Digital Prosopography Information in Virtual Record Treasury of Ireland Knowledge Graph

Beyza Yaman[1,*], Lucy McKenna[1], Alex Randles[1], Lynn Kilgallon[2], Peter Crooks[2] and Declan O'Sullivan[1]

[1]*ADAPT Centre, SCSS, Trinity College Dublin, Dublin, Ireland*
[2]*Department of History, Trinity College Dublin, Dublin, Ireland*

Abstract

The Virtual Record Treasury of Ireland (VRTI), a virtual archive currently in the third phase of its research programme (2023-2025), places a key emphasis on utilising Semantic Web technologies to further construct a comprehensive knowledge graph (KG) of historical Irish biographical data. To do this, computer scientists and historians are collaborating closely to enhance the VRTI Knowledge Graph for Irish History (VRTI-KG) in a multidisciplinary manner. Biographical data was uplifted to RDF format as a result of the collaboration, and the prosopographical graph currently includes 965 women and 8807 notable men from Irish history. Apart from employing the out-of-the-box tools, a new tailored VRTI-KG Editor has been introduced to facilitate the exploration of the VRTI-KG. This paper describes the VRTI-KG's technical architecture and provides an overview of the tools being used to interact with and explore VRTI-KG, as illustrated by a use case using a notable people whose graph is displayed.

Keywords

Linked Data, Visualization Tools, Prosopography, Digital Humanities

1. Introduction

The Virtual Record Treasury of Ireland (VRTI)[1] [1, 2, 3] is an all-island and international legacy from Ireland's Decade of Centenaries. The VRTI digital resource is the outcome of the Beyond 2022 project – a seven-year programme of State-funded research hosted at Trinity College Dublin, which combines historical investigation, archival conservation and technical innovation to re-imagine and reconstruct Ireland's national treasury of records lost in a catastrophic fire in 1922 [3]. The Irish Civil War commenced on June 28th, 1922, when the Four Courts in Dublin, which had been occupied by military forces opposed to the Anglo-Irish Treaty of 1921, was attacked by the National Army of the Provisional Government of the Irish Free State. On the third day of the bombardment of the Four Courts, a massive explosion caused significant damage to the buildings in the western part of the Four Courts complex [3]. The fire spread to nearby

SemDH'24: First International Workshop of Semantic Digital Humanities, May 28, 2024, Creete, Greece
*Corresponding author.
beyza.yaman@adaptcentre.ie (B. Yaman); lucy.mckenna@adaptcentre.ie (L. McKenna); alex.randles@adaptcentre.ie (A. Randles); kilgalll@tcd.ie (L. Kilgallon); pcrooks@tcd.ie (P. Crooks); declan.osullivan@adaptcentre.ie (D. O'Sullivan)
0000-0003-2130-0312 (B. Yaman); 0000-0002-6035-7656 (L. McKenna); 0000-0001-6231-3801 (A. Randles); 0000-0002-3075-8571 (L. Kilgallon); 0000-0001-6782-044X (P. Crooks); 0000-0003-1090-3548 (D. O'Sullivan)

[1]Virtual Record Treasury of Ireland at https://virtualtreasury.ie/

buildings, including the Public Record Office of Ireland (PROI), which held over seven centuries of Ireland's documentary heritage in a purpose-built archival repository known as the "Record Treasury". A century later, in June 2022, the VRTI was launched. The Beyond 2022 project decided to use Semantic Web technologies to support the powerful knowledge distribution and reasoning and has made available the VRTI Knowledge Graph (VRTI-KG) of Irish History.[2]. An initial presentation of the development process used to produce the VRTI-KG was published in 2022 [14]. The VRTI-KG contains knowledge of notable People, Places, Offices, Organisations, and Interests and their interconnections, from the records of Irish history. This new phase of research (2023-2025)[3] contains six research strands, five of which are enhancing the materials available through the VRTI and the sixth focusing on further enhancing the VRTI-KG itself and the development of better and more appropriate user interfaces to the KG.

The VRTI takes a multidisciplinary approach to KG management through the collaboration of computer scientists and historians in requirements gathering, interface design and project planning. Semantic Web technologies have been employed to address the project's requirements which include: i) using a structured data format for the purpose of interconnecting related entities based on time periods and entity types; ii) offering data lineage so that the origins of data may be tracked; iii) presenting data in an engaging manner to the general public; iv) offering historians a user interface for finding, updating, editing and publishing KG data.

In the scope of the first phase of the VRTI's research programme, an ontology extending CIDOC-CRM [4] was developed and published online as the B2022 ontology[4]. The KG is modelled on this ontology and a Named Graph infrastructure is also used to contextualise entities depending on the class types (e.g. place, person etc.). Data lineage is provided by adopting the PROV-O ontology to express when the KG was updated and by whom. Interaction with the VRTI-KG is currently enabled through the use of existing out-of-the-box Knowledge Graph tools including OSCAR [5, 6], Ontodia [7], LodView[5] and Beyond Timelines[6] which can be accessed through VRTI website[7]. However, these out-of-the-box tools have been designed primarily for use by knowledge engineers, as they have to cover a wide range of deployment scenarios, from event-based networks [8] through to classical data integration [9] scenarios. The contributions of this paper include the technical design considerations of the VRTI Knowledge Graph for Irish History, the implementation of the VRTI-KG Editor - a bespoke user interface designed for use by the public and historians to interact with the VRTI-KG and the presentation of how out-of-the-box tooling can be used to demonstrate prosopographical data for people represented in the KG.

