM2 process manager is to ascertain that the web application runs continuously without downtime. If the web application crashes or experiences downtime, it will restart the web application, whereas the latter aims to automatically distribute incoming application traffic across multiple Amazon EC2 instances, thus allowing the development of more fault-tolerant and highly available web architectures.

6.5.1 Front-end application for users and administrators

KeystoneJS provides an auto-generated admin user interface that handles the authentication system; the related subdiagram is shown in the "front-end application for administrators" section in Figure 32. This admin interface allows administrators to upload, edit, and publish content, including 3D digital cultural heritage models on the architecture. The front-end application for the users was implemented by utilising several front-end frameworks including a web-GIS and 3D visualisation framework; the related subdiagram is shown in the "front-end application for users" section in Figure 32.

As the front-end offers a web mapping application, an interactive web map was accomplished using the latest version of the web mapping framework of

OpenLayers 6, while the user interface buttons and other widgets were achieved using Bootstrap and jQuery frameworks. OpenLayers 6 was used for web mapping in the architecture, as it is an open-source and mature web mapping library that supports various vector and raster formats including geospatial web mapping standards of Open Geospatial Consortium (OGC) such as web mapping service (WMS), web feature service (WFS), web map tile service (WMTS), etc.

Bootstrap was used for implementing the graphical user interface of the architecture because it is the world's most popular framework for the front-end development of web applications, which is also known for its responsive and mobile-first concepts. Furthermore, Bootstrap significantly simplifies the development of front-end web applications and offers several ready-to-use templates and components such as navigation, forms, and typography.

Three.js was used to visualise 3D digital cultural heritage models, which offers numerous customisations and additions in addition to the default features. Three.js was chosen because of its non-proprietary licence, integration with all web browsers that support webGL, being relatively well documented, etc. Interested readers may refer to the previous study of the authors, which discusses three.js customisations and additions in more detail [4].

6.5.2 External data providers

As mentioned previously, DBpedia offers a SPARQL endpoint, while GeoNames offers to query their linked open data using REST API, the architecture utilises those query endpoints and retrieves the geolocation of the cultural heritage places from the query result along with some other related metadata; the related sub-diagram is shown in the external data providers section in Figure 32. For instance, from the DBpedia SPARQL endpoint, the query retrieves the World Heritage Site located in Australia, as well as cultural heritage sites that are in the Australian National Heritage List.

An example of the SPARQL query that retrieves all World Heritage Sites in Australia is shown in Figure 33. As can be seen from Figure 33, the query requests place, heritage site name, geolocation in the form of longitude and latitude, and a short description of the Australian World Heritage Sites. In the query, a variable - “?place” held a Uniform Resource Identifier (URI) of the cultural heritage site. A URI could be used to retrieve additional information about the cultural heritage site. Thus, in the architecture, users could visit the provided URI to obtain additional information and links related to the cultural heritage site.

```
SELECT DISTINCT  ?place ?heritageSiteName ?longitude ?latitude ?shortDescription
 FROM <http://dbpedia.org>
 WHERE {
        ?place <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
        <http://dbpedia.org/class/yago/WikicatWorldHeritageSitesInAustralia> ;
        <http://dbpedia.org/ontology/country> <http://dbpedia.org/resource/Australia> ;
        <http://xmlns.com/foaf/0.1/name> ?heritageSiteName ;
        <http://dbpedia.org/ontology/abstract> ?shortDescription ;
        <http://www.w3.org/2003/01/geo/wgs84_pos#long> ?longitude ;
        <http://www.w3.org/2003/01/geo/wgs84_pos#lat> ?latitude .
   FILTER ( "en" =  lang( ?shortDescription) )
}
```

Figure 33. SPARQL query that retrieves all the World Heritage Sites in Australia.

A sample of the result that was obtained from the DBpedia SPARQL and is illustrated in Figure 34, while Figure 35 shows a sample result from the GeoNames. Both sample results are returned in an RDF format. GeoNames only returns a name, URI of the cultural heritage place, and the geolocation in the form of longitude and latitude, as it is a geographical database. Furthermore, for the GeoNames REST API, the query was restricted to Australia only, which means it only returned places that are geographically located in Australia. This GeoNames search restriction guaranteed that the user would not be forwarded to other regions outside Australia.

```xml
<rdf:RDF xmlns:res="http://www.w3.org/2005/sparql-results#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<rdf:Description rdf:nodeID="rset">
<rdf:type rdf:resource="http://www.w3.org/2005/sparql-results#ResultSet" />
    <res:resultVariable>place</res:resultVariable>
    <res:resultVariable>heritageSiteName</res:resultVariable>
    <res:resultVariable>longitude</res:resultVariable>
    <res:resultVariable>latitude</res:resultVariable>
    <res:resultVariable>shortDescription</res:resultVariable>
    <res:solution rdf:nodeID="r0">
      <res:binding rdf:nodeID="r0c0"><res:variable>place</res:variable>
      <res:value rdf:resource="http://dbpedia.org/resource/Macquarie_Island"/></res:binding>
      <res:binding rdf:nodeID="r0c1"><res:variable>heritageSiteName</res:variable>
      <res:value xml:lang="en">Macquarie Island</res:value></res:binding>
      <res:binding rdf:nodeID="r0c2"><res:variable>longitude</res:variable>
      <res:value datatype="http://www.w3.org/2001/XMLSchema#float">158.85</res:value></res:binding>
      <res:binding rdf:nodeID="r0c3"><res:variable>latitude</res:variable>
      <res:value datatype="http://www.w3.org/2001/XMLSchema#float">-54.64</res:value></res:binding>
      <res:binding rdf:nodeID="r0c4"><res:variable>shortDescription</res:variable>
      <res:value xml:lang="en">Macquarie Island, a UNESCO World Heritage Site, lies in the Southwestern Pacific Ocean
    </res:solution>
```

Figure 34. An extract from the World Heritage Sites in Australia SPARQL query response.

```xml
▼<rdf:RDF xmlns:cc="http://creativecommons.org/ns#" xmlns:dcterms="http://purl.org/dc/terms/"
 xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:gn="http://www.geonames.org/ontology#"
 xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:wgs84_pos="http://www.w3.org/2003/01/geo/wgs84_pos#">
  ▼<gn:Feature rdf:about="https://sws.geonames.org/10629645/">
     <rdfs:isDefinedBy rdf:resource="https://sws.geonames.org/10629645/about.rdf"/>
     <gn:name>Ningaloo</gn:name>
     <gn:featureClass rdf:resource="https://www.geonames.org/ontology#P"/>
     <gn:featureCode rdf:resource="https://www.geonames.org/ontology#P.PPL"/>
     <gn:countryCode>AU</gn:countryCode>
     <gn:population>185</gn:population>
     <wgs84_pos:lat>-22.83612</wgs84_pos:lat>
     <wgs84_pos:long>113.80124</wgs84_pos:long>
     <gn:parentCountry rdf:resource="https://sws.geonames.org/2077456/"/>
     <gn:nearbyFeatures rdf:resource="https://sws.geonames.org/10629645/nearby.rdf"/>
     <gn:locationMap rdf:resource="https://www.geonames.org/10629645/ningaloo.html"/>
   </gn:Feature>
```

Figure 35. An extract from GeoNames semantic web query response.

6.5.3 Clustering overlapped cultural heritage places

The architecture clusters cultural heritage places into a single clickable marker according to the geolocation of cultural heritage places, as shown in Figure 36. Upon the launch of the web application, the architecture sends a request to the local REST API to retrieve the metadata of the locally stored 3D digital cultural heritage models. Simultaneously, it sends the SPARQL query to the DBpedia SPARQL endpoint to retrieve Australian World Heritage Sites and cultural heritage places that are in the Australian National Heritage List. Afterwards, the architecture starts comparing the geolocations of the locally stored 3D cultural heritage models against the geolocations of the DBpedia cultural heritage sites. Since the 3D models are stored on the local server, and the DBpedia is an external data provider, the metadata including geolocations of the locally stored 3D models are retrieved slightly faster than the DBpedia data. However, this relatively delayed response from the DBpedia is not an issue in the architecture, as it can dynamically update the clustering of cultural heritage places.

The architecture clusters cultural heritage places as follows: If cultural heritage sites from DBpedia and the locally stored 3D cultural heritage models are overlapped or within 3 metres apart from each other, they are placed under a single marker on the digital interactive map. If the cultural heritage places and the locally stored 3D cultural heritage models are located more than 3 metres apart from each other, they are placed into individual markers and shown separately. In the architecture, users can also search for any cultural heritage place located in Australia which is retrieved from the GeoNames platform and placed on top of other markers. This was purposely designed in this way to make cultural heritage places retrieved from GeoNames search result easily noticeable on the interactive map.

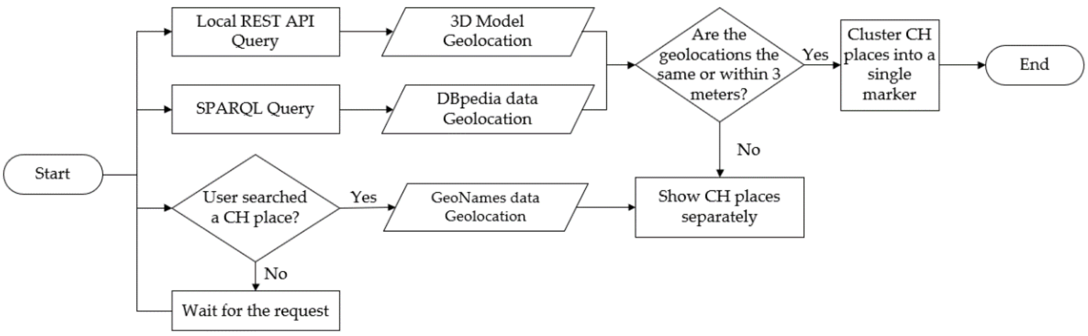


Figure 36. A flowchart for clustering overlapped cultural heritage places.

6.6 Results

The front-end interface of the architecture consists of three sections, as shown in Figure 37. The left sidebar provides map controls which allows changing the base map and switching on and off the cultural heritage sites and digital elevation model layers. A map in the central part of the user interface has clickable markers that represent cultural heritage sites from different data sources namely, DBpedia, local data with 3D models, and the GeoNames platform. A search box can be used to search for any Australian cultural heritage place which is retrieved from the GeoNames platform. The right section of the user interface displays cultural heritage data, which dynamically changes upon the selection of any marker on the map. Finally, the user interface has some other additional controls and map widgets such as an overview map that shows the current extent of the map, a scale bar that shows a drawing scale of the map, an info-box that acknowledges the data sources, and zoom buttons that makes the map content larger or smaller.

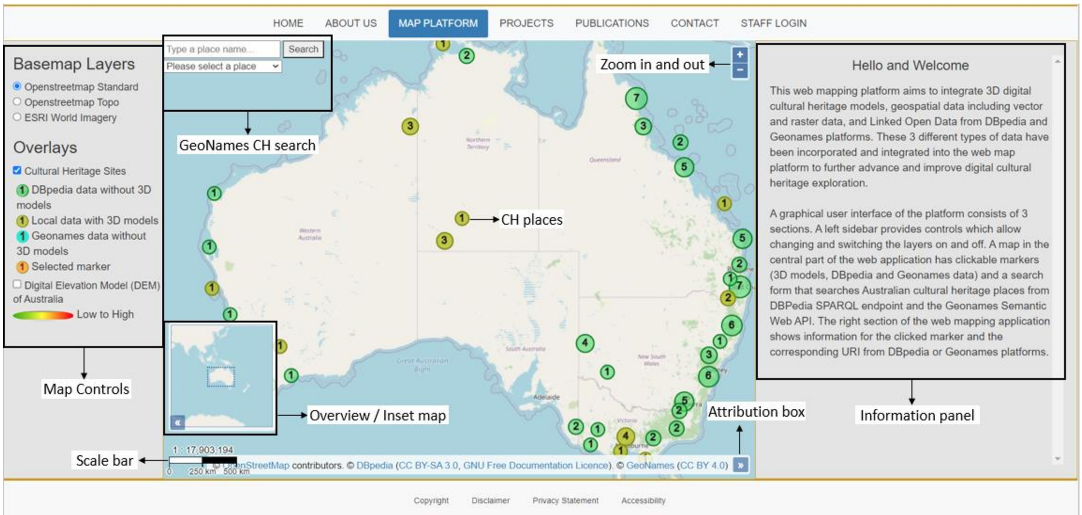


Figure 37. Front-end user interface of the architecture.

