The Knowledge Graph Explorer for the Virtual Record Treasury of Ireland

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

[1] *ADAPT Centre for Digital Content, Trinity College Dublin, Ireland*
[2] *Department of History, Trinity College Dublin, Ireland*

Abstract

The Irish civil war in 1922 resulted in the destruction of Ireland's central archive containing documents dating back seven centuries. A century later, the Virtual Record Treasury of Ireland (VRTI) created a Knowledge Graph (KG) of the recovered historical documents. However, accessing the information in the KG requires the definition of complex and time-consuming SPARQL queries. It was decided to create an application named the VRTI-KG explorer to facilitate searching of the VRTI-KG. The explorer includes customised views which provide natural language descriptions and visualisations of the underlying data in a format to allow non-technical users to easily interpret the information. In addition, the explorer is configurable to allow straightforward synchronisation of changes in the data model into the visualisations on the interface and queries involved in creating them. A user evaluation was conducted with 20 participants to measure the level of satisfaction, understanding and efficiency provided by the explorer.

Keywords

KG Search, VRTI, User Interface, User Testing

1. Introduction

The **Virtual Record Treasury of Ireland (VRTI)**[1] [1–3] is state-funded programme hosted at Trinity College Dublin. The VRTI is a digital recreation of archival records damaged during a fire in the 1922 Irish Civil war. The project involves a high level of interdisciplinary research with historians collecting information and computer scientists digitising the information. It was decided to create a **Knowledge Graph (KG)** to integrate heterogenous data collected by historians and facilitate information discovery. The project has adopted W3C standards to support standard internet technologies. The VRTI-KG[2] is represented using an extended version

VOILA 2024: The 9th International Workshop on the Visualization and Interaction for Ontologies, Linked Data and Knowledge Graphs co-located with the 23rd International Semantic Web Conference (ISWC 2024), Baltimore, USA, November 11-15, 2024.
* Corresponding author.
alex.randles@adaptcentre.ie (A. Randles); lucy.mckenna@adaptcentre.ie (L. McKenna); kilgall@tcd.ie (L. Kilgallon); beyza.yaman@adaptcentre.ie (B. Yaman); pcrooks@tcd.ie (P. Crooks); declan.osullivan@adaptcentre.ie (D. O'Sullivan)
0000-0001-6231-3801 (A. Randles); 0000-0002-6035-7656 (L. McKenna); 0000-0002-3075-8571 (L. Kilgallon); 0000-0003-2130-0312 (B. Yaman); 0000-0001-6782-044X (P. Crooks); 0000-0003-1090-3548 (D. O'Sullivan)
© 2024 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

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

CEUR-WS.org/Vol-3773/paper4.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

of the CIDOC-CRM [4] ontology, which was designed to model information in the cultural heritage domain. The extended version used by the VRTI-KG contains concepts to represent notable people, places and their interconnections from Irish history.

Navigating the VRTI-KG requires the creation of SPARQL [5] queries which facilitate retrieval of information represented in RDF. However, these queries are often complex and time-consuming, requiring a high-level of technical expertise. In addition, the RDF data which is retrieved is often difficult for non-technical users to interpret. In this paper we propose the **VRTI-KG Explorer**, which is a bespoke web interface to facilitate searching of the resources in the VRTI-KG. The explorer includes customised views which provide natural language descriptions and visualisations of the underlying data in a format to allow non-technical users to interpret the information. In addition, the explorer is configurable which allows changes within the VRTI-KG to be easily propagated into the views on the interface without extensive code changes. The configurability of the explorer will allow it to be applied to other KGs in the future.

This paper describes the design of the explorer and the user evaluation. The evaluation of the explorer provided a method to measure the level of satisfaction, understanding and efficiency of the interaction. The evaluation used standardized metrics which allows the results to be easily conveyed and compared with existing acceptable thresholds. The results were used to validate and refine the defined user requirements. Lessons learnt from the evaluation provide useful insights for researchers validating similar approaches.

This paper is structured as follows: Section 2 discusses the design and implementation of the VRTI-KG explorer. Section 3 describes a user evaluation which was completed on the explorer. Section 4 presents related work on interfaces designed to facilitate searching of KGs. Section 5 outlines future work and concludes the paper.

2. VRTI-KG Explorer

This section discusses the design[3] of the VRTI-KG Explorer (https://vrti-graph.adaptcentre.ie/) that includes multi-faceted search, which was inspired by the design of Sampo-UI [6]. However, we propose advancements to the Sampo-UI design by introducing novel natural language and visualisation elements, as well as data-driven configurability of the user interface. The initial design requirements were conceived from a focus group which consisted of the computer science and digital humanities researchers in the VRTI. Thereafter, an initial prototype was created, and an iterative process of refinement was completed where a demonstration of each version was conducted followed by feedback interviews. In addition, the early versions of the explorer were made accessible online to the team along with a link to a feedback spreadsheet, which enabled the explorer to be tested in diverse environments.

2.1. Implementation of VRTI-KG Explorer

The explorer was primarily implemented using Python libraries[4]. Flask is a customizable web framework which was used to create the web application. Folium is used to create maps for the geospatial data. SPARQLWrapper is used execute SPARQL [5] queries on the VRTI-KG. The

[3] https://github.com/alex-randles/Explorer-Evaluation/blob/main/component-diagram.png
[4] https://github.com/alex-randles/Explorer-Evaluation/blob/main/libraries.pdf

Open-AI library is used to facilitate the ChatGPT search. In addition, some Javascript was used for handling the dynamic content on the HTML pages.