The rest of the paper is structured as follows: Section 2 discusses related work focusing on prosopographical data. Section 3 provides a use-case from the VRTI-KG. Section 4 discusses the ontology and tools which are used to explore the VRTI-KG. Section 5 presents the technical architecture of the project. Finally, Section 6 concludes the paper and discusses future work.

[2]https://virtualtreasury.ie/knowledge-graph

[3]Deep History, Deepening Collaborations, to be found at: https://virtualtreasury.ie/backend/flipbooking/vrti-101-brochure/

[4]https://ont.virtualtreasury.ie/ontology/index-en.html

[5]https://lodview.it/

[6]https://github.com/Beyond-2022/Beyond-2022.github.io

[7]https://virtualtreasury.ie/knowledge-graph

2. Related Work

This paper presents how the VRTI is using out-of-the-box Linked Data tooling to provide users with the ability to explore biographical data in the VRTI-KG. In addition, work is underway on a customised VRTI-KG Editor to enhance that user experience. In this section, we first discuss the rationale for taking both an out-of-the-box approach and adding to it with a customised approach to support user interaction with the VRTI-KG.

AcademySampo [10, 11] discusses the initial outcomes of converting the Finnish registries "Ylioppilasmatrikkeli" 1640–1852 and 1853–1899, containing comprehensive biographical information on academic individuals in Finland, into a Linked Open Data service. This project adheres to the FAIR principles and uses named entity recognition and linking techniques, with the intention of creating a Semantic Web portal named "AcademySampo" for biographical and prosopographical research.

BiographySampo [12, 13] introduces a Semantic Web portal designed to illustrate and explore the National Biography of Finland (NBF). The foundation of the system is an automatically derived KG from a set of 13,100 textual NBF biographies, enhanced with information pointing to sixteen additional data sources obtained through data harvesting from external collections found in archives, museums, and libraries.

WarSampo Knowledge Graph [14, 15], is a Linked Open Data resource with detailed metadata on over 100,000 individuals related to Finland in World War II. The WarSampo KG showcases and facilitates the visualisation and analysis of prosopographical phenomena including the death records of WW2 casualties from a prosopographical perspective, provided by the various local military cemeteries where the dead were buried. The problem is solved using aggregated Linked Open Data provided by the WarSampo Data Service and SPARQL endpoint, and providing tools for data analysis and supporting digital humanities studies.

People of Medieval Scotland (POMS) project[8] [16] encompasses information on individuals involved in events in or related to Scotland between the death of Malcolm III in 1093 and Robert I's parliament in 1314, covering territory established as part of Scotland by the death of Alexander III (excluding Orkney and Shetland). The data is structured to reflect social interactions and relationships, providing insights into how these were mediated by the documents themselves using a general "factoid-oriented" model that links people to the information about them via references to primary sources that assert that information.

Digital Prosopography of the Roman Republic [17] aims to enhance prosopographical research on the Roman Republic's elite by creating a searchable KG database with a terminological approach providing information on careers, office holdings, family relationships and personal status, thereby facilitating new research possibilities through statistical and quantitative methods. The triple based "factoid model" of digital prosopography has been applied to major databases dealing with medieval England, Scotland, France, and Byzantium.

History of Vienna with Semantic MediaWiki project [18] examines the role of Semantic MediaWiki (SMW) in knowledge graph curation, focusing on the Vienna History Wiki as a case study. It discusses collaborative editing processes, linking unique identifiers, and integration with external sources like Wikidata and Schema.org. The project includes entities such as

[8] www.poms.ac.uk

people, events, organizations, and places from Vienna.

Within the project scope, in contrast to the works mentioned above, first out-of-the-box tools were introduced that did not require any data or tool transformation. They were quick to implement and deploy to the server. Furthermore, a customised tool has been implemented for VRTI-KG for a number of reasons. Other elements, such as maps and events with unique underlying structures, should be included in addition to prosopographical data. The decision to recently start a new user interface development for VRTI-KG was informed by the observations that; i) the out-of-the-box features, although usable by the historians in their engagement with the VRTI-KG, still had a very unnatural technical feel for them; ii) more appealing methods of presenting data are now possible thanks to recent advancements in visualisation technologies.

3. Use Case

The VRTI-KG helps explore the intertwined narratives of individuals and locations in Ireland, spanning from the medieval to the modern era. Currently, much of the data regarding historical people comes from the Dictionary of Irish Biography (DIB)[9] a project of the Royal Irish Academy, and the persons listed in Philomena Connolly's calendar of medieval Irish Exchequer Payments, 1270-1446 [19]. The DIB is an authoritative reference work of nearly 11,000 notable figures in Irish history, society and culture from the earliest times to the twenty-first century. Biographies range in length from 200 to 15,000 words, covering diverse figures across a broad range of human activity from scientists to sportspeople, or from suffragists to soldiers. The individuals drawn from Connolly's calendar of Irish Exchequer payments were those who received payments from the English government in Ireland between 1270-1446. The data sources mentioned above provide biographical data in text and CSV formats based on a structured database which is highly curated. Using an unstructured/semistructured data format has the following two drawbacks: Firstly, despite the sources having a wealth of important information, users may find it difficult to locate the precise fact or piece of information they are searching for. Secondly, finding the relationships between the individuals is difficult because the texts do not typically indicate how they are related to one another. In the scope of this work, only semistructured data sets are uplifted to RDF format.