As mentioned previously, the right sidebar of the interface changes according to the selected marker on the map. For example, Figure 38 shows that Uluru–Kata Tjuta National Park has been selected on the map. Hence, the marker has changed the colour to orange, and the right sidebar displays the content relating to this particular cultural heritage place. As this marker has clusters three cultural heritage places into a single one, the right sidebar displays content for these three cultural heritage places. The first and the second cultural heritage places are from the same data source (DBpedia) and do not have a 3D cultural heritage model in them, while the third cultural heritage place was retrieved from the locally stored data and includes a 3D model of the place. The first and the second cultural heritage places are duplicates because Uluru–Kata Tjuta National Park is on both the Australian World Heritage List and National Heritage List.

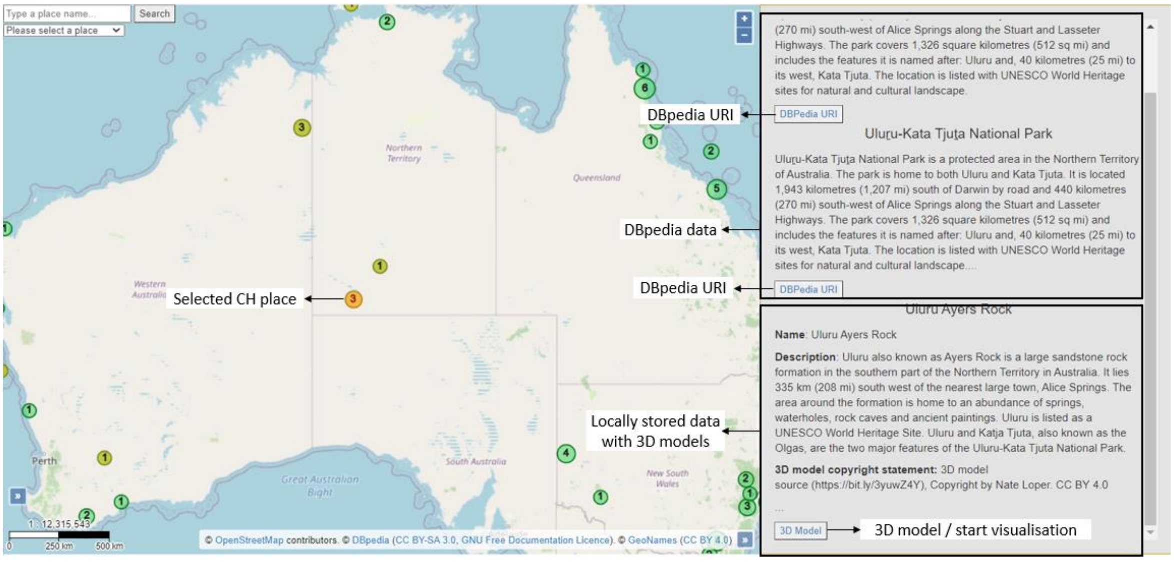


Figure 38. Selected cultural heritage place in the front-end interface.

A GeoNames search result for the same cultural heritage place (Uluru–Kata Tjuta National Park) is shown in Figure 39. It can be seen from Figure 39 that a GeoNames marker is placed on top of the other markers. Since GeoNames does not provide any additional content relating to a cultural heritage place, a right sidebar only shows a cultural heritage place name and a link to the corresponding DBpedia URI.

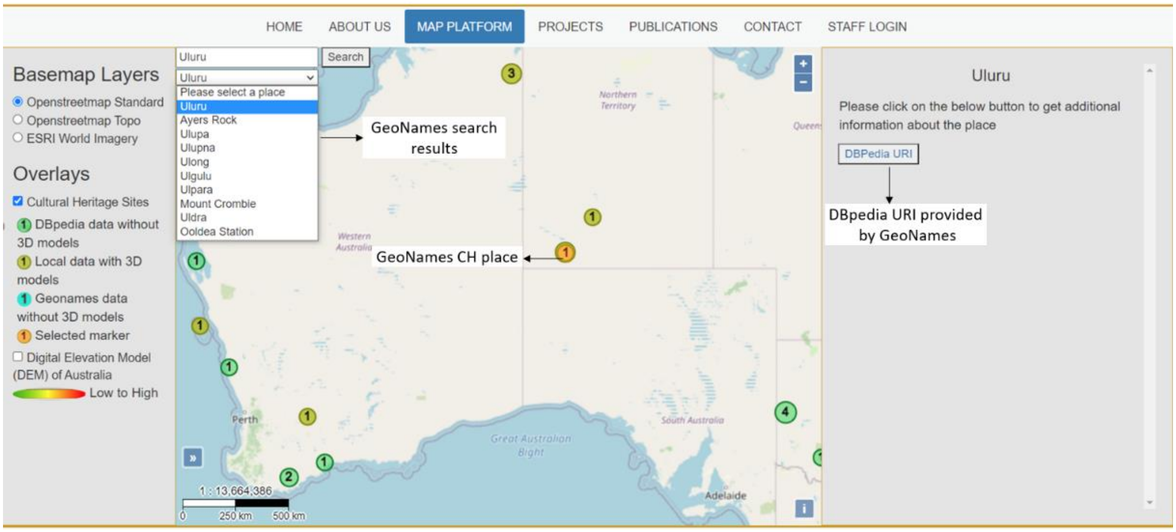


Figure 39. GeoNames result in the front-end interface.

Figure 40 shows the 3D visualisation of the cultural heritage place in the architecture. In the architecture, a 3D model can be zoomed in and out, panned around, and moved across the 3D scene. This 3D visualisation is achieved using the three.js 3D visualisation library. Three.js offers many easy-to-implement customisations, which could be implemented on top of the existing default features. Interested readers may refer to the previous study of the authors, which provides more detailed information on customisation and additions in the three.js framework [4].

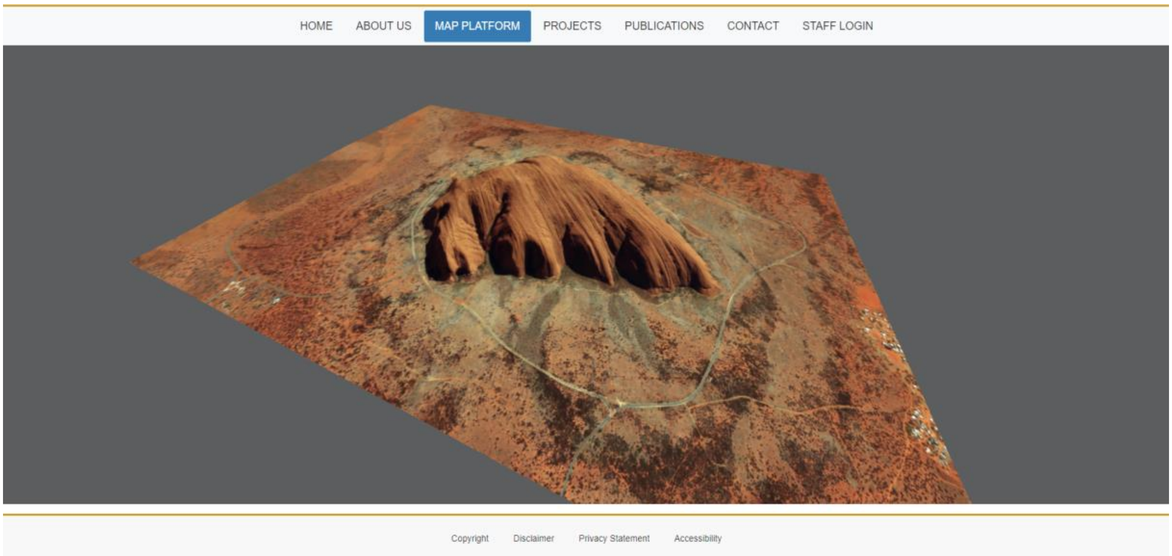


Figure 40. A 3D visualisation model in the proposed architecture.
3D model source (https://sketchfab.com/3d-models/uluru-ayers-rock-northern-territory-australia-89c0028622bc4aae82f08796a6295301, accessed on 22 August 2021), Copyright © by Nate Loper. CC BY 4.0.

As mentioned previously, the architecture provides a DBpedia URI to the cultural heritage place if a Wikipedia page is available for that particular cultural heritage place. DBpedia URI then allows users to explore web resources related to the cultural heritage place. Figure 41 shows the DBpedia URI for the Uluru–Kata Tjuta National Park. As can be seen from Figure 41, there are some external web links that users can visit to obtain some additional information about the cultural heritage place.

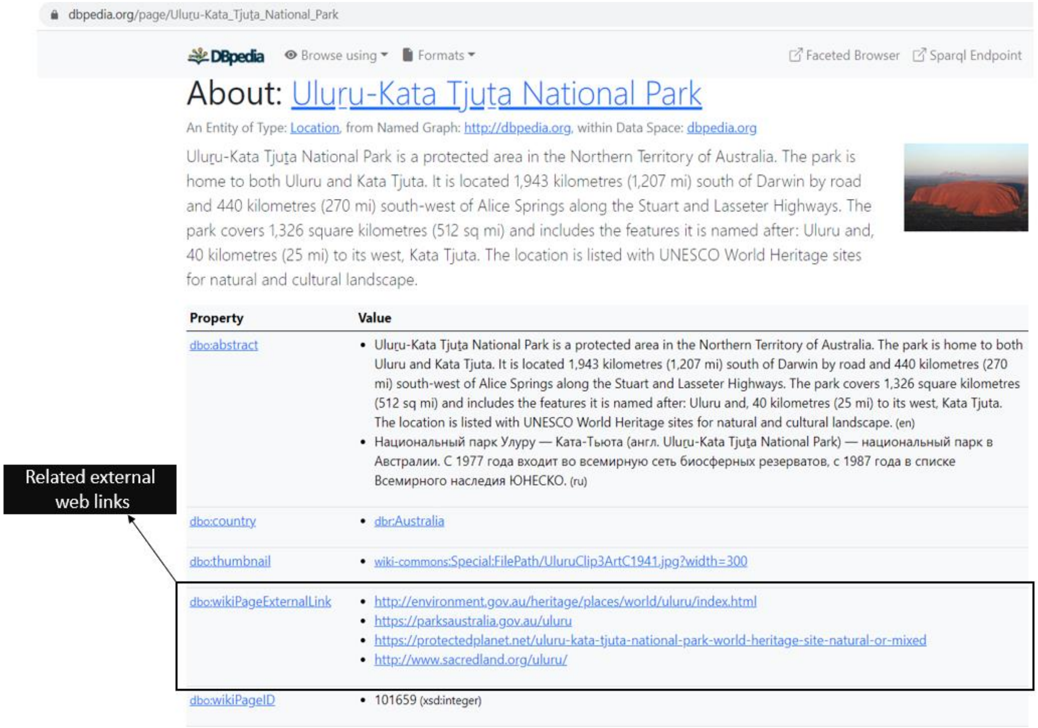


Figure 41. DBpedia data for the selected cultural heritage place. Copyright © 2021 by DBpedia Association. CC BY-SA 3.0 and GPLv2.

To better demonstrate the benefits and usefulness of the architecture, it was employed in a sample use case, which is described in the next paragraphs.

In this use case, it is assumed that a person or a user wants to explore cultural heritage places located in Australia. In such cases, this web-GIS application can easily fulfill this requirement. A user can visit the web application via any modern web browser. Upon the visit of the user, the web-GIS application visualises a digital interactive map with the cultural heritage places shown as a clickable marker. The user can then visualise 3D models available in the web-GIS application and read the provided short information about the cultural heritage places. Furthermore, if the user wants to obtain some more information about the cultural heritage place, they can visit the provided DBpedia page that presents additional information and context relating to the cultural heritage place. For instance, Figure 42 shows a 3D digital model of the previously discussed Uluru–Kata Tjuta National Park. Once the user has seen the 3D model and has read the text information of the cultural heritage place, they can visit the DBpedia page to further explore external web resources related to this cultural heritage place. As this cultural heritage place belongs to the Australian national parks category, they might be interested in opening the link in the DBpedia page that provides a list of all Australian national parks. Since a DBpedia page offers linked pages to other DBpedia pages, the user can continuously explore and gain more contextual information about the cultural heritage place.

If the digital interactive map does not have a cultural heritage place that the user is looking for, they can use a search box that finds Australian cultural heritage places from GeoNames API. The search result is then plotted on the map as a marker, and the DBpedia page is offered to the user.