2.2. Configuration of Explorer

The explorer was designed to be configurable to enable straightforward synchronisation with changes in the VRTI-KG without the need for extensive code changes. In addition, it is hoped that the configurability of the application allows it to be applied to other projects involving KGs (e.g. searobend.ie). The implementation contains several configuration files which allow front-end and back-end components, such as the SPARQL queries, grouping of query results, technical resources, suggested search terms and sample resources shown to be easily updated. The configuration files are represented in JSON format which allows the views displayed on the interface to be changed without needing to alter any code. **Listing 1** presents an extract of a configuration files used by the application to fetch and display the search results for the people in the KG.

```
1 "People": {
2         "query": """
3            PREFIX crm: <http://erlangen-crm.org/current/>
4            PREFIX vrti:  <https://ont.virtualtreasury.ie/ontology#>
5            PREFIX foaf: <http://xmlns.com/foaf/0.1/>
6            PREFIX owl: <http://www.w3.org/2002/07/owl#>
7            PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8            PREFIX geo: <http://www.opengis.net/ont/geosparql#>
9
10           SELECT DISTINCT ?Person ?Name ?DeathDate ?BirthDate ?Image
11           WHERE { .......  }
12           ORDER BY ?Name
13     ""
14         "filter_variables": ["Name", "Gender", "BirthPlace", "DeathPlace", "Interest", "Occupation", "Office"],
15         "viewable_variables": ["Birth", "Death"],
16         "dates": {
17             "lower_date": "BirthDate",
18             "upper_date": "DeathDate",
19         },
20         "redirect_variable": "Person",
21         "search_variable": "Name",
22         "tooltip_text": "People in the Dictionary of Irish Biography (DIB)",
23         "icon": "bi bi-people-fill",
24         "default_image": "https://upload.wikimedia.org/wikipedia/commons/thumb/b/bc/Unknown_person.jpg/694px-Unknown_person.jpg",
25
26 }
```

Listing 1: Extract of Configuration file for People in the VRTI-KG

The configuration file shown consists of the query (line #2) which is executed to retrieve people from the VRTI-KG. The retrieved results are grouped in a specific tab (#1) on the search results page. A redirect variable (#20) defines where the user is redirected when they click a particular result. Tooltip text (#22) and an icon ("bi bi-people-fill")[5] is used to distinguish the people result sets from others. The advanced search filters (#14) allow users to apply further searchers to their initial query. Each filter variable is mapped into a FILTER condition in a SPARQL [5] query once selected. A query variable (#21) is used to match the initial search results once a term has been entered into the main search bar. Then, certain variables (#15) are shown in the search results to allow users to distinguish between the retrieved people. A default image (#24) is shown for people with no associated image in the KG. Styling of the interface can

[5] https://icons.getbootstrap.com/icons/people/

be configured in other files by using classes available in the Bootstrap library[6]. Bootstrap is a front-end toolkit which includes configurable components used to design interfaces. For instance, the user wants to change the colour of the search buttons from blue to green. The class name for green ("success") is retrieved from the Bootstrap documentation and replaces the class for blue ("primary"). The interface will automatically change all button colours when the configuration is updated.

2.3. Main Search Functionality

The main search functionality allows the users to search using the queries defined in a configuration file. **Figure 1** presents a screenshot of sample search results for the search term "Thomas".

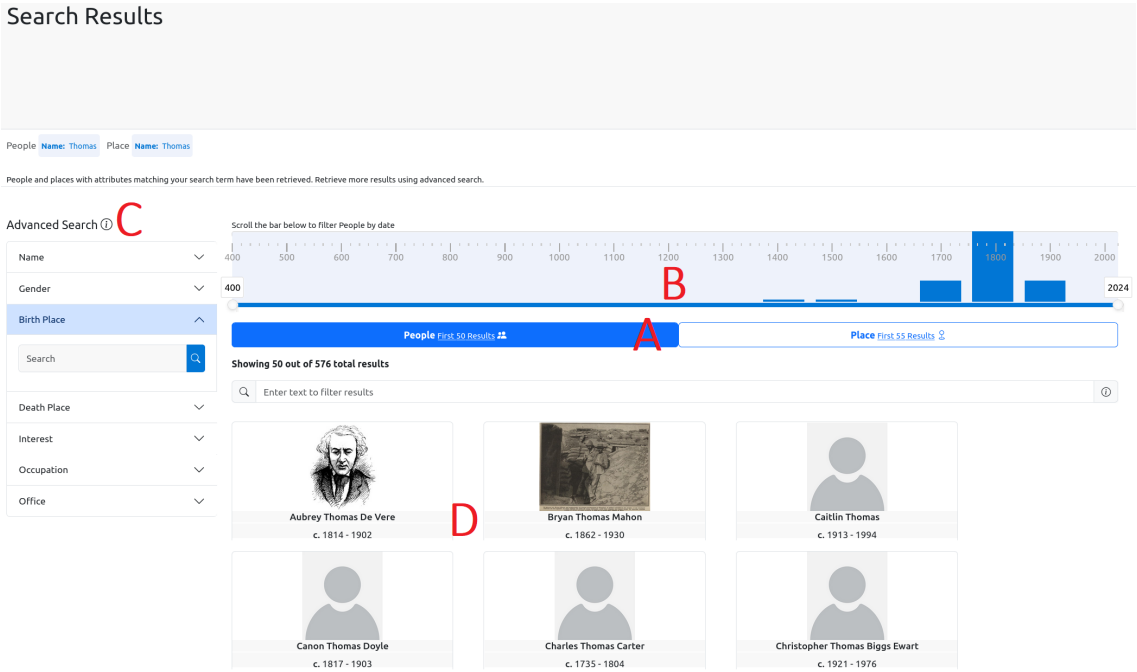


Figure 1: Screenshot of search results for search term "Thomas"