The VRTI team has created a person schema that allows for the CSV capture of biographical data for individuals represented in the sources in order to uplift the data into a rich graph-based structured data format that makes the information about individuals, and connections between individuals, more easily navigable by a user and allows for further information to be easily linked in as new data sources are processed. The current person schema includes concepts such as forename, surname and their variant spellings, gender, date and place of birth and death, family relations, religion, occupation, areas of interest and, where possible, interlinks to related records in the DIB and Wikidata[10]. Using the VRTI ontology as a data model, information about people is uplifted from the data sources into RDF[11], a standard Linked Data format, using the

[9] https://www.dib.ie/
[10] https://www.wikidata.org/wiki/Wikidata:Introduction
[11] https://www.w3.org/RDF/

R2RML[12] mapping language. In total, the VRTI-KG contains 8807 men and 965 women from Irish history uplifted from the DIB and Irish Exchequer Payments 1270-1446.

The poet and playwright Oscar Wilde[13] is one of the notable people in the DIB whose biographical data was uplifted to the VRTI-KG. Listing 1 presents a partial snippet of the uplifted RDF triples for Oscar Wilde. In the first set of triples, it can be seen that Oscar Wilde is classified as a cidoc:E21_Person who is listed in (cidoc:P71i_is_listed_in) the DIB at <https://www.dib.ie/biography/Wilde-Oscar-Fingal-OFlahertie-a9036>. Wilde's areas of interest (b2022:DIB_area_of_interest) have been declared as "Literature" and "Theatre, Film and TV". Properties from the VRTI ontology are represented using the prefix *"b2022"*. These properties were added in cases where a suitable property was not available in CIDOC-CRM.

The second set of triples provides information on the birth of Oscar Wilde. This birth event is given a date (cidoc:P4_has_time-span) and a geographic location (cidoc:P7_took_place_at). This event is attributed to Oscar Wilde using the inverse property (cidoc:P98_brought_into_life). CIDOC-CRM uses time-spans to describe dates as this allows for the provision of date ranges where there is historical uncertainty. However, as Oscar Wilde's birth and death dates are known, these specific dates are used as the range. Geospatial information can also be provided here by linking a birth or death event with a location. The VRTI-KG data for Oscar Wilde is available for exploration using a set of Linked Data tools discussed in Section 4.

```
@prefix cidoc: <http://erlangen-crm.org/current/> .
@prefix xsd:   <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#> .
@prefix b2022: <https://ont.virtualtreasury.ie/ontology#> .


<https://kb.virtualtreasury.ie/person/Wilde_Oscar-Fingal-OFlahertie-Wills_c19_dib_a9036
    > a cidoc:E21_Person ;
cidoc:P71i_is_listed_in <https://www.dib.ie/biography/Wilde-Oscar-Fingal-OFlahertie-
    a9036> ;
b2022:DIB_area_of_interest
    <https://kb.virtualtreasury.ie/dib-area-of-interest/IE/Theatre-Film-and-TV> ,
    <https://kb.virtualtreasury.ie/dib-area-of-interest/IE/Literature> .


<https://kb.virtualtreasury.ie/birth/Wilde_Oscar-Fingal-OFlahertie-Wills_c19_dib_a9036>
    a cidoc:E67_Birth ;
cidoc:P4_has_time-span <https://kb.virtualtreasury.ie/time-span/1854-10-16_1854-10-16>;
cidoc:P7_took_place_at <https://kb.virtualtreasury.ie/place/IE/Dublin> ;
cidoc:P98_brought_into_life <https://kb.virtualtreasury.ie/person/Wilde_Oscar-Fingal-
    OFlahertie-Wills_c19_dib_a9036> .
```

Listing 1: Partial set of triples pertaining to Oscar Wilde

4. VRTI Ontology and KG Exploration Tools

This section describes the VRTI Ontology and exploration tools to visualise the prosopographical information in the VRTI-KG

[12] https://www.w3.org/TR/r2rml/

[13] https://www.dib.ie/biography/wilde-oscar-fingal-oflahertie-a9036

4.1. VRTI Ontology: Use of Standard Ontologies

The VRTI ontology[14] primarily consists of named individuals for qualifying entities in the CIDOC Conceptual Reference Model (CIDOC-CRM). CIDOC-CRM is a high-level, event-based ontology of human activity, things and events, which provides a formal structure for describing concepts and relationships commonly used in the cultural heritage domain [20]. It is recognised as an international standard (ISO 21127:2014)[15] for the controlled exchange of cultural heritage information. CIDOC-CRM (version 7.1.2) consists of 81 classes and 160 properties. In order to accurately model Irish historical data, the VRTI ontology introduced bespoke types that are related to entities via the cidoc:P2_has_type predicate (instead of creating instances with rdf:type). Examples of these bespoke types, all of which are instances of cidoc:E55_Type, include *b2022:Floruit* to represent the period in which a person "flourished" (Latin: floreo, 'to bloom, flourish'), b2022:Office, b2022:Occupation, and b2022:Rank for qualifying some of the professional groups to which a person belonged. Other types describe different kinds of places such as b2022:Chapel, b2022:Manor, b2022:Harbour and b2022:Friary. Types were used to ensure maximum backwards compatibility with the CIDOC-CRM model. Twelve object properties were also created as part of the VRTI Ontology. These properties were added where there was no existing CIDOC-CRM property available that was suitable to express a given relationship. For example, the properties b2022:ofHusband and b2022:ofWife were added in order to link a marriage event to the individuals being married. The ontology has been published according to best practices in Linked Data and its documentation and behaviour according to Linked Data principles have been generated with WIDOCO [21].