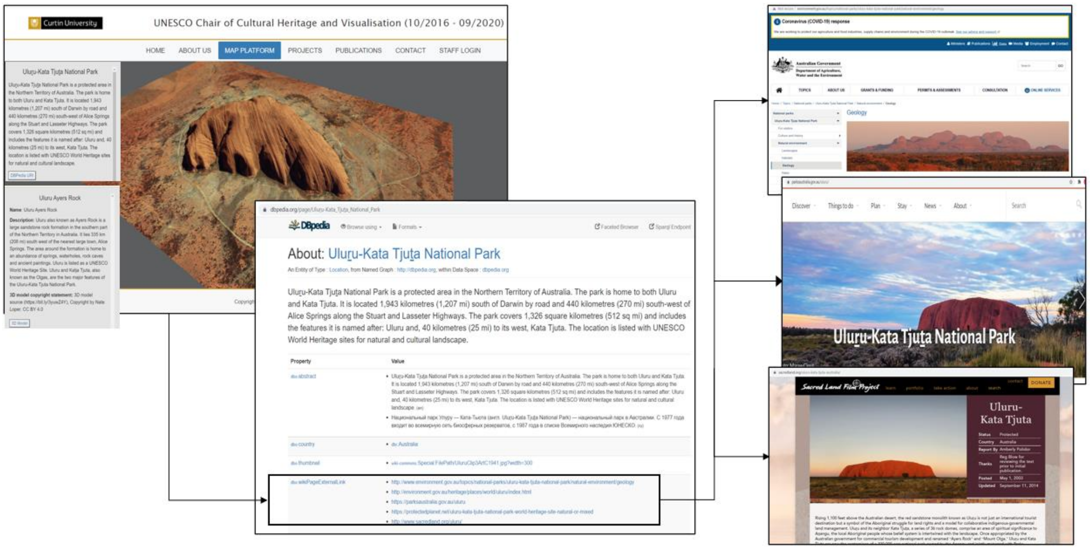


Figure 42. From 3D models to linked open data.
3D model source (https://sketchfab.com/3d-models/uluru-ayers-rock-northern-territory-australia-89c0028622bc4aae82f08796a6295301, accessed on 22 August 2021), Copyright © by Nate Loper. CC BY 4.0. Copyright © 2021 by DBpedia Association. CC BY-SA 3.0 and GPLv2.

6.7 Discussion

This study presented a novel architecture that integrates a digital interactive map and linked open data from DBpedia and GeoNames platforms with the locally stored 3D digital cultural heritage models. Although this integration works well for clustering 3D digital CH models with different CH data sources using geolocation, it might not work for cases in which a more granular clustering is needed. For instance, if the cultural heritage place occupies a large area, and the task is to link DBpedia resources to certain areas of the 3D model of the CH place, it might require linking DBpedia resources to those areas of the 3D model.

The clustering approach presented in this article can handle a broad clustering based on the geolocation; however, it does not offer a more granular linking on the 3D model level. Nevertheless, it could be feasible to use a polygon geometry on the interactive map rather than the point geometry to denote the areas of interest that should be linked to related DBpedia URI. This approach might require a text-based linking using names of the cultural heritage places rather than the geolocation, as the geolocation-based clustering might not provide sufficiently accurate matchings in this case. Once the clustering is achieved, the 3D model should only show the area to which the linking has been made by restricting the view of the 3D scene.

The architecture presented in this study utilises linked open data from DBpedia and GeoNames platforms. As mentioned previously, DBpedia is one of the greatest projects relating to linked open data, which derives the data from Wikipedia. Hence, DBpedia shares many similarities with Wikipedia, including varying content quality. In the past, there were many research articles that analysed the content quality of Wikipedia articles using various methodologies such as natural language processing and deep learning [169], assessment based on the historical edits that employ both deep neural networks and feature engineering [170], etc. [171,172].

As Wikipedia is a community-driven project, it receives content contributions from many people all over the world. Hence, the content quality of the articles in Wikipedia varies from article to article. Thus, the geolocation retrieved from the DBpedia in the architecture is not verified and might include inaccuracies including imprecise geolocation, incorrect description of places, and obsolete links to the related places. Nevertheless, many Wikipedia articles have a quality grading class assigned to them such as FA, FL, A, GA, B, C, Start, Stub, and List.

For instance, FA class is assigned to articles that have passed an in-depth quality examination, whereas GA for good quality articles and Stub for basic articles which may include significant content issues. Arguably, this Wikipedia article quality classification could be employed to improve the cultural heritage content in the platform. For instance, the architecture should only retrieve articles with at least GA—good article—grading and eliminate the ones with lower quality gradings.

On the other hand, GeoNames is a geographical database that has been deployed in many research and commercial projects. The GeoNames platform is generally assumed to contain sufficiently accurate geolocation of places. However, Ahlers [173] conducted research on the quality and accuracy of the GeoNames data and concluded that it may contain data inaccuracies including imprecise geocoordinates, overlapped places, repetitions, and misclassifications of places. Thus, a dynamic search from the GeoNames API in the architecture might also include data inaccuracies including plotting the places in the wrong location.

6.7.1 Recommendations and suggestions on the architecture by cultural heritage researchers

As mentioned previously, the authors contacted cultural heritage researchers who have experience in GIS, the semantic web, and digital cultural heritage management to collect feedback and recommendations on the architecture. In total, seven cultural heritage researchers provided constructive feedback on the architecture. The feedback and recommendation were collected through an online anonymous survey[88], as well as an online meeting with cultural heritage researchers who have expertise in GIS, linked open data, and digital cultural heritage management. The following paragraphs present these construct feedback and recommendations.

6.7.2 Interlinking words in the unstructured text document to DBpedia resources

Currently, the architecture clusters overlapped cultural heritage places into a single marker. The source of these overlapped cultural heritage places might be DBpedia or locally stored 3D digital cultural heritage models. In addition to this clustering, the authors were suggested to consider implementing entity-

88 https://curtin.au1.qualtrics.com/jfe/form/SV_9mfIkNL6e9Nm7FH (last accessed on 7 October 2021)

based interlinking of text to DBpedia resources. DBpedia has Spotlight Entity Linking which automatically interlinks unstructured text into relevant DBpedia resources.

For instance, Figure 43 shows a demo of the DBpedia Spotlight which has the unstructured text for Uluru-Kata Tjuta National Park. As can be seen from Figure 43, wherever there is a match between the words in the text and DBpedia resources, DBpedia Spotlight found and interlinked the relevant DBpedia pages to the words. This functionality of the DBpedia Spotlight could be implemented in the architecture, which could potentially interlink unstructured text to DBpedia pages upon the publication of the 3D digital cultural heritage models. Specifically, once 3D models are uploaded to the architecture and the related metadata of the 3D models entered, the architecture should run the DBpedia Spotlight to interlink the text information in the metadata with the relevant DBpedia resources. Afterwards, users of the architecture can view 3D digital cultural heritage models and visit the DBpedia pages interlinked to the words, which could potentially provide more contextualised information about the cultural heritage place.

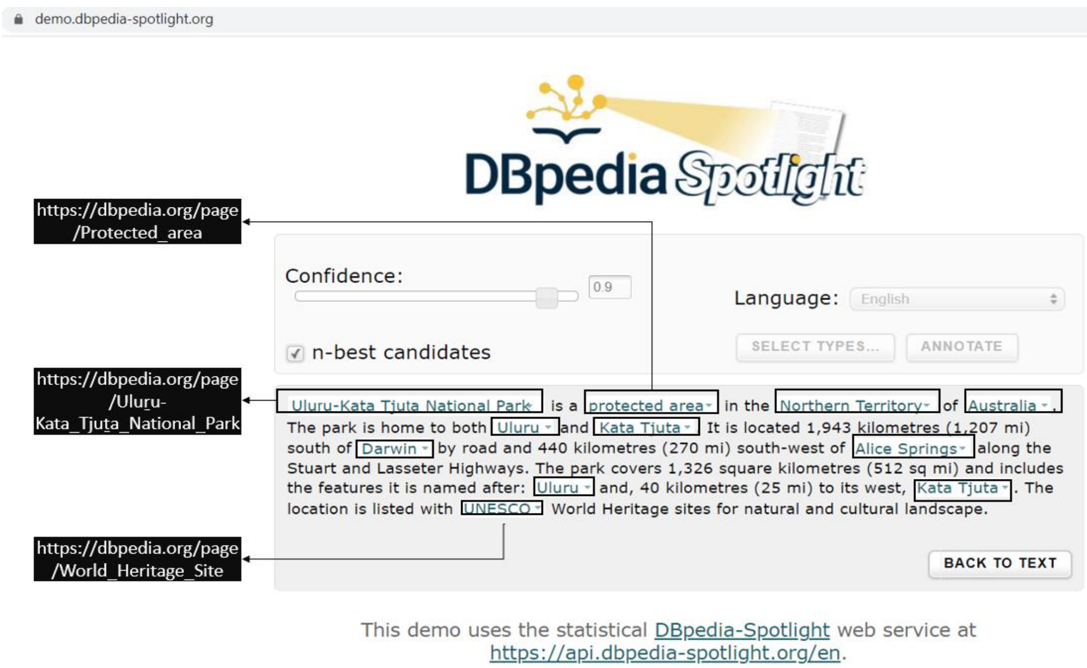


Figure 43. DBpedia Spotlight.

6.7.3 Geospatial search and filter

Currently, the architecture allows users to search for any cultural heritage place located in Australia. In addition to the text-based search available in the architecture, the authors were suggested developing geospatial search functionalities. Geospatial search based on the buffer is one type of geospatial search, which works by creating a buffer around the point. For instance, a user enters a 10 km buffer and pins a point on the map. The architecture then searches for cultural heritage places located within 10 km around the point and displays the results on the map.

Another geospatial search is finding the cultural heritage places based on the drawn polygon. This search works by asking the user to draw a polygon using a free-hand drawing tool, which enables a user to draw arbitrarily any polygon geometry on the map. Once the polygon geometry is drawn, the architecture should find the cultural heritage places that are geographically within the drawn polygon and display the results. Furthermore, the authors were also suggested including filter options to enable users to filter out the cultural heritage places. For instance, filter by certain types of cultural heritage (e.g., natural, or human-made cultural heritage), filter by year or era to which the cultural heritage places belong, etc.

6.7.4 DBpedia links in the architecture

As discussed in the previous sections, the architecture currently provides links to DBpedia pages, which are retrieved from the GeoNames search result and the DBpedia SPARQL result. In turn, these DBpedia pages include links to external web resources which could be visited to gain more contextualised information about the cultural heritage place. However, the DBpedia pages have a relatively complicated structure, which might not be user friendly to many people. Thus, people might have difficulty finding those external web resources and other related links.

For future development, the authors of this research project were advised to include more web links relating to cultural heritage sites. For instance, most DBpedia pages of cultural heritage places include information such as related external web links, the nearest city to the cultural heritage place, a subject topic of the cultural heritage place (e.g., cultural landscapes in Australia, World Heritage Sites in Australia, national parks managed by the Australia government, etc.). These additional links and information could be retrieved from the DBpedia pages and displayed in the architecture.

6.8 Conclusion

This study presented a novel architecture that aimed to enhance digital cultural heritage exploration by integrating cloud computing concepts, 3D web-based visualisation, and digital interactive map along with linked open data from the well-known RDF knowledge base of DBpedia and the geographical database of GeoNames. These distinct technologies were brought together to enable users of the architecture to interact with the digital map, visualise locally stored 3D cultural heritage models, and explore linked open data from GeoNames and DBpedia platforms, which offers additional information related to the selected cultural heritage place including external web resources.

All frameworks used in the architecture are FOSS which allows easy and cost-free replication of the architecture for CH institutions such as museums, galleries, and libraries. Furthermore, the study discussed some limitations of the architecture such as data inaccuracies that might be present in the received DBpedia and GeoNames data and the limitations of clustering of overlapped cultural heritage places presented in the article. We also presented expert recommendations and suggestions, which may also be helpful to other cultural heritage researchers and professionals working on similar research topics related to digital cultural heritage exploration and dissemination.

Author Contributions: Conceptualization, Methodology, Investigation, and Software implementation – Ikrom Nishanbaev; Survey implementation – Ikrom Nishanbaev, Erik Champion, David A. McMeekin; Writing – Review and Editing: Ikrom Nishanbaev, Erik Champion; Supervision: Erik Champion, David A. McMeekin. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The web platform source code, and the survey results data presented in the research article are available on request from the corresponding author.

Acknowledgments: We would like to thank Digital and Technology Solutions at Curtin university for providing access to Amazon Web Services cloud computing services.

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

Research Ethics Approval: Curtin University Human Research Ethics Committee (HREC) has approved this study (HREC Number HRE2021-0155).

Chapter 7 Discussion

7.1 Discussion

7.1.1 Discussion on the research objective one

As discussed in Chapter 2, an ever-increasing number of cultural heritage institutions are adopting semantic web and geospatial semantic web concepts as they have promising benefits such as better data interoperability and integration than in traditional data publishing practices. However, since the semantic web and the geospatial semantic web are relatively new concepts, at the time of publishing my research article entitled 'A Survey of Geospatial Semantic Web for Cultural Heritage' no one had performed a conceptual survey of geospatial semantic web concepts pertinent to the cultural heritage domain. However, this type of survey is crucial to a cultural heritage audience as it may equip them with an overview of all the necessary concepts, tools and frameworks, and FOSS semantic web and the geospatial semantic web platforms that can be exploited to implement geospatial semantic web-based cultural heritage platforms.