Suggested terms which exist in the VRTI-KG are provided to the users once they start to type their search term. In addition, a description of each suggested term is included. The results page contains tabs (A – Figure 1) for each result group defined in the configuration file allowing them to toggle between the result sets. The timeline (B) provides a visual representation of the defined date variables. The slider allows users to filter the results based on start and end dates of when people existed. The advanced search (C) options allow another query to be executed using the defined filter variables. The entered search term and respective variables are mapped into a SPARQL query template which allows them to discover relevant results. The initial queries are limited to 50 results to improve efficiency. However, a button at the bottom of the page allows

[6] https://getbootstrap.com/docs/5.3/getting-started/introduction/

a user to retrieve more results from the existing query until no more results are available to display. The functionality is accomplished by incrementing the offset and limit conditions in the SPARQL query while comparing them to the total number of results which is measured using a count query. The viewable variables from the query results are displayed for each result (D) and once clicked are redirected to the defined redirect variable. For this configuration, users are redirected to a summary page of the person[7], which includes a natural language summary of the associated data in the VRTI-KG along with visualizations. Similar pages are available for places[8]. The natural language summaries are created by configurable templates which insert specific query variables into sentences. Early experimentation [7] was conducted with LLMs to investigate how prompts can be used to generate syntactically and semantically correct SPARQL [5] queries. The lessons learnt provided useful insights for creating queries from natural language questions designed to retrieve answers from the VRTI-KG.

2.4. Geospatial Search

The explorer provides several map views[9] of the geospatial data in the VRTI-KG, which is represented in a nested hierarchy. The hierarchy consists of Townland (lowest), Parish, Barony and County (highest). **Figure 2** presents a screenshot of the page where users can toggle between maps representing places in each level of the hierarchy.

The map views can be toggled by clicking the respective tab name (A – Figure 2). Each map is generated by querying the VRTI-KG with a query designed to retrieve the respective place and associated information. The result set is iterated, and each coordinate is plotted onto the map using a marker (B), along with the associated information as hover text. The markers are clickable which redirects them to the summary page of the selected place. A specific place on the map can be found by using the search bar (C). In addition, the users can complete a separate search[10] for nearby places by selecting a location on a map, which then will create and execute a SPARQL [5] inserting the selected coordinates.

[7] Sample summary page for person at https://vrti-graph.adaptcentre.ie/entity-card/person/Wellesley_Arthur_c19_dib_a8961
[8] Sample summary page for place at https://vrti-graph.adaptcentre.ie/entity-card/geo/modern-townland/C_22_B_10_P_12_160116_BALLYBIN-ED-Ratoath
[9] https://vrti-graph.adaptcentre.ie/place-homepage
[10] https://vrti-graph.adaptcentre.ie/place-homepage#county-search-form

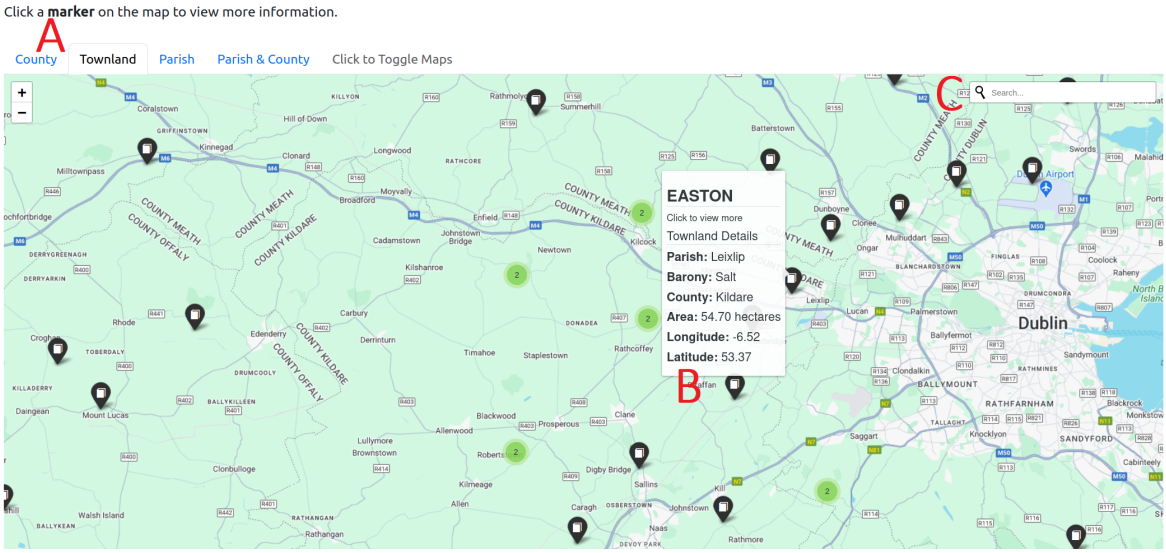


Figure 2: Screenshot of map visualisation of geospatial data

2.5. Natural Language Search

A natural language search[11] [8] has been integrated into the explorer to facilitate searching through natural language questions and answers. We experimented with a natural language querying tool by Ontotext[12] before deciding to create a bespoke solution. The tool was setup with the VRTI-KG and provided the ontology. However, it struggled to create syntactically correct queries for most of the test cases, which is likely because of the complex CIDOC-CRM [4] based structure of the VRTI ontology. Thus, we decided to use ChatGPT [9] to facilitate the extraction of entities from the question, which are then inserted into SPARQL [5] query templates defined in a configuration file. It was decided to use ChatGPT as it performed best compared to other LLMs in early experimentation. A natural language response is formed from the query results using ChatGPT. For instance, a user can ask "Who is Michael Collins", "Where and when was Michael Collins born?" or "Was Michael Collins in the Army?". **Figure 3** presents a screenshot of the results for the question "Tell me about Michael Collins".