4.2. Prosopography Exploration in VRTI-KG

As mentioned in Section 2, a prosopography of the Irish poet and playwright Oscar Wilde will be presented using data from the VRTI-KG via the KG exploration tools employed by the project.

4.2.1. Oscar

Oscar[16] is an OpenCitations RDF Search Application, which can be used to search any RDF triplestore providing a SPARQL endpoint. Oscar provides a user-friendly search interface in that it hides the complexities of SPARQL queries from non-Linked Data experts. For VRTI, Oscar has been configured to search for terms entered by users in a search bar in the rdfs:label attribute associated with an entity. Fig.1 shows the results of a search for Oscar Wilde in the VRTI-KG using the Oscar tool. The results include the HTTP URI (unique resource identifier) for Oscar Wilde's appellation as well as his entry in the DIB. By clicking on the URI, users are redirected to the LodView page describing the entity.

[14] https://ont.virtualtreasury.ie/ontology/index-en.html
[15] https://www.iso.org/standard/57832.html
[16] https://oscar.virtualtreasury.ie/oscar/index.html

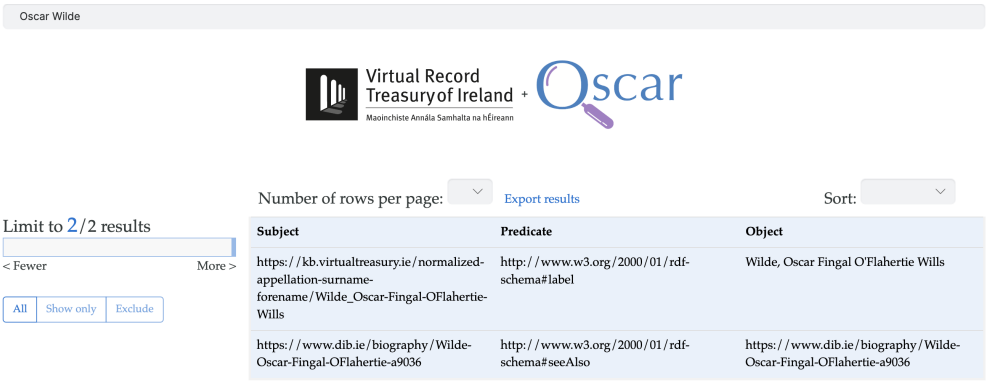

Figure 1: OSCAR search results for Oscar Wilde.

4.2.2. LodView

LodView[17] is a web application used to dereference the URIs of RDF resources. When a user requests the URI of an RDF resource, LodView retrieves the data about the resource from the SPARQL endpoint and presents this data as a human-readable HTML page.

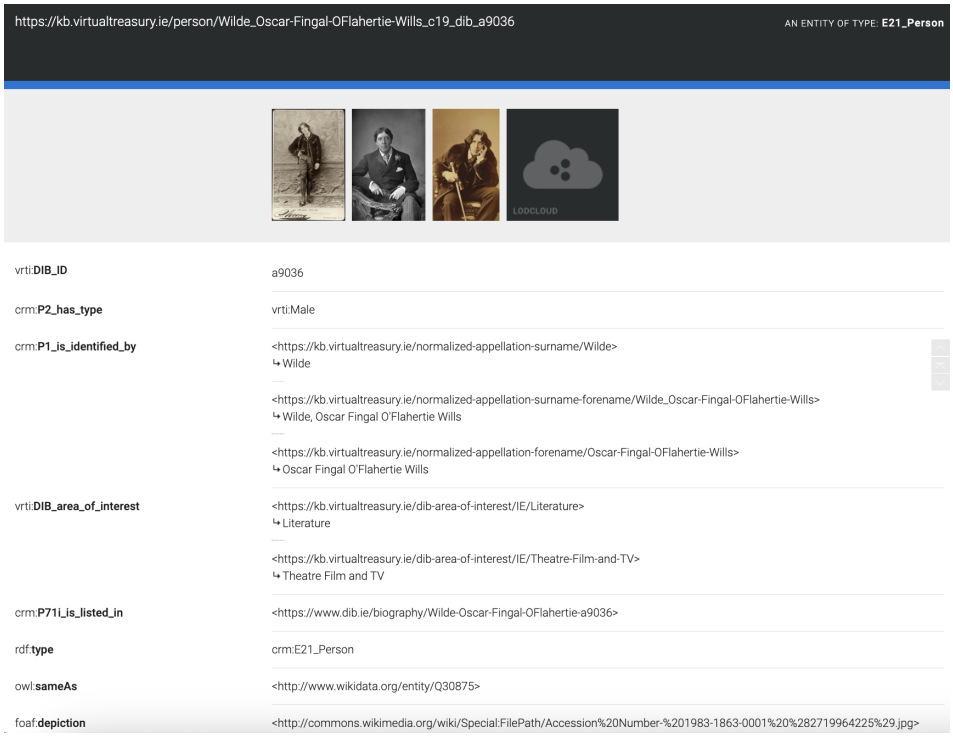

Figure 2: LodView generated HTML page for Oscar Wilde.