In Chapter 2, the survey provided a review on the state-of-the-art geospatial semantic web concepts pertinent to the cultural heritage field, followed by a methodology to convert geospatial cultural heritage data in a vector data model into machine-readable and processable RDF data with a case study to demonstrate its applicability. Then, leading cultural heritage projects that successfully deployed geospatial semantic web concepts were presented. Also, major FOSS semantic web and geospatial semantic web-based purpose-built platforms for cultural heritage institutions were discussed. Finally, major technical limitations of the geospatial semantic web were analysed and discussed.

According to research findings from the survey, the cultural heritage domain has already implemented some domain-specific ontologies such as CIDOC-CRM, CRMgeo to describe and document cultural heritage concepts and relationships on the semantic web. Furthermore, the cultural heritage domain has some FOSS semantic web and geospatial semantic web-based purpose-built platforms such as WissKI, Arches, ResearchSpace, and Omeka S among others, which allow cultural heritage professionals and researchers to easily publish cultural heritage data as (geo)-linked open data to the web. However, currently, these platforms and the geospatial semantic web in general have some technical limitations, constraining the use of the geospatial semantic web in cultural heritage.

Raster data representation on the geospatial semantic web is one of the major technical limitations. Raster data is a geospatial data model which is used to represent spatial phenomena. Raster data consists of a matrix of pixels where each pixel value represents information such as temperature, elevation, precipitation among others. Sources of raster data may include images from UAVs (unmanned aerial vehicles), scanned maps, satellites, and others. In cultural heritage, raster data is widely applied for analysing and monitoring cultural heritage sites. Furthermore, many web-based geospatial data publishing platforms such as GeoServer (http://geoserver.org/) support various formats of raster data such as GeoTIFF, WorldImage, GeoPackage, and others. Nevertheless, raster data representation on the geospatial semantic web is still challenging as there is no consistent way of encoding raster data for the geospatial semantic web. Recently, there have been some research works to overcome this challenge. For instance, one proposed methodology was to build vector objects using image segmentation and create RDF data for the vector objects [76]. However, more research projects are needed to consolidate the raster data representation challenge of the geospatial semantic web.

Another technical challenge is 3D geospatial semantic web, in particular encoding massive 3D models in RDF, and querying 3D geospatial semantic web. Currently, there are some standardized ways of representing large 3D models such as virtual 3D city models using a CityGML (https://www.ogc.org/standards/citygml, last accessed on 1 November 2021) data model by OGC (Open Geospatial Consortium), 3D Tiles (https://github.com/CesiumGS/3d-tiles, last accessed on 1 November 2021) by the developers of the Cesium and others. Although there have recently been some promising research works that propose encoding and querying large 3D models on the geospatial semantic web [174,175], there are still many challenges associated with those approaches such as slow performance of the resulting RDF data, transformation issues that may occur during the conversion from the CityGML data model to the proposed OWL ontology among others [174,175]. Hence, more research projects are needed to define consistent ways to encode, interlink, query, process, and visualise massive 3D data on the geospatial semantic web.

There are also some other challenges such as handling big geospatial data on the geospatial semantic web, a plethora of data models to choose from for encoding geospatial data, issues with entity matching also known as instance matching in the geospatial semantic web among others.

7.1.2 Discussion on the research objective two

In Chapter 3, this thesis compared and evaluated the features and functionality of the current state-of-the-art geospatial RDF generation tools and interlinking frameworks. This evaluation was specifically performed for people, including cultural heritage researchers and professionals, with limited expertise in computer programming. Hence, the tools and frameworks were chosen based on a pre-defined set of criteria. Moreover, Chapter 3 presented a complete methodology to generate RDF data using CIDOC-CRM ontology, and to interlink the resulting RDF data with a leading knowledge base of DBpedia. However, it is worth noting that this comparative evaluation did not attempt to perform performance benchmarking of RDF generation tools and frameworks.

This evaluation was performed due to two reasons: a) according to the research findings of Schmachtenberg, Bizer and Paulheim [16], 44% of published linked open datasets on the web are not linked to other datasets at all; b) many RDF generation tools and interlinking frameworks require solid knowledge in programming and the semantic web to be deployed successfully.

However, my research findings show that it is still achievable to generate RDF data using CIDOC-CRM ontology and interlink the resulting RDF data to knowledge bases such as DBpedia without programming or scripting, as the compared tools and frameworks offer a graphical user interface (GUI) to the users. As a result, this may encourage cultural heritage researchers and professionals to publish cultural heritage data as linked open data and most importantly interlink the resulting RDF data with other related data sources on the semantic web.

Although the compared RDF generation tools provide GUI and do not require computer programming, they still require knowledge of the chosen ontology to be able to map the concepts and relationships in the data to the ontology. Regarding the interlinking frameworks, they also do not require programming. However, in most cases, the process of interlinking is not automatic and requires someone to specify the path or class and property for the datasets being interlinked. This means they require users to have a solid knowledge of the ontology of the datasets to specify those attributes. This may cause significant challenges to cultural heritage researchers and professionals if the task is to interlink the resulting RDF data with large knowledge bases such as DBpedia where there are a great many classes and properties.

7.1.3 Discussion on the research objective three

In Chapter 4, the thesis presented a reusable, extendable, novel web repository including methodology based on FOSS, easy-to-implement frameworks for long-term archiving, and dissemination of 3D cultural heritage models. This initiative was inspired by the fact that it is difficult to find, use and re-use 3D cultural heritage models, especially the ones that resulted from research projects. Furthermore, many previous related projects developed a web repository from the ground up, which might cause challenges for replication.

At the core of the web repository and methodology is a FOSS, easy-to-implement, database-driven, content management system (CMS) namely KeystoneJS. According to the developers of the KeystoneJS, “Keystone helps you build faster and scale further than any other CMS or App Framework.", (https://keystonejs.com/, last accessed on 1 November 2021).

All other frameworks, such as web mapping, 3D web-based visualisation, etc, in the web repository and methodology, are FOSS, which allows great flexibility for extensions and cost-free implementation. These two characteristics are imperative for many cultural heritage researchers and professionals as research and applied projects may have various requirements on the platform including a limited budget.

As discussed in Chapter 4, although this methodology and web repository provide a solution for long-term archiving and dissemination of 3D cultural heritage models, it does not deal with part-based annotation of 3D cultural heritage models to knowledge bases. For instance, a 3D digital model of an ancient building might consist of parts such as a roof, walls, windows, and doors, etc. These specific parts of the building could be interlinked to knowledge bases such as DBpedia or GeoNames using annotations drawn on top of the 3D model. This functionality may allow users to get information about each annotated part of the building individually.

7.1.4 Discussion on the research objective four

Chapter 5 extended the previous web repository and methodology to integrate cloud computing concepts into the repository. As discussed in Chapter 5, cloud computing offers many benefits over traditional IT infrastructures such as almost unlimited computing and storage resources, backup solutions for data, low maintenance costs, flexibility on hardware and software, and others. Thus, this extended web repository and methodology presented an AWS-based cloud architecture that integrates cloud computing concepts, digital interactive map, and 3D digital cultural heritage models. This integration is based on AWS cloud services; however, the presented methodology could also be deployed into any other IT infrastructure which supports Node.js environment as KeystoneJS is based on the Node.js runtime environment.

7.1.5 Discussion on the research objective five

As discussed in Chapter 6, in the past there were plenty of research studies that prove the great benefits of web-GIS for the dissemination and online representation of cultural heritage data. Nevertheless, there are not many research projects that present a web-GIS architecture that integrates cloud computing concepts, linked open data from DBpedia and GeoNames platforms, and 3D cultural heritage models to enhance cultural heritage exploration. Thus, Chapter 6 further extended the previous web platform and methodology to integrate digital interactive map, 3D digital cultural heritage models, and linked open data from DBpedia and GeoNames platforms into a novel cloud-based web-GIS architecture. Also, seven cultural heritage researchers, who have expertise in GIS, linked open data and digital cultural heritage, provided expert feedback and recommendations on the web-GIS architecture. These feedback and recommendations were collected using an online anonymous survey and direct online communication with the researchers.

According to my research findings and survey analysis, this integration can enhance cultural heritage exploration as the users of the web-GIS architecture can visualize 3D digital cultural heritage models and explore linked open data from DBpedia and GeoNames platforms, which can provide additional web resources including links about the cultural heritage places. According to the feedback and recommendation results, additional features in the web-GIS architecture that could further enhance the digital cultural heritage exploration are: to interlink words in the unstructured text document to DBpedia resources; to develop additional geospatial search and filter functionalities; provide more DBpedia links related to the cultural heritage place.

Chapter 8 Conclusion and Future Work

8.1 Conclusion

This PhD project had the following aims:

- To equip cultural heritage researchers and professionals with an overview of the geospatial semantic web concepts pertinent to cultural heritage including tools, frameworks, and FOSS purpose-built semantic web and the geospatial semantic web platforms that can be exploited to implement geospatial semantic web cultural heritage repositories.

This aim was achieved in Chapter 2. In Chapter 2, the thesis presented a survey on the state-of-the-art geospatial semantic web concepts pertinent to the cultural heritage field such as domain-specific ontologies, leading cultural heritage projects that successfully employed geospatial semantic web concepts, and FOSS purpose-built semantic web and the geospatial semantic web platforms that can be exploited to implement geospatial semantic web cultural heritage repositories. It also presented a methodology for converting geospatial cultural heritage data in a vector data model into machine-readable and processable RDF data. Chapter 2 concluded with a discussion on the major technical limitations of the geospatial semantic web that are limiting the use of the geospatial semantic web in many fields including in cultural heritage.

- To develop guidelines and methodology, which do not involve programming or scripting, for generating geospatial linked open data including RDF data generation using CIDOC-CRM ontology and interlinking the resulting RDF data to knowledge bases such as DBpedia.

This aim was achieved in Chapter 3. In Chapter 3, the thesis compared and evaluated the features and functionality of the current state-of-the-art geospatial RDF generation tools and interlinking frameworks. As this evaluation was performed for people with limited expertise in computer programming, the tools and frameworks were chosen based on a pre-defined set of criteria. Chapter 3 also presented a complete methodology to generate

RDF data using CIDOC-CRM ontology, and to interlink the resulting RDF data with the leading knowledge base of DBpedia. Although the compared RDF generation tools and interlinking frameworks do not require computer programming, they still require some knowledge of semantic web concepts.

For instance, RDF generation tools require knowledge of the chosen ontology as mapping the concepts and relationships in the data to the chosen ontology needs to be done by the user. The interlinking frameworks also require the user to specify the path or class and property for the datasets being interlinked. This may cause significant challenges to cultural heritage researchers and professionals if the task is to interlink the resulting RDF data with large knowledge bases such as DBpedia with a great many classes and properties.

- To develop a reusable, completely extendable, cloud-based, novel web repository as well as methodology based on FOSS, easy-to-implement frameworks for long-term archiving and dissemination of 3D cultural heritage models.

This aim was achieved in Chapter 4 and Chapter 5.

In Chapter 4, the thesis presented a reusable, extendable, novel web repository as well as methodology based on FOSS, easy-to-implement frameworks for long-term archiving and dissemination of 3D cultural heritage models. More specifically, the interactive map, user interface widgets, and 3D visualisation were achieved using OpenLayers, Bootstrap, and Three.js respectively, while a content management system (CMS) named KeystoneJS were used to manage (e.g., publish, edit, update, delete) all the content in the web repository. KeystoneJS comes with MongoDB database by default. According to the developers of KeystoneJS, “Keystone helps you build faster and scale further than any other CMS or App Framework”. Thus, the methodology presented in Chapter 4 can help cultural heritage organizations and cultural heritage professionals facilitate the rapid implementation of the web repository for geo-located 3D digital cultural heritage models at no cost. It is worth mentioning that cultural heritage projects may have some additional technical requirements on the web repository. For instance, 3D cultural heritage models can have time-related information (temporal information) and even different 3D digital representations for different time periods as the physical representation of cultural heritage objects, sites, etc may change over time. Since this web repository does not use any commercial frameworks or tools, it can easily be extended to accommodate project requirements including the above-mentioned temporal information.