The entities are retrieved from the initial question entered in the search bar (A – Figure 3) in a JSON dictionary format, which allows them to be mapped into templates for FITLER conditions within SPARQL queries. For instance, the user asks "Tell me about Michael Collins". The extracted entity ("Michael Collins") is inserted into a FILTER condition ("FILTER (CONTAINS(?Name, 'Michael Collins'))") which targets the variable associated with the names of people (?Name). Then, the query is executed on the VRTI-KG and the results output into a dictionary format. Finally, the initial question and query results and input into a prompt template which asks ChatGPT to form a natural language answer (B) from only the information in the provided dictionary.

[11] https://vrti-graph.adaptcentre.ie/gpt-search
[12] https://www.ontotext.com/blog/natural-language-querying-of-graphdb-in-langchain/

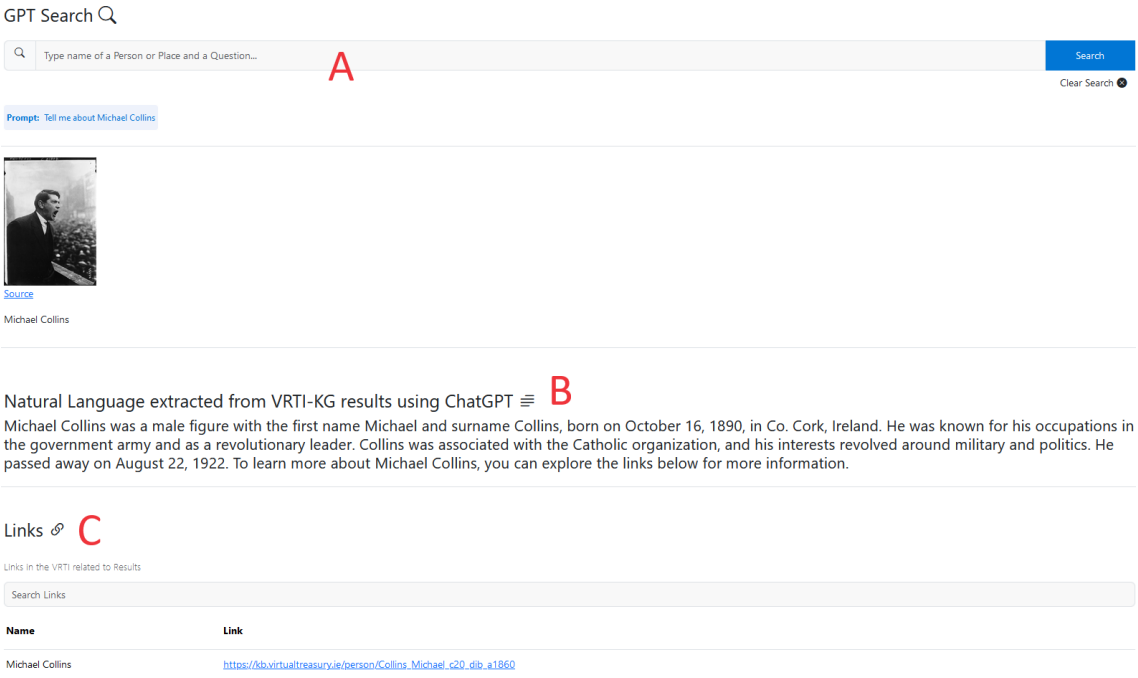
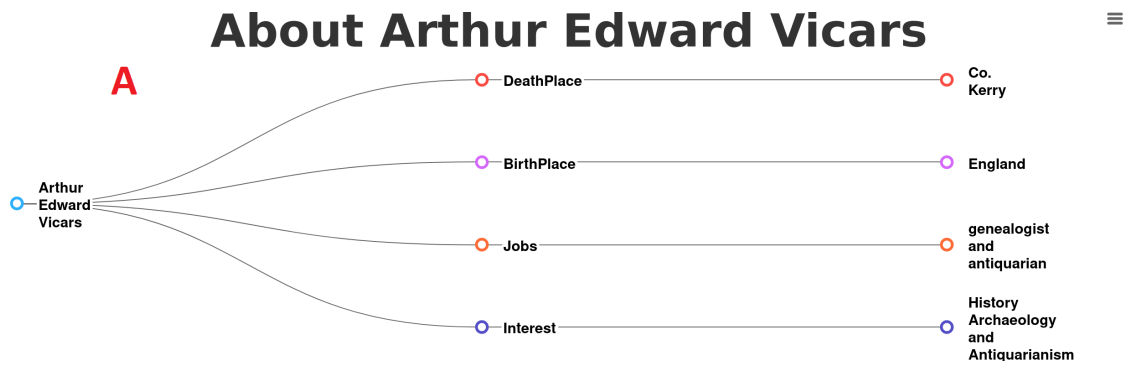


Figure 3: Screenshot of sample natural language search result

In this way, the knowledge presented is only that which exists in the KG, and essentially we are using ChatGPT only for its natural language processing capabilities, rather than its generative text capabilities. In addition, links (C) to the retrieved resources are presented to allow further exploration using the explorer.