[17] https://kb.virtualtreasury.ie/lodview/

On the HTML page, the resource properties are represented in rows and columns. In each row, the first column contains the property name and the second column contains the property value. If a property value is the URI of another resource, then the value is provided as a clickable link that takes users to another page describing that resource. Fig.2 shows the LodView representation of some of the data pertaining to resource for Oscar Wilde in the VRTI-KG LodView interface. Here you can see information such as a link to Oscar Wilde's WikiData entry, gender, interests and DIB identifier.

4.2.3. Ontodia

Ontodia[18] is a visualisation tool that allows users to interactively explore entities in a Knowledge Graph and dynamically navigate their connections. For the VRTI, a SPARQL query has been designed that retrieves all resources in the KG and Ontodia presents these results in a node and edge visualisation. Fig.3 displays how a user could explore information about Oscar Wilde[19] in the VRTI-KG. The nodes in the image represent Oscar Wilde, his name, his date of birth and his date of death. Users can click on these nodes and discover other attributes and resources associated with Oscar Wilde in the KG.

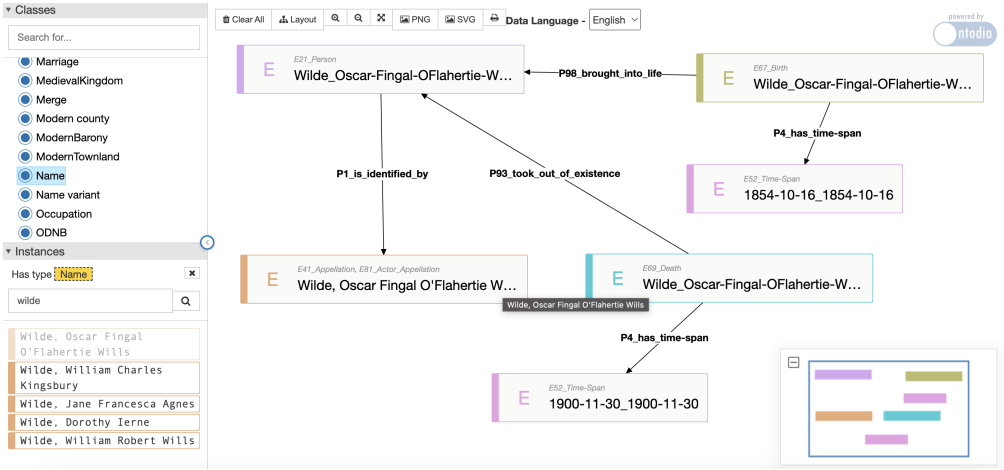

Figure 3: Ontodia View of Oscar Wilde.

4.2.4. BeyondTimelines

Beyond Timelines[20] is a bespoke data visualisation tool developed for the VRTI[21]. By querying the VRTI SPARQL endpoint, the Beyond Timelines tool can generate a visualisation of people born within a specified time-frame as selected by the user. The resulting image displays all persons in the VRTI-KG born within the selected dates, and also colour-codes these individuals

[18] https://ontodia.virtualtreasury.ie/ontodia/
[19] https://kb.virtualtreasury.ie/person/Wilde_Oscar-Fingal-OFlahertie-Wills_c19_dib_a9036
[20] https://github.com/Beyond-2022/Beyond-2022.github.io
[21] https://timelines.virtualtreasury.ie/timelines/

based on their associated interests. Fig.4 displays the result of a search for people born between 1854 and 1855. Oscar Wilde's entry can be seen and is colour-coded according to the areas of interest of literature and theatre. Using this tool, users can discover other persons born around the same time as Oscar Wilde and they can get a sense of the areas of interests of these individuals.

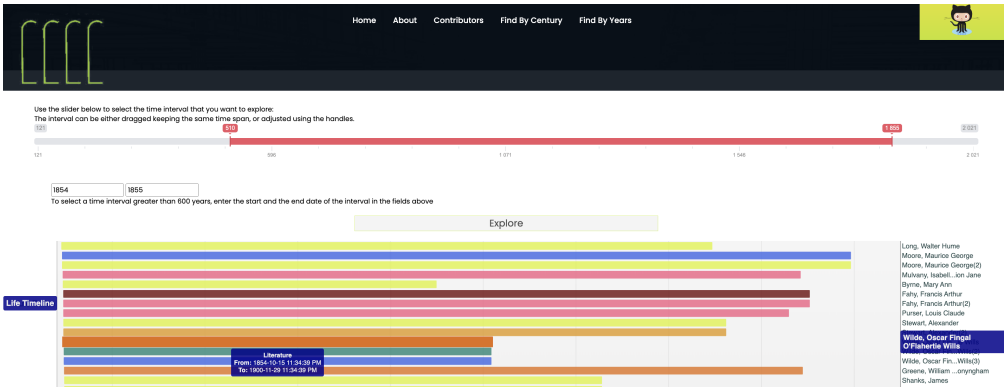

Figure 4: Beyond Timelines view for people born between 1854 and 1855.