In Chapter 5, the thesis presented an extended methodology and web repository to integrate cloud computing concepts into the repository from Chapter 4. Cloud computing offers many benefits such as almost unlimited computing and storage resources, backup solutions for data, low maintenance costs, and flexibility on hardware and software among others. Thus, Chapter 5 presents a cloud architecture based on AWS which includes cloud services such as Amazon EBS storage for 3D cultural heritage models, Elastic Load Balancing, PM2 process manager, the web server of NGINX, and cloud installed MongoDB database among others. AWS has been selected as a cloud computing solutions provider as it is the largest provider in the cloud computing market. It is important to note that although this cloud architecture is based on AWS cloud services, the presented methodology can also be deployed into any other IT infrastructure which supports Node.js environment as KeystoneJS is based on Node.js runtime environment.

- To enhance digital cultural heritage exploration by integrating cloud computing concepts, digital interactive maps, linked open data from DBpedia and GeoNames, and 3D digital cultural heritage models into a web-GIS-based cloud architecture.

This aim was achieved in Chapter 6, which further extended the previous web platform and methodology to integrate digital interactive map, 3D digital cultural heritage models, and linked open data from DBpedia and GeoNames platforms into a novel cloud-based web-GIS architecture. As demonstrated in Chapter 6, this integration can enhance cultural heritage exploration as the users of the web-GIS architecture can visualise 3D digital cultural heritage models and explore linked open data from DBpedia and GeoNames platforms.

8.2 Future work

Interlinking parts of 3D cultural heritage models to knowledge bases

Chapter 6 presented a novel architecture that integrates a digital interactive map and linked open data from DBpedia and GeoNames platforms with the locally stored 3D digital cultural heritage models. As mentioned in Chapter 6, future work on this part of the thesis may include integrating a 3D geospatial interactive map with a part-based annotation tool that allows interlinking parts of the 3D model with knowledge bases such as DBpedia and GeoNames. This would facilitate more granular interlinking of 3D digital cultural heritage models to knowledge bases by allowing the users to draw so-called regions of interest (RoS) on the 3D model. Thus, it may potentially open up new ways to integrate and interact with cultural heritage information.

As I mentioned in Chapter 6, there have already been some promising research projects in cultural heritage and other fields that proposed a similar tool [78,176,177]; however, there are still many technical issues associated with these proposed tools. There is no doubt that there is a reliable method for implementing a tool (e.g., ray casting) for drawing the above mentioned RoS on 3D models. Nevertheless, the challenge might occur with storing the part-based annotations and sharing them across platforms. In the past there were some research works on standardization of storing and sharing 3D semantic annotations through Open Annotation Data model [78,176]. Initial results of these research works are promising but they have not tested sharing the 3D semantic annotations across platforms. Thus, future research work might include implementing a framework that allows sharing 3D semantic annotations across platforms.

Interlinking words in the unstructured text document to DBpedia resources, and geospatial search and filter

As mentioned in the survey results in Chapter 6, this web-GIS architecture could further be extended by implementing interlinking words in the unstructured text document to DBpedia resources. This would allow then interlinking unstructured text to DBpedia pages upon the publication of 3D digital cultural heritage models in the architecture, which could in turn further enhance cultural heritage exploration.

Some other future developments may include geospatial search and filter functionalities such as buffer search; providing more DBpedia links related to the cultural heritage places such as the nearest city to the cultural heritage place; a subject topic of the cultural heritage place and others.

RDF generations tools and interlinking frameworks for non-programmers

In Chapter 3 I presented a comparative evaluation on features and functionality of the current state-of-the-art geospatial RDF generation tools and frameworks, which was specifically performed for people with limited expertise in computer programming including cultural heritage. As I discussed in Chapter 3, there have been many relatively large-scale research projects that developed RDF generation tools and interlinking frameworks, most of which support a graphical user interface and do not require computer programming. However, they still require a solid knowledge of ontology to which the RDF data is being interlinked. Whereas interlinking frameworks require to specify the path or class and property for the datasets being interlinked, which means they can't interlink RDF datasets without human intervention. This could be challenging if the task is to interlink RDF data with a large knowledge base such as DBpedia. Thus, future research work could be on further easing and simplifying the RDF generation and interlinking process. This could allow cultural heritage researchers and professionals to publish linked open data more easily and enrich linked open data cloud among others.

3D geospatial semantic web

Recent developments in 3D modelling and visualisation have now enabled cultural heritage professionals and researchers to easily create a 3D digital replica of cultural heritage objects and display them in web browsers without any plugins. On the other hand, geospatial semantic web technologies are being advanced with the aim of making Internet geospatial data machine-readable and processable, thus enabling machines to understand the meaning behind the data. However, the 3D geospatial semantic web remains an active research area, particularly representation, encoding, and querying 3D digital models in the geospatial semantic web. Recent research projects have now presented very promising results such as semantic 3D city database [178] and semantic web 3D city agents [179] among others. Hence, future research work may include on the integration of 3D geospatial semantic web technologies with cultural heritage including encoding, storing, and representation of country-scale 3D cultural heritage models in geospatial RDF data.

Raster data representation in the geospatial semantic web

As mentioned in Chapter 1, raster data is a data model for representing spatial phenomena, which can be used for representing various types of geospatial data such as continuous data (e.g. elevation, precipitation, temperature), thematic data with classifications (e.g. land use with different classifications such as vegetation, soil) [71]. The sources of raster data include digital images from satellites, UAV (unmanned aerial vehicles), scanned maps, and many others. In the cultural heritage domain, raster data can be employed in many case studies, such as analysing and monitoring CH sites [72,73]. However, raster data representation is currently not supported in the geospatial semantic web, more specifically RDF data models GeoSPARQL, stRDF, and others do not support raster data yet. Hence, there is no consistent way of encoding raster data for the geospatial semantic web. The main reason for that is raster data consists of a matrix of cells each containing a value or values that represent information (e.g., spectral value, category, magnitude, and height). Therefore, converting raster files into RDF would increase the size of the triples substantially or it might even require a few triples to encode a raster file, which in the end could create performance issues [74,75]. Nevertheless, there have already been several research projects attempting to solve this issue. For example, there is a methodology that builds vector objects using image segmentation and then creates RDF for vector objects [76]. This workaround is arguably not a complete solution to represent raster files in RDF as it still does not represent raster geometry in RDF. Thus, future work might include on encoding and representation of raster data in the geospatial semantic web for cultural heritage.

Appendix

Human ethics approval

Curtin University

Research Office at Curtin

GPO Box U1987
Perth Western Australia 6845

Telephone +61 8 9266 7863
Facsimile +61 8 9266 3793
Web research.curtin.edu.au

30-Mar-2021

Name: David McMeekin
Department/School: Curtin University
Email: David.Mcmeekin@curtin.edu.au

Dear David McMeekin

RE: Ethics Office approval
Approval number: HRE2021-0155

Thank you for submitting your application to the Human Research Ethics Office for the project **Web GIS-based integration of 3D cultural heritage models with Linked Open Data.**

Your application was reviewed through the Curtin University Low risk review process.

The review outcome is: **Approved.**

Your proposal meets the requirements described in the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research (2007).*

Approval is granted for a period of one year from **30-Mar-2021** to **29-Mar-2022.** Continuation of approval will be granted on an annual basis following submission of an annual report.

Personnel authorised to work on this project:

Name	Role
Nishanbaev, Ikrom	Co-Inv
Champion, Erik	Co-Inv
McMeekin, David	CI

Approved documents:

Document

Standard conditions of approval

1. Research must be conducted according to the approved proposal
2. Report in a timely manner anything that might warrant review of ethical approval of the project including:
   - proposed changes to the approved proposal or conduct of the study
   - unanticipated problems that might affect continued ethical acceptability of the project
   - major deviations from the approved proposal and/or regulatory guidelines
   - serious adverse events
3. Amendments to the proposal must be approved by the Human Research Ethics Office before they are implemented (except where an amendment is undertaken to eliminate an immediate risk to participants)

4. An annual progress report must be submitted to the Human Research Ethics Office on or before the anniversary of approval and a completion report submitted on completion of the project
5. Personnel working on this project must be adequately qualified by education, training and experience for their role, or supervised
6. Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, that bears on this project
7. Changes to personnel working on this project must be reported to the Human Research Ethics Office
8. Data and primary materials must be retained and stored in accordance with the Western Australian University Sector Disposal Authority (WAUSDA) and the Curtin University Research Data and Primary Materials policy
9. Where practicable, results of the research should be made available to the research participants in a timely and clear manner
10. Unless prohibited by contractual obligations, results of the research should be disseminated in a manner that will allow public scrutiny; the Human Research Ethics Office must be informed of any constraints on publication
11. Approval is dependent upon ongoing compliance of the research with the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research, applicable legal requirements, and with Curtin University policies, procedures and governance requirements
12. The Human Research Ethics Office may conduct audits on a portion of approved projects.

Special Conditions of Approval

It is the responsibility of the Chief Investigator to ensure that any activity undertaken under this project adheres to the latest available advice from the Government or the University regarding COVID-19.

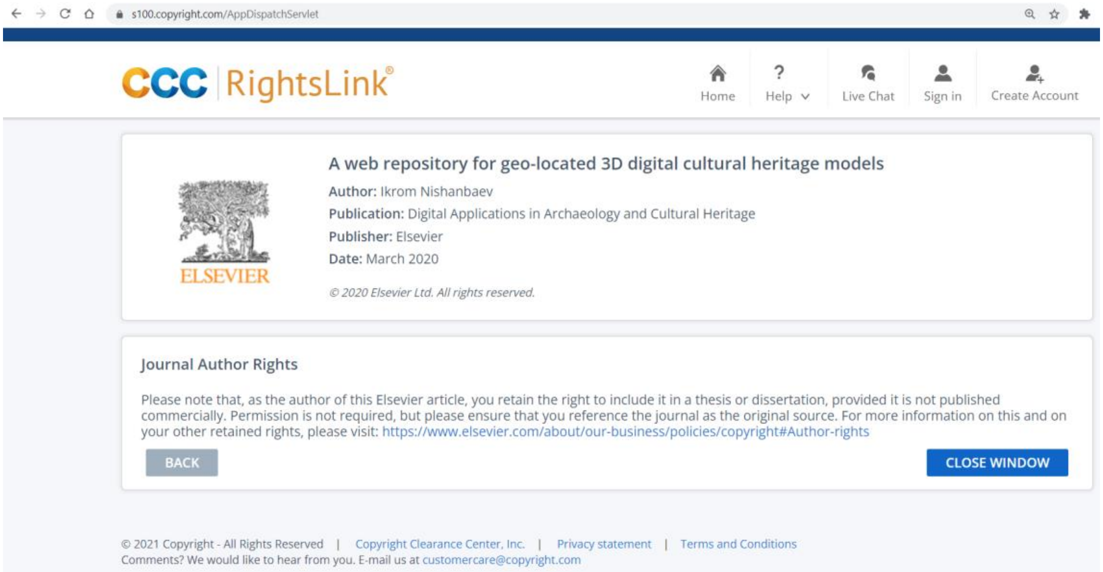
This letter constitutes low risk/negligible risk approval only. This project may not proceed until you have met all of the Curtin University research governance requirements.

Should you have any queries regarding consideration of your project, please contact the Ethics Support Officer for your faculty or the Ethics Office at hrec@curtin.edu.au or on 9266 2784.

Yours sincerely

Amy Bowater
Ethics, Team Lead

Copyright information to use a research article – "A web repository for geo-located 3D digital cultural heritage models" in the thesis.


s100.copyright.com/AppDispatchServlet

CCC RightsLink®

Home | Help | Live Chat | Sign in | Create Account

A web repository for geo-located 3D digital cultural heritage models

Author: Ikrom Nishanbaev

Publication: Digital Applications in Archaeology and Cultural Heritage

Publisher: Elsevier

Date: March 2020

© 2020 Elsevier Ltd. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK | CLOSE WINDOW

© 2021 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions

Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Best Student Paper Award

INSTICC

Institute for Systems and Technologies of Information, Control and Communication

Best Student Paper Award Certificate

for the paper entitled:

A Cloud Architecture for Processing and Visualization of Geo-located 3D Digital Cultural Heritage Models

authored by:

Ikrom Nishanbaev

received at the

6th International Conference on Geographical Information Systems Theory, Applications and Management
(GISTAM)

held from May 7 - 9, 2020

On behalf of the Organizing Committee,

Lemonia Ragia
GISTAM Conference Chair

e-mail: secretariat@insticc.org

website: http://www.insticc.org

References

1. Comes, R.; Buna, Z.; Badiu, I. Creation and preservation of digital cultural heritage. *Journal of Ancient History and Archaeology* 2014, *1*, 50-56.

2. Santos, P.; Ritz, M.; Fuhrmann, C.; Fellner, D. 3D mass digitization: a milestone for archeological documentation. *Virtual Archaeology Review* 2017, *8*, 1-11.

3. Barsanti, S.G.; Guidi, G. 3D digitization of museum content within the 3D-ICONS project. In Proceedings of the XXIV International CIPA Symposium, Strasbourg, France, 2-6 September, 2013; pp. 151-156.