2.6. Additional Visualisations

Figure 4 presents a screenshot of the tree-graph and hierarchical-graph visualisations available on the summary pages for a person[13].



[13] https://vrti-graph.adaptcentre.ie/entity-card/person/Vicars_Arthur-Edward_c20_dib_a8813

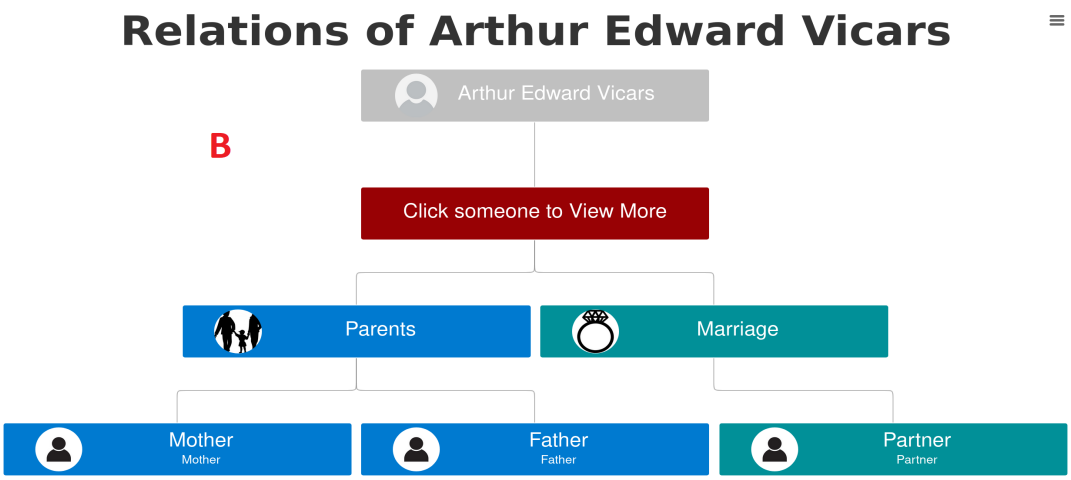


Figure 4: Tree-graph of person attributes (A) and hierarchical graph of relations (B)

Similar visualisations are configured for offices, organisations, interests, among others where information for key attributes is presented. The visualisations are created by executing a SPARQL [5] query on the VRTI-KG, which is specifically designed to retrieve all of related information. The result set is iterated to retrieve variables which contain key attributes involved in creating the visualisations. Then, the retrieved information is fed into methods that use a Javascript library named Highcharts[14] to create the resulting visualisations. The visualisations include clickable elements to explore related links within the explorer.

3. User Testing of VRTI-KG Explorer

This section discusses the user evaluation which was completed on the explorer.

3.1. Methodology

An evaluation was completed on the explorer to validate the design with respective end users. The participants were asked to complete several tasks using the explorer followed by questionnaires. It was decided to measure understanding, satisfaction and efficiency in order to identify if participants could effectively navigate the explorer while understanding the presented information. A standardized metric was used to measure satisfaction. Understanding was measured using a bespoke questionnaire which include questions relevant to information shown on the explorer.

3.2. Metrics

The following metrics were used to measure the perceived **understanding**, **satisfaction** and **efficiency.**

Post Study Usability Questionnaire (PSSUQ). The PSSUQ [10] was used to measure the user's **satisfaction** of the explorer. The PSSUQ includes a Likert scale in the range of 1-7, to

[14] https://www.highcharts.com/

rate various aspects of the application, such as the quality of the interface and information. The PSSUQ includes four subscales, which are system usefulness, information quality, interface quality and overall. A score of 1 means the user was highly satisfied, 3 means neutral and 7 means highly dissatisfied.

Understanding Questionnaire. The **understanding** questionnaire[15] consisted of 11 questions which asked the participants to provide information presented on the interface. The questions were designed to capture information which they would interact with during the tasks. The information includes details of notable people and places, along with geospatial data on the maps.

Time for completion. The total time to complete the experiment was used to measure **efficiency**. The participants were asked to provide self-reported times for completion given the experiment ran asynchronously.

3.3. Experiment Setup

The experiment setup includes the participants and tasks which they were asked to complete.

Participants. The experiment consisted of 20 participants with unspecified background knowledge. The participants were not required to have any computer science, semantic web or historical knowledge. The participants were recruited by sharing the experiment details in public email threads and at recent conferences (e.g. at ESWC workshops).

Task sheet. A focus group was conducted with the researchers in the VRTI to determine which tasks should be completed during the experiment that would allow the users an initial experience of the interface and diversity of content of the KG. The tasks were designed to mimic expected user interaction. The resulting task sheet[16] consisted of 10 tasks. Tasks 1-3 asked the participants to examine notable people in different time periods. Tasks 4-7 asked them to examine the maps and place summary pages. Tasks 7-10 asked them to complete a search and navigate from the results to specific summary pages and examine the information shown.

3.4. Experiment Execution

The experiment received approval from the TCD Research Ethics Committee before commencing. This section discusses how the experiment was conducted.

Completion of Experiment. The experiment was completed asynchronously by the participants where each of them accessed a web link. The link included the informed consent, background information on the explorer, task sheet, understanding questionnaire and PSSUQ.