4.2.5. VRTI-KG Editor

The VRTI knowledge graph editor has been designed by the VRTI-KG team to allow subject-matter experts (historians) with limited relevant technical background knowledge to search and edit entities in the VRTI graph. The web-based editor provides a graphical user interface (GUI) that transforms form input provided by end users into SPARQL queries, which are executed on the graph. The editor provides sections to allow editing of entities representing people, places, offices and organisations. Moreover, it allows to visualize data

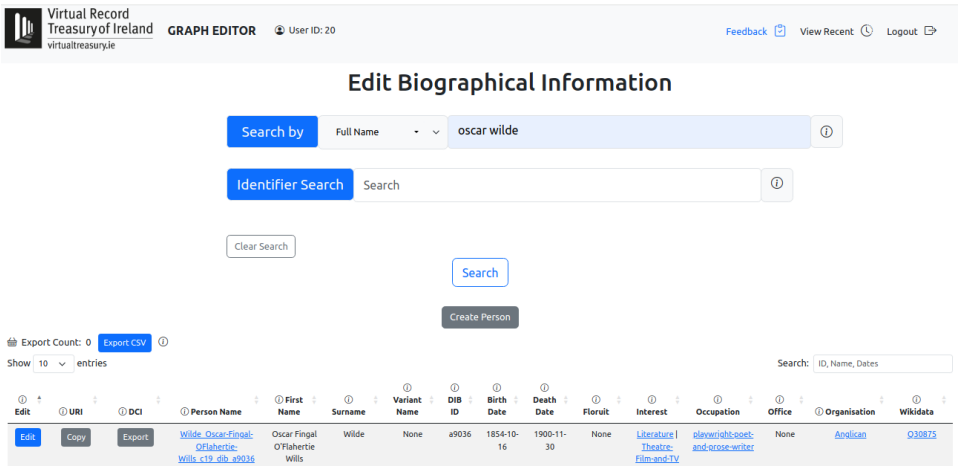

Figure 5: Search results for people with the name "Oscar Wilde" in the editor.

As can be seen in Fig.5, several people are retrieved from the execution of the SPARQL query which was generated by inserting the string “Oscar Wilde” into a filter condition. The retrieved people are presented in tabular format with columns detailing distinguishable information, which is hoped to aid in identifying the correct entity to edit. In addition, the table can be searched by keyword to find the respective entity.

Fig.6 presents a screenshot of the main editing page for a person in the editor. The editing page contains several headings at the top, which group editable attributes based on similarities hoped to improve navigation. The inputs presented on the page are a combination of free text, dropdown menus and restricted text. Validation is completed on the input to help prevent inaccurate data from being inserted into the graph. A status is associated with each entity to provide an indication of the level of associated information and whether it requires further editing.

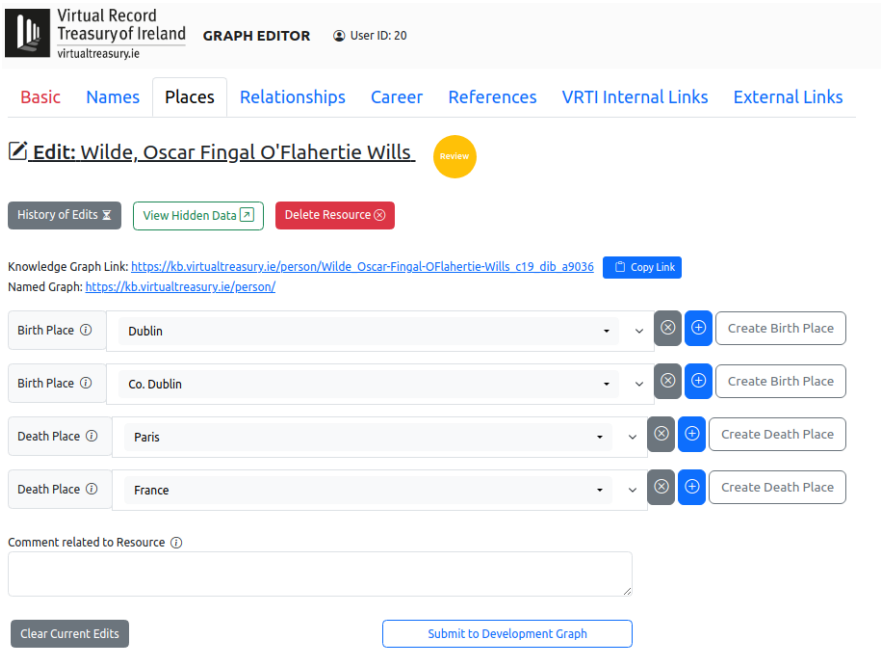

Figure 6: Edit page for a person in the editor.

Producing high-quality data is essential when inserting triples using the VRTI-KG Editor since errors in the data source can lead to unreliability [22]. Thus, SHACL Shapes are used to validate that data graphs meet a set of requirements. To guarantee that the resources produced are of a high-quality, SHACL shapes have been created for the entities that the graph editor has produced as part of this work. Graphs which do not satisfy these conditions result in debugging information being shown to users and the creation will be halted. The SHACL rules are currently generated manually, however, we are investigating automatic rule generation methods that leverage the underlying ontology structure to construct SHACL Shapes [23]. A user evaluation of the VRTI-KG editor with historians is currently underway and it is hoped to publish the results when available.

5. Technical Architecture

This section discusses the technical architecture and technical decisions made during the project.

5.1. KG Architecture Diagram

In this project, two different virtual machines are being used to provide a secure environment for the production and deployment of the systems, as well as providing a seamless infrastructure for public users.

The first virtual machine is called the production server, where web applications using Tomcat and Apache are provided with access to a write-protected Blazegraph triplestore. Users can access and visualise the data through several applications, namely, a Blazegraph query interface[22], LodView, Ontodia, Oscar and Beyond Timelines (see Section 4). Fig.7 presents the technical diagram of the KG structure in the VRTI project. The applications and triplestore are started by the developer on the web server. From the VRTI website[23], hyperlinks are provided to users to the KG exploration tools described in Section 4. Upon system launch, the user can interact with the visualization tools by visiting the VRTI website.