4. Nishanbaev, I. A web repository for geo-located 3D digital cultural heritage models. *Digital Applications in Archaeology and Cultural Heritage* 2020, *16*, e00139.

5. Cofrancesco, P.; Capolupo, C.; Frisetti, A. Mapping Medieval Christianity. CARE-GIS: Reusing Web Digital Data in a webGIS Application. *Journal on Computing and Cultural Heritage (JOCCH)* 2019, *12*, 1-12.

6. Agapiou, A.; Lysandrou, V.; Alexakis, D.D.; Themistocleous, K.; Cuca, B.; Argyriou, A.; Sarris, A.; Hadjimitsis, D.G. Cultural heritage management and monitoring using remote sensing data and GIS: The case study of Paphos area, Cyprus. *Computers, Environment and Urban Systems* 2015, *54*, 230-239.

7. 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). *International journal of architectural heritage* 2016, *10*, 9-19.

8. Dore, C.; Murphy, M. Integration of Historic Building Information Modeling (HBIM) and 3D GIS for recording and managing cultural heritage sites. In Proceedings of the 18th International Conference on Virtual Systems and Multimedia, Milan, Italy, 2012; pp. 369-376.

9. Brovelli, M.A.; Magni, D. An archaeological web GIS application based on Mapserver and PostGIS. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences* 2003, *34*, 89-94.

10. Duran, Z.; Doğru, A.G.; Toz, G. Web-based multimedia GIS for historical sites. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2004, *35*.

11. Dąbrowski, J.; Kulawiak, M.; Moszyński, M.; Bruniecki, K.; Kamiński, Ł.; Chybicki, A.; Stepnowski, A. Real-time web-based GIS for analysis,

visualization and integration of marine environment data. In *Information Fusion and Geographic Information Systems*, Vasily Popovich, M.S., Jean-Claude Thill, Christophe Claramunt, Tianzhen Wang, Ed.; Springer: 2009; pp. 277-288.

12. Vacca, G.; Fiorino, D.R.; Pili, D. A spatial information system (SIS) for the architectural and cultural heritage of sardinia (Italy). *ISPRS International Journal of Geo-Information* 2018, *7*, 49.

13. Costantino, D.; Angelini, M.; Alfio, V.; Claveri, M.; Settembrini, F. Implementation of a system WebGIS open-source for the protection and sustainable management of rural heritage. *Applied Geomatics* 2020, *12*, 41-54.

14. Bizer, C.; Heath, T.; Berners-Lee, T. Linked data-the story so far. *International journal on semantic web and information systems* 2009, *5*, 1-22.

15. Guerrini, M.; Possemato, T. Linked data: a new alphabet for the semantic web. *Italian Journal of Library, Archives, and Information Science* 2013, *4*, 67-90.

16. 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, October 19-23, 2014; pp. 245-260.

17. Champion, E.; Rahaman, H. 3D Digital Heritage Models as Sustainable Scholarly Resources. *Sustainability* 2019, *11*, 2425.

18. Varghese, B.; Buyya, R. Next generation cloud computing: New trends and research directions. *Future Generation Computer Systems* 2018, *79*, 849-861.

19. Remondino, F. Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Remote Sensing* 2011, *3*, 1104-1138.

20. Xiao, W.; Mills, J.; Guidi, G.; Rodríguez-Gonzálvez, P.; Barsanti, S.G.; González-Aguilera, D. Geoinformatics for the conservation and promotion of cultural heritage in support of the UN Sustainable Development Goals. *ISPRS Journal of Photogrammetry and Remote Sensing* 2018, *142*, 389-406.

21. Petrescu, F. The use of GIS technology in cultural heritage. In Proceedings of the XXI International CIPA Symposium, Athens, Greece, 2007; pp. 01-06.

22. Solter, A.; Gajski, D. Towards the virtual museum–exploring tools and methods for 3D Digitalization and visualization. *Opvscvla archaeologica* 2018, *39*, 117-124.

23. Preuss, U. Sustainable Digitalization of Cultural Heritage - Report on Initiatives and Projects in Brandenburg, Germany. *Sustainability* 2016, *8*, 891.

24. Dhonju, H.; Xiao, W.; Mills, J.; Sarhosis, V. Share Our Cultural Heritage (SOCH): Worldwide 3D Heritage Reconstruction and Visualization via Web and Mobile GIS. *ISPRS International Journal of Geo-Information* 2018, *7*, 360.

25. Kondo, Y.; Miki, T.; Kuronuma, T.; Oguchi, T. On-site digital heritage inventory development at Bat, Oman. In Proceedings of the 25th International CIPA Symposium, Taipei, Taiwan, 2015; pp. 145-152.

26. Scott, R.E. Europeana. *Music Reference Services Quarterly* 2013, *16*, 118-120, doi:10.1080/10588167.2013.782221.

27. Valtysson, B. EUROPEANA. *Information, Communication & Society* 2012, *15*, 151-170, doi:10.1080/1369118X.2011.586433.

28. Fresa, A. A data infrastructure for digital cultural heritage: characteristics, requirements and priority services. *International Journal of Humanities and Arts Computing* 2013, *7*, 29-46.

29. Zhang, C.; Zhao, T.; Li, W. Conceptual Frameworks of Geospatial Semantic Web. In *Geospatial Semantic Web*; Springer: Cham, Switzerland, 2015; pp. 35-56.

30. 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, November 8-9, 2002, 2002; pp. 1-4.

31. Hyvönen, E. Cultural heritage linked data on the semantic web: Three case studies using the sampo model. In Proceedings of the VIII Encounter of documentation centres of contemporary art: open linked data and integral management of information in cultural centres., Vitoria-Gasteiz, Spain, 19-20 October, 2016; pp. 19-20.

32. Hyvönen, E.; Mäkelä, E.; Kauppinen, T.; Alm, O.; Kurki, J.; Ruotsalo, T.; Seppälä, K.; Takala, J.; Puputti, K.; Kuittinen, H. CultureSampo: A national publication system of cultural heritage on the semantic Web 2.0. In Proceedings of the European Semantic Web Conference, Crete, Greece, 31 May - 4 June, 2009; pp. 851-856.

33. 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, Indiana, USA, 23-26 June, 2014; pp. 197-201.

34. Li, W.; Zhao, T.; Zhang, C. Big Geospatial Data and the Geospatial Semantic Web: Current State and Future Opportunities. In *Big Data*

and Computational Intelligence in Networking; CRC Press: Boca Raton, Florida, USA, 2017; pp. 59-80.

35. Kuhn, W.; Kauppinen, T.; Janowicz, K. Linked data -a paradigm shift for geographic information science. In Proceedings of the International Conference on Geographic Information Science, Vienna, Austria, 24-26 September, 2014; pp. 173-186.

36. Nalepa, G.J.; Furmańska, W. Review of semantic web technologies for GIS. *Automatyka/Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie* 2009, *13*, 485-492.

37. Faye, D.C.; Curé, O.; Blin, G. A survey of RDF storage approaches. *Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées* 2012, *15*, 11-35.

38. Rohloff, K.; Dean, M.; Emmons, I.; Ryder, D.; Sumner, J. An Evaluation of Triple-Store Technologies for Large Data Stores. On the Move to Meaningful Internet Systems 2007: OTM 2007 Workshops, Berlin, Heidelberg, 2007; pp. 1105-1114.

39. Battle, R.; Kolas, D. Enabling the geospatial semantic web with parliament and geosparql. *Semantic Web* 2012, *3*, 355-370.

40. Buccella, A.; Cechich, A.; Fillottrani, P. Ontology-driven geographic information integration: A survey of current approaches. *Computers & Geosciences* 2009, *35*, 710-723.

41. Ballatore, A.; Wilson, D.C.; Bertolotto, M. A survey of volunteered open geo-knowledge bases in the semantic web. In *Quality issues in the management of web information*, Pasi G., B.G., Jain L., Ed.; Springer: Berlin, Heidelberg, 2013; pp. 93-120.

42. Berners-Lee, T.; Dimitroyannis, D.; Mallinckrodt, A.J.; McKay, S. World Wide Web. *Computers in Physics* 1994, *8*, 298-299.

43. O'Reilly, T.; Battelle, J. *Web squared: Web 2.0 five years on*; 144936571X; O'Reilly Media, Inc.: Sebastopol, California, 2009.

44. Murugesan, S. Understanding Web 2.0. *IT Professional Magazine* 2007, *9*, 34.

45. Gandon, F.; Krummenacher, R.; Han, S.-K.; Toma, I. The Resource Description Framework and its Schema. In *Handbook of Semantic Web Technologies*, John Domingue, D.F., James A. Hendler, Ed.; Springer: 2011.

46. Gruber, T.R. A translation approach to portable ontology specifications. *Knowledge acquisition* 1993, *5*, 199-220.

47. Noy, N.F.; McGuinness, D.L. Ontology development 101: a guide to creating your first ontology Stanford University. Available online:

https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html (accessed on 8 April, 2019).

48. Xu, Z.; Lee, Y. Semantic heterogeneity of geodata. *International archives of photogrammetry remote sensing and spatial information sciences* 2002, *34*, 216-224.

49. Zhang, C.; Zhao, T.; Li, W. *Geospatial semantic web*; Springer: Cham, Switzerland, 2015.

50. Battle, R.; Kolas, D. Geosparql: enabling a geospatial semantic web. *Semantic Web Journal* 2011, *3*, 355-370.

51. Perry, M.; Herring, J. OGC GeoSPARQL-A geographic query language for RDF data. OGC Implementation Standard. 2012.

52. Nick, C.; Doerr, M.; Gill, T.; Stead, S.; Stiff, M. Definition of the CIDOC Conceptual Reference Model. CIDOC CRM Implementation Standard. 2011.

53. Felicetti, A.; Samaes, M.; Nys, K.; Niccolucci, F. AnnoMAD: a semantic framework for the management and the integration of full-text excavation data and geographic information. In Proceedings of the 11th International conference on Virtual Reality, Archaeology and Cultural Heritage, September 21 - 24, 2010, 2010; pp. 123-130.

54. Hiebel, G.; Doerr, M.; Eide, Ø. CRMgeo: A spatiotemporal extension of CIDOC-CRM. *International Journal on Digital Libraries* 2017, *18*, 271-279.

55. Emmons, I. Parliament. User Manual. *Raytheon BBN Technologies* 2014.

56. Ravada, S. Spatial and Graph Analytics with Oracle Database 12c Release 2. White Paper. 2016.

57. Erling, O.; Mikhailov, I. Towards Web-Scale RDF. White Paper. n.d.

58. Koubarakis, M.; Karpathiotakis, M.; Kyzirakos, K.; Nikolaou, C.; Sioutis, M. Data Models and Query Languages for Linked Geospatial Data. In *Reasoning Web. Semantic Technologies for Advanced Query Answering. Reasoning Web 2012. Lecture Notes in Computer Science*, Eiter, T., Krennwallner, T., Eds.; Springer: Berlin, Heidelberg, 2012; Volume 7487, pp. 290-328.

59. Scholz, M.; Goerz, G. WissKI: A Virtual Research Environment for Cultural Heritage. ECAI, 2012; pp. 1017-1018.

60. Carlisle, P.; Avramides, I.; Dalgity, A.; Myers, D. The Arches Heritage Inventory and Management System: a standards-based approach to the management of cultural heritage information. In Proceedings of the CIDOC (International Committee for Documentation of the

International Council of Museums) conference: access and understanding–networking in the digital era, Dresden, Germany, 2014; pp. 6-11.

61. Cohen, D. Introducing Omeka. Available online: http://mars.gmu.edu/bitstream/handle/1920/6089/2008-02-20_IntroOmeka.pdf?sequence=2&isAllowed=y (accessed on 4 April).

62. Lagoze, C.; Payette, S.; Shin, E.; Wilper, C. Fedora: an architecture for complex objects and their relationships. *International Journal on Digital Libraries* 2006, *6*, 124-138.

63. 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, Sep 23-27, 2012, 2012.

64. Meghini, C.; Scopigno, R.; Richards, J.; Wright, H.; Geser, G.; Cuy, S.; Fihn, J.; Fanini, B.; Hollander, H.; Niccolucci, F.; et al. ARIADNE: A Research Infrastructure for Archaeology. *Journal on Computing and Cultural Heritage (JOCCH)* 2017, *10*, 1-27, doi:10.1145/3064527.

65. Bardi, A.; Frosini, L. Building a Federation of Digital Humanities Infrastructures. *Digital Humanities* 2017, 28.

66. Hill, L.L. Core Elements of Digital Gazetteers: Placenames, Categories, and Footprints. In *Research and Advanced Technology for Digital Libraries*, Borbinha, J., Baker, T., Eds.; Springer: Berlin, Heidelberg, 2000; Volume 1923, pp. 280-290.