Assistance. No assistance was provided to the participants during the experiment. In addition, the participants were not provided with any documentation or video describing the explorer prior to the experiment, as the interface itself provides tool tips that are intended the user to understand the interface, and the intention was to evaluate experiences of the users coming to the interface “cold”.

---

[15] https://github.com/alex-randles/Explorer-Evaluation/blob/main/understanding-questionnaire.pdf
[16] https://github.com/alex-randles/Explorer-Evaluation/blob/main/task-sheet.pdf

3.5. Experiment Results

This section discusses the results of the PSSUQ, understanding questionnaire and timings recorded.

3.5.1. Results of PSSUQ

The mean scores of each PSSUQ question and subscale (see Section 3.2) have been plotted on the boxplot presented in **Figure 5**. In addition, the scores were compared with acceptable research thresholds [10].

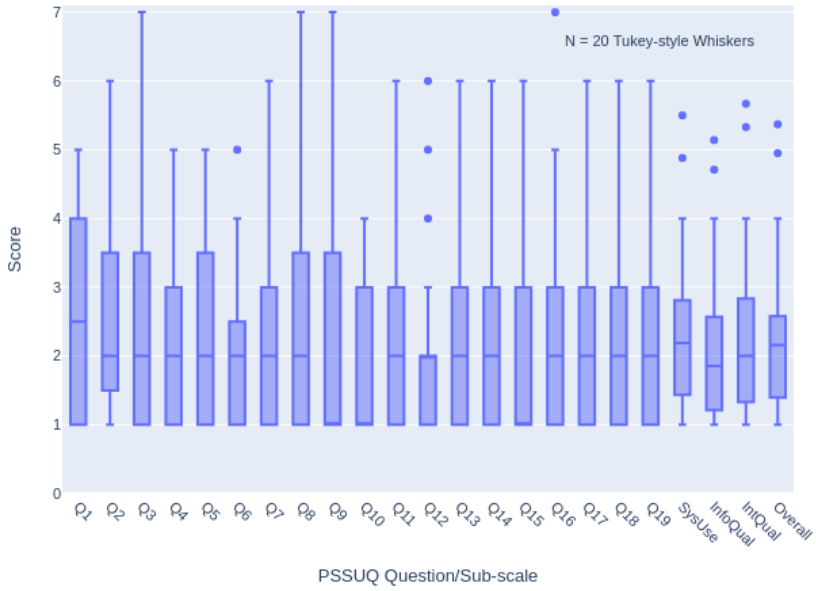


Figure 5: Scores of the PSSUQ questions and sub-scales

The scores indicate that the subscales are between 5.21% and 39.08% better than their respective acceptable research threshold. Information quality (InfoQual) scored best with a score 39.08% better than its threshold. Comments which supported sufficient information quality included “I believe I could quickly get used to navigating this system to find the information I need.”, “I was able to follow all of the instructions and find the information.” and “The information seems of high quality and helped me achieve the task”. However, it was noted that some headings and sections could be reorganised to help find information easier. Related comments include “The headings I was looking for weren't always immediately obvious” and “I had to look closely for 'tabs', 'headers' and 'labels'”. System usefulness (SysUse) and Overall of the PSSUQ scored similar with 15.17% and 22.05, respectively better than their threshold. Related comments include “I found it intuitive and easy to use.”, “The system was quick and responsive” and “Really looks good and works very well”. Interface quality (IntQual) scored worst with 5.21% better than its respective threshold. Lack of diverse colours on the interface could have been the cause of the poor score as 2 participants stated “Using more colours would help the interface look more professional.” and “The distinction between 'Events' and 'Links' on the Places–Parish page was unclear.”. However, others stated they liked to interface with comments such as “Interface is nice”, “Stylish and modern interface” and “The interface has a

pleasant, modern and professional feel to its design". The interface was configured with neutral colours, which could be changed in future to include more colours to clearly distinguish different areas. The overall results indicate that sufficient satisfaction was observed with the interaction as all sub-scales scored better than their respective research thresholds.

3.5.2. Results of Understanding Questionnaire

11 questions were included in the understanding questionnaire and 20 responses were received for each question. The number of correct answers for each question are presented in **Figure 6** .