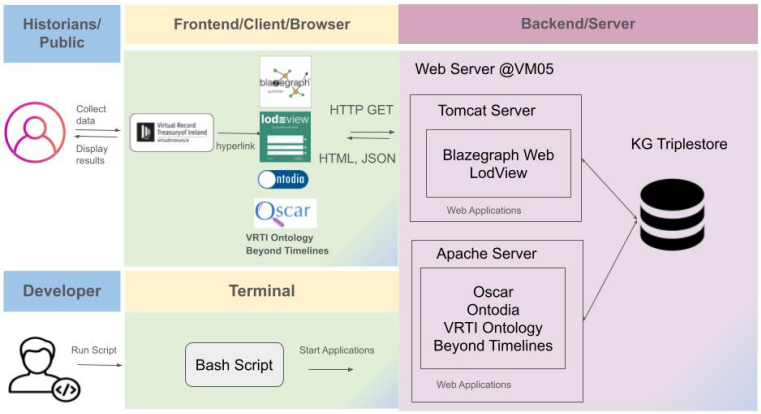

Figure 7: Production Server Diagram.

The second server (Fig 8) is called the development server and this server is used to host tools that are under development. In order to safeguard and maintain the security of the data on the production server, a separate server was deployed and access to programmes running on this server is restricted to the VRTI historians. This server hosts the KG visualisation and editing interface, called VRTI-KG Editor, which is currently being developed. Users who have authenticated with the system can access, add and edit data in the development server triplestore as soon as the developer launches updates to the applications in order to test the tooling.

Furthermore, data can be visualised using tables and maps if it has a geospatial aspect connected to the biographical information. The user can also export data from the system as

[22] https://blazegraph.virtualtreasury.ie/blazegraph/#query
[23] https://virtualtreasury.ie/knowledge-graph

spreadsheets (CSVs) whenever they are needed locally. The data is systematically backed up to the GitHub repository to avoid data loss.

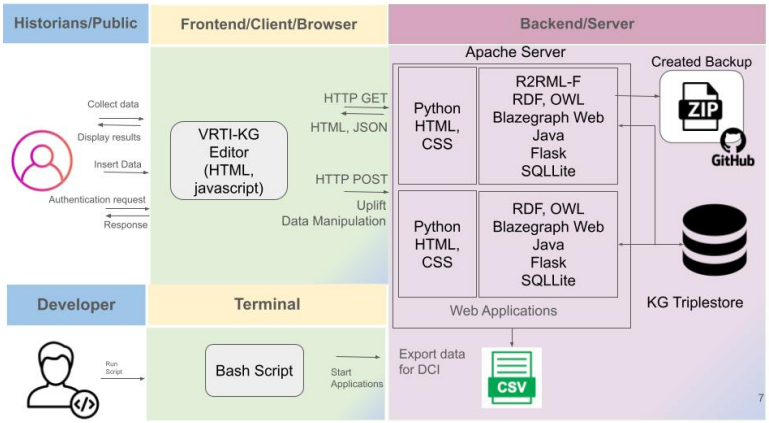

Figure 8: Development Server Diagram.

5.2. Triplestore Selection

Currently, we are using Blazegraph triplestore for our facilities however due to several reasons we decided to migrate from Blazegraph to another triplestore: i) Due to the acquisition of the Blazegraph by Amazon Neptune, Blazegraph is not maintained anymore as an open source tool ii) Blazegraph does not support GeoSPARQL extension fully and visualisation requirements of the project need GeoSPARQL queries to be posed to the endpoint and iii) Blazegraph has been resulting with timeouts for some query types and crashes unexpectedly which causes instability. Taking into account these issues, we decided to migrate to another triplestore. We looked at the literature for the comparison of several features[24] [24] and decided to migrate to Virtuoso triplestore. Even though Fuseki[25] seems to be compliant to more features, it was too slow for our queries which required some OPTIONAL script. As such, Virtuoso Open-Source Edition[26] is currently being trialled.

6. Conclusions and Future Work

This paper presented a KG for prosonography information in VRTI-KG. In this work, we address data visualisation for domain specialists, technical discussions, and biographical data structure by balancing searchability and performance taking into account user experience without overwhelming them.

An interdisciplinary project requires close collaboration from different domains and it is crucial to meet the demands and provide solutions. It was discovered that balancing performance

[24]https://upload.wikimedia.org/wikipedia/commons/e/ea/WDQS_Backend_Alternatives_working_paper.pdf
[25]https://jena.apache.org/documentation/fuseki2/
[26]https://vos.openlinksw.com/owiki/wiki/VOS

and searchability is an important point in selecting the right tools. Using out-of-the-box tools only requires minimal configuration, quick setup and deployment, they are economical both in terms of time and money. They also arrive ready to install. They are not, however, made to meet the particular requirements of the domain experts, hence they are limited in customisation and lack some functionality. While these technologies provided a temporary solution, longer-term implementation is requiring more reliable solutions to be developed.

It is hoped that the VRTI historians will now have a platform to quickly add, modify, and visualise data thanks to the VRTI-KG Editor tool. As a future work, we will continue to develop this editor and conduct a Post-Study System Usability Questionnaire to measure the usability of the interface and apprehend the satisfaction and understanding of the users.