67. Acheson, E.; De Sabbata, S.; Purves, R.S. A quantitative analysis of global gazetteers: Patterns of coverage for common feature types. *Computers, Environment and Urban Systems* 2017, *64*, 309-320, doi:10.1016/j.compenvurbsys.2017.03.007.

68. Laurini, R. Geographic Ontologies, Gazetteers and Multilingualism. *Future Internet* 2015, *7*, 1.

69. Cobb, J. The Journey to Linked Open Data: The Getty Vocabularies. *Journal of Library Metadata* 2015, *15*, 142-156, doi:10.1080/19386389.2015.1103081.

70. Oguz, D.; Ergenc, B.; Yin, S.; Dikenelli, O.; Hameurlain, A. Federated query processing on linked data: a qualitative survey and open challenges. *The Knowledge Engineering Review* 2015, *30*, 545-563, doi:10.1017/S0269888915000107.

71. Baumann, P. Management of multidimensional discrete data. *The VLDB Journal* 1994, *3*, 401-444.

72. Abrate, M.; Bacciu, C.; Hast, A.; Marchetti, A.; Minutoli, S.; Tesconi, M. GeoMemories—A Platform for Visualizing Historical, Environmental

and Geospatial Changes in the Italian Landscape. *ISPRS International Journal of Geo-Information* 2013, *2*, doi:10.3390/ijgi2020432.

73. Agapiou, A.; Lysandrou, V.; Alexakis, D.; Themistocleous, K.; Cuca, B.; Argyriou, A.; Sarris, A.; Hadjimitsis, D. Cultural heritage management and monitoring using remote sensing data and GIS: The case study of Paphos area, Cyprus. *Computers, Environment and Urban Systems* 2015, *54*, 230-239, doi:10.1016/j.compenvurbsys.2015.09.003.

74. Yue, P.; Guo, X.; Zhang, M.; Jiang, L.; Zhai, X. Linked Data and SDI: The case on Web geoprocessing workflows. *ISPRS Journal of Photogrammetry and Remote Sensing* 2016, *114*, 245-257, doi:10.1016/j.isprsjprs.2015.11.009.

75. Bereta, K.; Xiao, G.; Koubarakis, M. ANSWERING GEOSPARQL QUERIES OVER RELATIONAL DATA. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 2017, *42*.

76. Usery, E.L.; Varanka, D. Design and development of linked data from the national map. *Semantic web* 2012, *3*, 371-384.

77. Serna, S.P.; Schmedt, H.; Ritz, M.; Stork, A. Interactive Semantic Enrichment of 3D Cultural Heritage Collections. In Proceedings of the 13th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, Brighton, UK, 2012.

78. Yu, D.; Hunter, J. X3D Fragment Identifiers — Extending the Open Annotation Model to Support Semantic Annotation of 3D Cultural Heritage Objects over the Web. *International Journal of Heritage in the Digital Era* 2014, *3*, 579-596, doi:10.1260/2047-4970.3.3.579.

79. Maietti, F.; Di Giulio, R.; Piaia, E.; Medici, M.; Ferrari, F. Enhancing Heritage fruition through 3D semantic modelling and digital tools: the INCEPTION project. In Proceedings of the Florence Heri-Tech - The Future of Heritage Science and Technologies, Florence, Italy, 2018; pp. 1-8.

80. Doerr, M.; Tzompanaki, K.; Theodoridou, M.; Georgis, C.; Axaridou, A.; Havemann, S. A Repository for 3D Model Production and Interpretation in Culture and Beyond. In Proceedings of the 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2010), Paris, France, 21-24 September, 2010; pp. 97-104.

81. Pan, X.; Schröttner, M.; Havemann, S.; Schiffer, T.; Berndt, R.; Hecher, M.; Fellner, D.W. A Repository Infrastructure for Working with 3D Assets in Cultural Heritage. *International Journal of Heritage in the Digital Era* 2013, *2*, 143-166, doi:10.1260/2047-4970.2.1.143.

82. Hor, A.; Jadidi, A.; Sohn, G. BIM-GIS integrated geospatial information model using semantic web and RDF graphs. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* 2016, *3*, 73-79.

83. Amato, F.; Moscato, V.; Picariello, A.; Colace, F.; Santo, M.; Schreiber, F.; Tanca, L. Big Data Meets Digital Cultural Heritage: Design and Implementation of SCRABS, A Smart Context-awaRe Browsing Assistant for Cultural EnvironmentS. *Journal on Computing and Cultural Heritage (JOCCH)* 2017, *10*, 1-23, doi:10.1145/3012286.

84. Castiglione, A.; Colace, F.; Moscato, V.; Palmieri, F. CHIS: A big data infrastructure to manage digital cultural items. *Future Generation Computer Systems* 2018, *86*, 1134-1145, doi:10.1016/j.future.2017.04.006.

85. Ekbia, H.; Mattioli, M.; Kouper, I.; Arave, G.; Ghazinejad, A.; Bowman, T.; Suri, V.R.; Tsou, A.; Weingart, S.; Sugimoto, C.R. Big data, bigger dilemmas: A critical review. *Advances in Information Science* 2015, *66*, 1523-1545, doi:10.1002/asi.23294.

86. Anuradha, J. A brief introduction on Big Data 5Vs characteristics and Hadoop technology. *Procedia computer science* 2015, *48*, 319-324.

87. Koubarakis, M.; Kyzirakos, K.; Nikolaou, C.; Garbis, G.; Bereta, K.; Dogani, R.; Giannakopoulou, S.; Smeros, P.; Savva, D.; Stamoulis, G.; et al. Managing Big, Linked, and Open Earth-Observation Data: Using the TELEIOS\/LEO software stack. *IEEE Geoscience and Remote Sensing Magazine* 2016, *4*, 23-37, doi:10.1109/MGRS.2016.2530410.

88. Campanaro, D.M.; Landeschi, G.; Dell'unto, N.; Leander Touati, A.-M. 3D GIS for cultural heritage restoration: A 'white box' workflow. *Journal of Cultural Heritage* 2016, *18*, 321-332.

89. Soler, F.; Melero, F.J.; Luzón, M.V. A complete 3D information system for cultural heritage documentation. *Journal of Cultural Heritage* 2017, *23*, 49-57.

90. 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, Crete, Greece, May 29 - June 2, 2016, 2016; pp. 758-773.

91. Trizio, I.; Savini, F.; Giannangeli, A. Integration of three-dimensional digital models and 3D GIS: The documentation of the medieval Burials of Amiternum. *The International Archives of the Photogrammetry* 2018, *42*, 1121-1128.

92. 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, July 3-6, 2017, 2017; pp. 152-168.

93. GeoSPARQL - A Geographic Query Language for RDF Data. Available online: https://www.ogc.org/standards/geosparql (accessed on 8 June 2020).

94. CIDOC CRM. Available online: http://www.cidoc-crm.org/Version/version-5.0.4 (accessed on 8 June 2020).

95. Nentwig, M.; Hartung, M.; Ngonga Ngomo, A.-C.; Rahm, E. A survey of current link discovery frameworks. *Semantic Web* 2017, *8*, 419-436.

96. Knoblock, C.A.; Szekely, P.; Ambite, J.L.; Goel, A.; Gupta, S.; Lerman, K.; Muslea, M.; Taheriyan, M.; Mallick, P. Semi-automatically mapping structured sources into the semantic web. In Proceedings of the 9th Extended Semantic Web Conference, Heraklion, Crete, Greece, May 27-31, 2012; pp. 375-390.

97. 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. *Journal of Web Semantics* 2018, *52*, 16-32.

98. Bereta, K.; Xiao, G.; Koubarakis, M. Answering GeoSPARQL queries over relational data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 2017, *XLII-4/W2*, 43-50.

99. Patroumpas, K.; Alexakis, M.; Giannopoulos, G.; Athanasiou, S. TripleGeo: an ETL Tool for Transforming Geospatial Data into RDF Triples. In Proceedings of the Workshops of the EDBT/ICDT 2014 Joint Conference (EDBT/ICDT 2014), Athens, Greece, March 28, 2014; pp. 275-278.

100. Gali, N.; Mariescu-Istodor, R.; Fränti, P. Similarity measures for title matching. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, December 4-8, 2016; pp. 1548-1553.

101. Nishanbaev, I.; Champion, E.; McMeekin, D.A. A Survey of Geospatial Semantic Web for Cultural Heritage. *Heritage* 2019, *2*, 1471-1498.

102. Remondino, F. 3D recording for cultural heritage. In Proceedings of the 11th EAC Heritage Management Symposium, Reykjavík, Iceland, 25-27 March, 2010; pp. 107-116.

103. Yastikli, N. Documentation of cultural heritage using digital photogrammetry and laser scanning. *Journal of Cultural Heritage* 2007, *8*, 423-427.

104. D'Andrea, A.; Niccolucci, F.; Bassett, S.; Fernie, K. 3D-ICONS: World Heritage sites for Europeana: Making complex 3D models available to everyone. 2012 18th International Conference on Virtual Systems and Multimedia, 2012; pp. 517-520.

105. Petros Patias, P.G., Klaus Hanke. Applications in cultural heritage documentation. In *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*, Zhilin Li, J.C., Emmanuel Baltsavias, Ed.; CRC Press: London, the UK, 2007; pp. 225-257.

106. Scopigno, R.; Callieri, M.; Cignoni, P.; Corsini, M.; Dellepiane, M.; Ponchio, F.; Ranzuglia, G. 3D models for cultural heritage: beyond plain visualization. *Computer* 2011, 48-55.

107. Michela, M.; Chiara Eva, C. 3D Virtual environments as effective learning contexts for cultural heritage. *Italian Journal of Educational Technology* 2018, *26*, doi:10.17471/2499-4324/1026.

108. Rahaman, H.; Champion, E.; Bekele, M. From photo to 3D to mixed reality: A complete workflow for cultural heritage visualisation and experience. *Digital Applications in Archaeology and Cultural Heritage* 2019, *13*, e00102.

109. Felicetti, A.; Lorenzini, M. Metadata and tools for integration and preservation of cultural heritage 3D information. *Geoinformatics FCE CTU* 2011, *6*, 118-124.

110. Auer, M.; Billen, N.; Loos, L.; Zipf, A.; Agugiaro, G. Web-based visualization and query of semantically segmented multiresolution 3D models in the field of cultural heritage. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 2014, *2*.

111. Agugiaro, G.; Remondino, F.; Girardi, G.; Von Schwerin, J.; Richards-Rissetto, H.; De Amicis, R. A web-based interactive tool for multi-resolution 3D models of a Maya archaeological site. In Proceedings of the ISPRS Trento 2011 Workshop, Trento, Italy, 2011; pp. 23-30.

112. Agugiaro, G.; Remondino, F.; Girardi, G.; von SCHWERIN, J.; Richards-Rissetto, H.; de Amicis, R. Queryarch3D: Querying and visualising 3D models of a Maya archaeological site in a Web-based interface. *Geoinformatics FCE CTU* 2011, *6*, 10-17.

113. Kiourt, C.; Koutsoudis, A.; Pavlidis, G. DynaMus: A fully dynamic 3D virtual museum framework. *Journal of Cultural Heritage* 2016, *22*, 984-991.

114. Guarnieri, A.; Pirotti, F.; Vettore, A. Cultural heritage interactive 3D models on the web: An approach using open source and free software. *Journal of Cultural Heritage* 2010, *11*, 350-353.

115. Zhu, C.; Tan, E.C.; Chan, K. 3D Terrain visualization for Web GIS. *Map Asia* 2003, 13-15.

116. McHenry, K.; Bajcsy, P. An overview of 3d data content, file formats and viewers. *National Center for Supercomputing Applications* 2008, *1205*, 22.

117. Barnes, M.; Finch, E.L. COLLADA–Digital Asset Schema Release 1.5.0. Specification Manual. *Khronos Group, Sony Computer Entertainment Inc* 2008.

118. Horsman, G. Web Content Management Systems: An Analysis of Forensic Investigatory Challenges. *Journal of Forensic Sciences* 2018, *63*, 1392-1400, doi:10.1111/1556-4029.13763.

119. Di Giorgio, S.; Felicetti, A.; Martini, P.; Masci, E. Dati. CulturaItalia: a Use Case of Publishing Linked Open Data Based on CIDOC-CRM. In Proceedings of the EMF-CRM@ TPDL, Poznan, Poland, 2015; pp. 44-54.

120. Hiebel, G.; Doerr, M.; Hanke, K.; Masur, A. How to put archaeological geometric data into context? Representing mining history research with CIDOC CRM and extensions. *International Journal of Heritage in the Digital Era* 2014, *3*, 557-577.