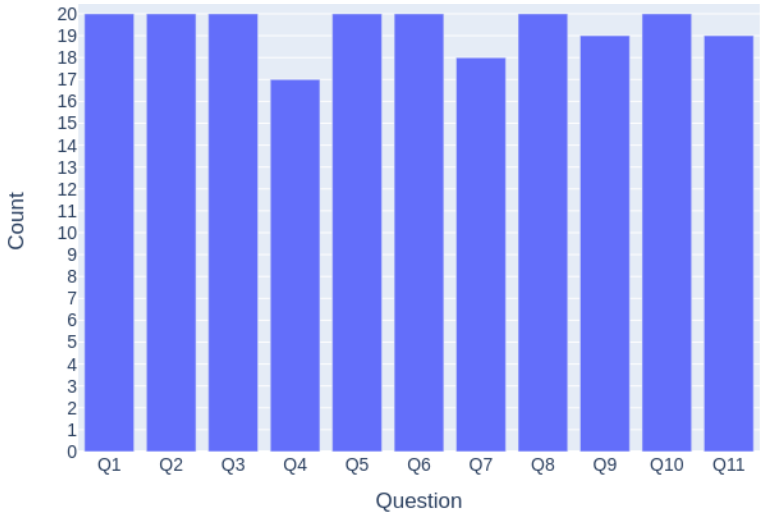


Figure 6: Number of correct answers in the understanding questionnaire

9 out of 11 questions scored at least 19 out of 20 correct (95%), which indicates a sufficient level of understanding for most information. The scores indicate that the participants could understand the time period categorisation of people (Q1-2), the place information presented in text and on maps (Q2-6), the number of search results (Q7-8) and the information presented on the summary pages of people and places. However, some questions scored worse than others. The worst scoring question (Q4) had a score of 17 out of 20 (85%) correct and related to the information provided by the hover text on the map markers. 1 of the incorrect answers were as a result of the participants using a different map and 2 others misunderstood the question with one of them stating "I dont know the
difference between district, area and elevation.", while the question asked them to select the only attribute shown in the hover text. Another lower scoring question (Q7) had a score of 18 out 20 (90%) correct, which asked them to provide the number of people results returned. The incorrect answers provided the total number of available results, rather than the number of initial results. Thus, there are indications that refinement of the related text on the interface is worth considering.

3.5.3. Completion Time Results

Figure 7 presents a boxplot of the self-reported completion times. The minimum time was 5 minutes, maximum was 26 minutes and the mean was 14.3 minutes. A standard deviation of 6.8 minutes indicates that the times were spread around the mean and efficiency was not equal for all participants. Spearman's correlation [11] is a test designed to measure the strength of a relationship between variables. The test was applied to identify if there was a correlation between timing, satisfaction and understanding. It was decided to use this test as it is less sensitive to outliers. A confidence level of 0.05 was applied to indicate a statically signification score.

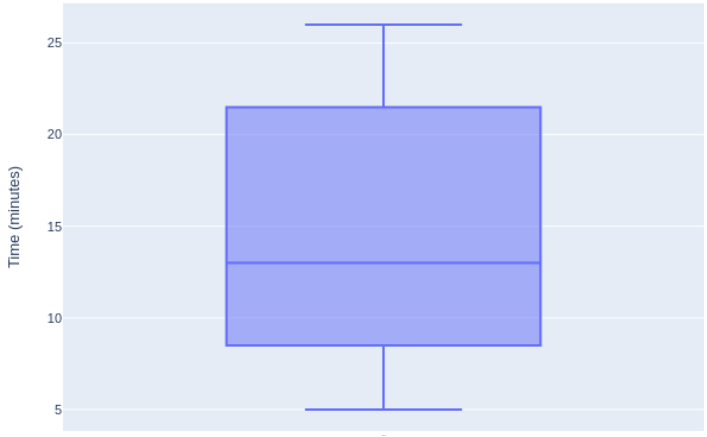


Figure 7: Completion times of participants

The test showed a statistically significant relationship (-0.128) between time and understanding, which indicates that as time decreases, understanding increases meaning more efficient participants had better understanding. Similarly, participants who scored the PSSUQ better also had better efficiency as the test showed a statistically significant relationship (-0.064) between them. These results indicate that more efficient participants also had better understanding and satisfaction.

3.6. Overall analysis of result

The results indicate sufficient understanding, efficiency and satisfaction for first time users of the explorer with positive quantitative and qualitative data observed. The understanding questionnaire scored high with a mean of 19 out of 20 (95%) correct answers. The qualitative data supported this finding with comments such as "Very easy for a first attempt. With an hour or two of use, I would imagine it would become easily intuitive", "I found it intuitive and easy to use." and "Really well designed and presented, with it also being fast". Each PSSUQ sub-scales measured better than its respective threshold between 5.21% and 39.08%, which indicates sufficient overall satisfaction. Comments which supported this finding were "The system is very useful for searching places and people.", "It was a nice system." and "Well done with the map exploration, impressive". The mean completion time for the experiment was 14.3 minutes which indicates a sufficient level of efficiency as the participants were asked to complete 10 tasks which involved them navigating to multiple pages and examining various pieces of information.

Comments which supported this finding include “The layout enabled me to complete the tasks quickly.”, “I felt I did so quite quickly.” and “yes, it was quick to check each page”. These results provide both quantitative and qualitative measurements which indicate that the explorer is a useful tool to facilitate exploration of KGs. It is hoped that the explorer can build upon the existing knowledge and advancements of technology to provide diverse visualisations and conversational search of data in KGs. The proposed approach provides a base model that has been shown to allow respective end users to navigate and identify relevant information in a timely manner. The flexibility of the approach provided by the configurability allows the exploration methods to be easily changed for the intended audience.

4. Related Work

This section discusses approaches to facilitate searching of data in KGs.

A recent survey [12] was completed which investigates tools designed to support searching of KGs. The survey compared 28 approaches which facilitate searching of KGs. The interfaces are compared based on interaction paradigm, information being displayed, and strategies used to improve understanding of information. The survey concluded that many of these approaches still require some level of technical expertise to be used effectively, which some domain experts may lack.

LodView[17] is a tool which facilitates browsing of relationships of resources in a KG in HTML format. The approach is configurable by RDF files, which allows the tool to be applied to different KGs. Dereferencing a selected URI presents a tabular listing of direct and indirect links. These links can be selected and further explored using the tool. However, the presentation of information is restricted to RDF terms, which means users need to understand the underlying schema. The configuration of the tool inspired the explorer, however, it was decided to use JSON rather than RDF to decrease workload.

Sampo-UI [6] is a tool which provides developers with a modular set of customisable and reusable components to interact with a KG. The tool provides multi-faceted search functionalities which can be targeted at different pre-defined perspectives. Each perspective provides a different entry point for navigating through the resources in a KG. . The tool includes configuration files which allows it to be customised for different KGs. The multi-faceted search functionality of the tool provides inspiration for a similar approach by the explorer. However, it does not include support for mapping polygons which was required to visualise the boundaries of places in the VRTI-KG. In addition, it does not support natural language question and answering.