Acknowledgments

Virtual Record Treasury of Ireland (VRTI) is funded by the Government of Ireland, through the Department of Tourism, Culture, Arts, Gaeltacht, Sport and Media, under the Project Ireland 2040 framework. The project is also partially supported by the ADAPT Centre for Digital Content Technology under the SFI Research Centres Programme (Grant 13/RC/2106_P2).

References

[1] P. Crooks, E. Johnston, T. Murtagh, For fear of oblivion, archival fragility and persistence from the middle ages to 1922 - and beyond, Analecta Hibernica 53 (2023) 1–14.

[2] P. Crooks, Z. Reid, C. Wallace, The virtual record treasury of ireland: A century of recovery from the 1922 four courts blaze - and beyond, History Ireland 30 (2022) 38–41.

[3] C. Debruyne, G. Munnelly, L. Kilgallon, D. O'Sullivan, P. Crooks, Creating a knowledge graph for ireland's lost history: knowledge engineering and curation in the beyond 2022 project, ACM Journal on Computing and Cultural Heritage (JOCCH) 15 (2022) 1–25.

[4] M. Doerr, The cidoc crm, an ontological approach to schema heterogeneity, in: Dagstuhl Seminar Proceedings, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2005.

[5] I. Heibi, S. Peroni, D. Shotton, Oscar: a customisable tool for free-text search over sparql endpoints, in: Semantics, Analytics, Visualization: 3rd International Workshop, SAVE-SD 2017, Perth, Australia, April 3, 2017, and 4th International Workshop, SAVE-SD 2018, Lyon, France, April 24, 2018, Revised Selected Papers 3, Springer, 2018, pp. 121–137.

[6] I. Heibi, S. Peroni, D. Shotton, Enabling text search on sparql endpoints through oscar, Data Science 2 (2019) 205–227.

[7] D. Mouromtsev, D. S. Pavlov, Y. Emelyanov, A. V. Morozov, D. S. Razdyakonov, M. Galkin, The simple web-based tool for visualization and sharing of semantic data and ontologies., in: ISWC (Posters & Demos), 2015.

[8] D. Lewis, J. Keeney, D. O'Sullivan, S. Guo, Towards a managed extensible control plane for knowledge-based networking, in: Large Scale Management of Distributed Systems: 17th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, DSOM 2006, Dublin, Ireland, October 23-25, 2006. Proceedings 17, Springer, 2006, pp. 98–111.

[9] B. Yaman, K. McGlinn, L. Hederman, D. O'Sullivan, M. A. Little, Towards a rare disease registry standard: Semantic mapping of common data elements between fairvasc and the european joint programme for rare disease (2022).
[10] P. Leskinen, E. Hyvönen, Linked open data service about historical finnish academic people in 1640–1899, in: Proceedings of the Digital Humanities in the Nordic Countries 5th Conference, CEUR-WS. org, 2020.
[11] P. Leskinen, H. Rantala, E. Hyvönen, Analyzing the lives of finnish academic people 1640–1899 in nordic and baltic countries: Academysampo data service and portal, in: Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022), CEUR-WS. org, 2022.
[12] E. Hyvönen, P. Leskinen, M. Tamper, H. Rantala, E. Ikkala, J. Tuominen, K. Keravuori, Biographysampo–publishing and enriching biographies on the semantic web for digital humanities research, in: European Semantic Web Conference, Springer, 2019, pp. 574–589.
[13] H. Rantala, E. Hyvönen, J. Tuominen, Finding and explaining relations in a biographical knowledge graph based on life events: Case biographysampo, in: CEUR Workshop Proceedings, volume 3443, RWTH Aachen University, 2023.
[14] E. Ikkala, M. Koho, E. Heino, P. Leskinen, E. Hyvönen, T. Ahoranta, et al., Prosopographical views to finnish ww2 casualties through cemeteries and linked open data., in: WHiSe@ ISWC, 2017, pp. 45–56.
[15] M. Koho, H. Rantala, E. Hyvönen, Digital humanities and military history: Analyzing casualties of the warsampo knowledge graph, in: Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022), CEUR-WS. org, 2022.
[16] M. Pasin, J. Bradley, Factoid-based prosopography and computer ontologies: towards an integrated approach, Digital Scholarship in the Humanities 30 (2015) 86–97.
[17] M. Hammond, From digital prosopography to social network analysis, Medieval People: Social Bonds, Kinship, and Networks 36 (2021) 235–262.
[18] B. Krabina, Building a knowledge graph for the history of vienna with semantic mediawiki, Journal of Web Semantics 76 (2023) 100771.
[19] P. Connolly, Irish exchequer payments 1270-1446, Irish Manuscripts Commission, 1998.
[20] M. Doerr, The cidoc conceptual reference module: an ontological approach to semantic interoperability of metadata, AI magazine 24 (2003) 75–75.
[21] D. Garijo, Widoco: a wizard for documenting ontologies, in: The Semantic Web–ISWC 2017: 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings, Part II 16, Springer, 2017, pp. 94–102.
[22] B. Catania, G. Guerrini, B. Yaman, Exploiting context and quality for linked data source selection, in: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, 2019, pp. 2251–2258.
[23] H. J. Pandit, D. O'Sullivan, D. Lewis, Using ontology design patterns to define shacl shapes., in: WOP@ ISWC, 2018, pp. 67–71.
[24] M. Jovanovik, T. Homburg, M. Spasić, A geosparql compliance benchmark, ISPRS International Journal of Geo-Information 10 (2021) 487.