121. Kakali, C.; Lourdi, I.; Stasinopoulou, T.; Bountouri, L.; Papatheodorou, C.; Doerr, M.; Gergatsoulis, M. Integrating Dublin Core metadata for cultural heritage collections using ontologies. In Proceedings of the International conference on Dublin core and metadata applications, Singapore, 2007; pp. 128-139.

122. Attene, M.; Robbiano, F.; Spagnuolo, M.; Falcidieno, B. Semantic Annotation of 3D Surface Meshes Based on Feature Characterization. In Proceedings of the 2nd International Conference on Semantic and Digital Media Technologies, Genoa, Italy, 2007; pp. 126-139.

123. Attene, M.; Robbiano, F.; Spagnuolo, M.; Falcidieno, B. Characterization of 3D shape parts for semantic annotation. *Computer-Aided Design* 2009, *41*, 756-763, doi:https://doi.org/10.1016/j.cad.2009.01.003.

124. Singh, A.; Malhotra, M. Agent based framework for scalability in cloud computing. *International Journal of Computer Science & Engineering Technology (IJCSET)* 2012, *3*, 41-45.

125. Lee, J.Y.; Kim, S.D. Software approaches to assuring high scalability in cloud computing. In Proceedings of the 2010 IEEE 7th International

Conference on E-Business Engineering, Shanghai, China, 2010; pp. 300-306.

126. Falatah, M.M.; Batarfi, O.A. Cloud scalability considerations. *International Journal of Computer Science and Engineering Survey* 2014, *5*, 37.

127. Bharadwaj, S.S.; Lal, P. Exploring the impact of Cloud Computing adoption on organizational flexibility: A client perspective. In Proceedings of the 2012 International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM), 2012; pp. 121-131.

128. Sultan, N. Making use of cloud computing for healthcare provision: Opportunities and challenges. *International Journal of Information Management* 2014, *34*, 177-184.

129. De Assunção, M.D.; Di Costanzo, A.; Buyya, R. Evaluating the cost-benefit of using cloud computing to extend the capacity of clusters. In Proceedings of the 18th ACM international symposium on high performance distributed computing, Garching, Germany, 2009; pp. 141-150.

130. Chandra, D.G.; Borah, M.D. Cost benefit analysis of cloud computing in education. In Proceedings of the 2012 International Conference on Computing, Communication and Applications, Dindigul, India, 22-24 February, 2012; pp. 1-6.

131. Kats, P.; Mielnicki, M.; Knoth, P.; Muhr, M.; Mamakis, G.; Werla, M. Design of Europeana Cloud technical infrastructure. In Proceedings of the IEEE/ACM Joint Conference on Digital Libraries, London, UK, 2014; pp. 491-492.

132. Benardou, A.; Dallas, C.; Dunning, A. From Europeana Cloud to Europeana Research: The challenges of a community-driven platform exploiting Europeana content. In Proceedings of the Euro-Mediterranean Conference, Limassol, Cyprus, 2014; pp. 802-810.

133. Sacher-Project. Sacher project homepage. Available online: http://www.eng.sacherproject.com/ (accessed on 28 November 2019).

134. Bertacchi, S.; Jawarneh, A.; Mashhour, I.; Apollonio, F.I.; Bertacchi, G.; Cancilla, M.; Foschini, L.; Grana, C.; Martuscelli, G.; Montanari, R. SACHER Project: A Cloud Platform and Integrated Services for Cultural Heritage and for Restoration. In Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good, Bologna Italy, 2018; pp. 283-288.

135. Apollonio, F.I.; Rizzo, F.; Bertacchi, S.; Dall'Osso, G.; Corbelli, A.; Grana, C. SACHER: smart architecture for cultural heritage in Emilia Romagna. In Proceedings of the Italian Research Conference on Digital Libraries, Bari, Italy, 2017; pp. 142-156.

136. Ariadne-Infrastructure-Project. Ariadne Infrastructure project homepage. Available online: https://ariadne-infrastructure.eu/ (accessed on 28 November 2019).

137. E-RIHS-Project. E-RIHS project homepage. Available online: (accessed on 28 November 2019).

138. ownCloud. ownCloud project homepage. Available online: (accessed on 28 November 2019).

139. Virtual-Planet-Builder. Virtual Planet Builder project homepage. Available online: (accessed on 28 November 2019).

140. Fanini, B.; Pescarin, S.; Palombini, A. A cloud-based architecture for processing and dissemination of 3D landscapes online. *Digital Applications in Archaeology and Cultural Heritage* 2019, e00100.

141. EUDAT. EUDAT project homepage. Available online: (accessed on 28 November 2019).

142. EUDAT. EUDAT-Cloud-like-services-to-improve-the-preservation-of-digital-cultural-heritage. Available online: https://eudat.eu/communities/cloud-like-services-to-improve-the-preservation-of-digital-cultural-heritage (accessed on 28 November 2019).

143. Pisu, C.; Casu, P. Cloud GIS and 3d modelling to enhance sardinian late gothic architectural heritage. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2013, *5*, W2.

144. Salesforce. Salesforce homepage. Available online: https://www.salesforce.com/ (accessed on 28 November 2019).

145. Heroku. Heroku homepage. Available online: https://www.heroku.com/ (accessed on 28 November 2019).

146. Amazon-EC2. Amazon EC2 homepage. Available online: https://aws.amazon.com/ec2/ (accessed on 28 November 2019).

147. Google-Compute-Engine. Google Compute Engine homepage. Available online: https://cloud.google.com/compute/ (accessed on 28 November 2019).

148. DigitalOcean. DigitalOcean homepage. Available online: https://www.digitalocean.com/ (accessed on

149. Zhang, Q.; Cheng, L.; Boutaba, R. Cloud computing: state-of-the-art and research challenges. *Journal of Internet Services and Applications* 2010, *1*, 7-18, doi:10.1007/s13174-010-0007-6.

150. Ali, M.; Khan, S.U.; Vasilakos, A.V. Security in cloud computing: Opportunities and challenges. *Information sciences* 2015, *305*, 357-383.

151. Zafar, F.; Khan, A.; Malik, S.U.R.; Ahmed, M.; Anjum, A.; Khan, M.I.; Javed, N.; Alam, M.; Jamil, F. A survey of cloud computing data integrity schemes: Design challenges, taxonomy and future trends. *Computers & Security* 2017, *65*, 29-49.

152. McKeever, S. Understanding Web content management systems: evolution, lifecycle and market. *Industrial management & data systems* 2003, *103*, 686-692.

153. WebGL. WebGL project homepage. Available online: https://www.khronos.org/webgl/ (accessed on 28 November 2019).

154. W3Techs. W3Techs - Usage of JavaScript libraries for websites website. Available online: https://w3techs.com/technologies/overview/javascript_library (accessed on 28 November 2019).

155. Hyvönen, E. Linked Open Data Infrastructure for Digital Humanities in Finland. In Proceedings of the Digital Humanities in Nordic Countries (DHN 2020), Riga, Latvia, 2020; pp. 254-259.

156. Chanhom, W.; Anutariya, C. TOMS: A linked open data system for collaboration and distribution of cultural heritage artifact collections of National Museums in Thailand. *New Generation Computing* 2019, *37*, 479-498.

157. Candela, G.; Escobar, P.; Carrasco, R.C.; Marco-Such, M. Migration of a library catalogue into RDA linked open data. *Semantic Web* 2018, *9*, 481-491.

158. Simou, N.; Chortaras, A.; Stamou, G.; Kollias, S. Enriching and publishing cultural heritage as linked open data. In *Mixed Reality and Gamification for Cultural Heritage*; Springer: 2017; pp. 201-223.

159. Daquino, M.; Mambelli, F.; Peroni, S.; Tomasi, F.; Vitali, F. Enhancing semantic expressivity in the cultural heritage domain: exposing the Zeri Photo Archive as Linked Open Data. *Journal on Computing and Cultural Heritage (JOCCH)* 2017, *10*, 1-21.

160. de Boer, V.; Wielemaker, J.; van Gent, J.; Oosterbroek, M.; Hildebrand, M.; Isaac, A.; van Ossenbruggen, J.; Schreiber, G. Amsterdam museum linked open data. *Semantic Web* 2013, *4*, 237-243.

161. Nishanbaev, I.; Champion, E.; McMeekin, D.A. A Comparative Evaluation of Geospatial Semantic Web Frameworks for Cultural Heritage. *Heritage* 2020, *3*, 875-890.

162. Nishanbaev, I. A Cloud Architecture for Processing and Visualization of Geo-located 3D Digital Cultural Heritage Models. In Proceedings of the 6th International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM 2020), Prague, Czech Republic, 2020; pp. 51-61.

163. Machidon, O.-M.; Tavčar, A.; Gams, M.; Duguleană, M. CulturalERICA: A conversational agent improving the exploration of European cultural heritage. *Journal of Cultural Heritage* 2020, *41*, 152-165.

164. De Fino, M.; Ceppi, C.; Fatiguso, F. Virtual Tours and Informational Models for Improving Territorial Attractiveness and the Smart Management of Architectural Heritage: the 3D-IMP-ACT Project. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 2020, *44*, 473-480.

165. Candela, G.; Escobar, P.; Carrasco, R.C.; Marco-Such, M. A linked open data framework to enhance the discoverability and impact of culture heritage. *Journal of Information Science* 2019, *45*, 756-766.

166. Quattrini, R.; Pierdicca, R.; Morbidoni, C. Knowledge-based data enrichment for HBIM: Exploring high-quality models using the semantic-web. *Journal of Cultural Heritage* 2017, *28*, 129-139.

167. Lehmann, J.; Isele, R.; Jakob, M.; Jentzsch, A.; Kontokostas, D.; Mendes, P.N.; Hellmann, S.; Morsey, M.; Van Kleef, P.; Auer, S. Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web* 2015, *6*, 167-195.

168. Mendes, P.N.; Jakob, M.; García-Silva, A.; Bizer, C. DBpedia spotlight: shedding light on the web of documents. In Proceedings of the 7th international conference on semantic systems, Graz, Austria, 2011; pp. 1-8.

169. Dang, Q.V.; Ignat, C.-L. Quality assessment of wikipedia articles without feature engineering. In Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries, New Jersey, USA, 2016; pp. 27-30.

170. Zhang, S.; Hu, Z.; Zhang, C.; Yu, K. History-based article quality assessment on Wikipedia. In Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), Shanghai, China, 2018; pp. 1-8.

171. Shen, A.; Qi, J.; Baldwin, T. A hybrid model for quality assessment of Wikipedia articles. In Proceedings of the Australasian Language Technology Association Workshop 2017, Brisbane, Australia, 2017; pp. 43-52.

172. di Sciascio, C.; Strohmaier, D.; Errecalde, M.; Veas, E. WikiLyzer: interactive information quality assessment in Wikipedia. In

Proceedings of the 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 2017; pp. 377-388.

173. Ahlers, D. Assessment of the accuracy of GeoNames gazetteer data. In Proceedings of the 7th Workshop on Geographic Information Retrieval, New York, United States, 2013; pp. 74-81.

174. Vinasco-Alvarez, D.; Samuel, J.; Servigne, S.; Gesquière, G. From CityGML to OWL. Technical Report. 2020.

175. Vinasco-Alvarez, D.; Samuel, J.; Servigne, S.; Gesquière, G. Towards a semantic web representation from a 3D geospatial urban data model. In Proceedings of the 16th Spatial Analysis and Geomatics Conference, La Rochelle, France, 5-7 May, 2021; pp. 227-238.

176. Banerjee, I.; Agibetov, A.; Catalano, C.E.; Patanè, G.; Spagnuolo, M. Semantic Annotation of Patient-Specific 3D Anatomical Models. In Proceedings of the 2015 International Conference on Cyberworlds (CW), Visby, Sweden, 7-9 October, 2015; pp. 22-29.

177. Banerjee, I.; Catalano, C.E.; Patané, G.; Spagnuolo, M. Semantic annotation of 3D anatomical models to support diagnosis and follow-up analysis of musculoskeletal pathologies. *International Journal of Computer Assisted Radiology and Surgery* 2016, *11*, 707-720, doi:10.1007/s11548-015-1327-6.

178. Chadzynski, A.; Krdzavac, N.; Farazi, F.; Lim, M.Q.; Li, S.; Grisiute, A.; Herthogs, P.; von Richthofen, A.; Cairns, S.; Kraft, M. Semantic 3D City Database—An enabler for a dynamic geospatial knowledge graph. *Energy and AI* 2021, *6*, 100106.

179. Chadzynski, A.; Li, S.; Grisiute, A.; Farazi, F.; Lindberg, C.; Mosbach, S.; Herthogs, P.; Kraft, M. Semantic 3D City Agents—An intelligent automation for dynamic geospatial knowledge graphs. *Energy and AI* 2022, 100137.