Ontodia[18] is a tool which contains an interface with a graph-based visualisation of the resources in a KG. It allows users to search through a listing of the classes in the data and find related resources. A visualisation of the connections between the selected resources is shown to provide an understanding of how the data is connected. Ontodia provided inspiration for some of the visualisations available on the explorer.

OSCAR (OpenCitations RDF Search Application) [13] is a tool designed to facilitate searching of the RDF data by querying a SPARQL [5] endpoint. The tool is configured by JSON

[17] https://github.com/LodLive/LodView
[18] https://github.com/metaphacts/ontodia

files, which define how the queries are created from the users input and displayed on the resulting table. The tool is limited to free text search and provides only a tabular view of the results. The configuration of the tool using JSON files provided inspiration for the configurability of the explorer.

DBpedia [14] is a project to transform Wikipedia data into RDF. The project includes a tool to facilitate navigation of resources in the RDF representation. The tool provides a listing of related resources in a tabular format to allow further internal navigation. The traversing of the resources within the application inspired similar interaction in the explorer.

To the best of our knowledge, none of these approaches have published results of formal user evaluations. Thus, it is difficult to determine how well they satisfy user requirements. The use of standardized metrics provides a method to convey and compare results with acceptable thresholds to validate defined requirements. In addition, none of these approaches provide support to search using natural language questions and answers, which is hoped to facilitate straightforward retrieval of relevant information.

5. Future Work and Conclusion

Future work involves refinement of the explorer design based on the results of the user testing. The user requirements will be reviewed to determine which ones were validated by the evaluation and revised for the next iteration of development. Once it has been refined based on the evaluation results, we plan to configure the application for another KG and then publish it as an open-source resource for other researchers to use. In addition, we intend to introduce new visualisations into the interface, including charts created from the proposed natural language questioning.

The VRTI-KG Explorer proposed in this paper is hoped to facilitate easy interaction between the data in the VRTI-KG and diverse users. The previous method of retrieving information in the KG required a level of technical expertise which most of the potential users would not possess. It is hoped the explore can improve the uptake of information by interested parties. The configurability of the explorer provides flexibility, allowing it to be easily tailored to specific needs without extensive code changes. The results of the evaluation demonstrated sufficient levels of satisfaction, understanding and efficiency. However, areas for improvement were identified. It is hoped that sharing the open-source implementation will allow the approach to be customised for other KGs. In addition, it is hoped that the evaluation methodology can provide useful insights to guide the validation of similar approaches. Finally, the lessons learnt from completing the evaluation are hoped to prevent similar pitfalls occurring in other applications.

Acknowledgements

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] B. Yaman, L. McKenna, A. Randles, L. Kilgallon, P. Crooks, D. O'Sullivan, Digital Prosopography Information in Virtual Record Treasury of Ireland Knowledge Graph, in: Proceedings of the 1st International Workshop of Semantic Digital Humanities (SemDH) Co-Located with the 21st Extended Semantic Web Conference, 2024. https://ceur-ws.org/Vol-3724/paper2.pdf.

[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, Hist Irel 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, J. Comput. Cult. Herit. 15 (2022). https://doi.org/10.1145/3474829.

[4] G. Goerz, M. Oischinger, B. Schiemann, An implementation of the CIDOC conceptual reference model (4.2. 4) in OWL-DL, in: Proceedings of the 2008 Annual Conference of CIDOC-The Digital Curation of Cultural Heritage, 2008.

[5] S. Harris, A. Seaborne, E. Prud'hommeaux, SPARQL 1.1 Query Language, World Wide Web Consortium (W3C) Recommendation 21 (2013) 778. https://www.w3.org/TR/sparql11-query/ (accessed April 1, 2023).

[6] E. Ikkala, E. Hyvönen, H. Rantala, M. Koho, Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces, Semant Web 13 (2022) 69–84.

[7] A. Randles, D. O'Sullivan, R2 [RML]-ChatGPT Framework, in: Fifth International Workshop on Knowledge Graph Construction@ ESWC2024, 2024: pp. 35–50.

[8] A. Randles, D. O'Sullivan, P. Crooks, L. Kilgallon, L. McKenna, B. Yaman, Facilitating Search of the Virtual Record Treasury of Ireland Knowledge Graph using ChatGPT, in: Proceedings of Poster and Demo Track and Workshop Track of the 20th International Conference on Semantic Systems (SEMANTiCS), 2024.

[9] T. Wu, S. He, J. Liu, S. Sun, K. Liu, Q.-L. Han, Y. Tang, A brief overview of ChatGPT: The history, status quo and potential future development, IEEE/CAA Journal of Automatica Sinica 10 (2023) 1122–1136.

[10] J.R. Lewis, Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies, Int J Hum Comput Interact 14 (2002) 463–488. https://doi.org/10.1080/10447318.2002.9669130.

[11] C. Spearman, The proof and measurement of association between two things., Am J Psychol 15 (1904) 72–101. https://doi.org/10.2307/1412159.

[12] E. Bernasconi, C. Miguel, M. Mecella, Linked Data interfaces: a survey, in: 19th Conference on Information and Research Science Connecting to Digital and Library Science, 2023: pp. 1–16.

[13] 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, 2018: pp. 121–137.

[14] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Ives, Dbpedia: A nucleus for a web of open data, in: International Semantic Web Conference, 2007: pp. 722–735.