Usable Mobile Geographical Linked Data Visualisation

Balázs Pete

A dissertation submitted to the University of Dublin,
in partial fulfilment of the requirements for the degree of
Master of Science in Computer Science

2014

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed: ______________________

Balázs Pete
28 August 2014

Permission to Lend and/or Copy

I agree that Trinity College Library may lend or copy this thesis upon request.

Signed: ______________________

Balázs Pete
28 August 2014

Acknowledgments

I would like to thank my supervisor, Rob Brennan, for his guidance and help in completing this dissertation. I would also like to thank the participants of the experiments, whom were essential to achieve the conclusions of this work. And finally, I would like to thank my friends and family for their moral support.

Abstract

In this study, the usability of the visualisation of geographical Linked Data on a mobile device was investigated. A mobile application was developed through an iterative process where the development of later prototype iterations were guided by analysing the results of usability studies. After a paper-based study and three prototype iterations, this application was compared to a desktop-based Linked Data geographical visualisation. This work is important because the amount of available spatial Linked Data is growing every day and the need to visualise this information to help users understand it increases correspondingly. There has been less effort to date in mobile Linked Data research, and not many mobile applications focus on visualising geographical data.

The experiments showed that the usability of the mobile app was affected by the addition visual feedback, the increased number of features, the introduction of visual cues, and whether users preferred the mobile app over the desktop one. The results indicated that achieving equivalent usability to the desktop application on mobile is possible. These results show that both applications are suitable for the exploration of geographical Linked Data, however one application cannot fully replace the other as the suitability of an application ultimately depended on the users platform preferences.

The analysis of the results of the usability experiments enable readers to identify the main challenges to consider when designing a mobile Linked Data application. This is supported by the presented investigation of how different mobile and Linked Data visualisation challenges can apply to a tablet application. The findings of this study could help in making appropriate design decisions for future applications. However, it is important to note that the experiments have been carried out with a small number of participants and that more data would need to be gathered through user studies prior to making any final conclusions.

Table of Contents

List of Tables

List of Figures

Abbreviations

FOAF	Friend Of A Friend
HCI	Human-Computer Interaction
HTTP	Hypertext Transfer Protocol
OWL	Web Ontology Language
PHU	Practical Heuristics Usability
RDF	Resource Description Framework
RDFS	RDF Vocabulary Definition Language
RO	Research Objective
SCSS	School of Computer Science and Statistics
SPARQL	SPARQL Protocol And RDF Query Language
SUS	System Usability Scale
UI	User Interface
URI	Uniform Resource Identifier
US	United States
USPV	United States Political Violence
UX	User Experience
WWW	World Wide Web

Chapter 1

Introduction

The modern ubiquity of computers has enabled users who are less familiar with technology to consume more digital information. There is a large amount of data that is being published on the World Wide Web, and the number of data providers that make information available as Linked is increasing [5]. Linked Data is a set of guidelines for structuring, publishing, and interlinking data-sets on the World Wide Web [5]. With the recent advances in Linked Data, more and more of these users may be exposed to Linked Data resources [15] however raw Linked Data is designed for machine consumption and is accessible only to specialised knowledge engineers.

To enable users to easily access, understand, and interact with Linked Data, usable interfaces need to be created [15, 26]. This may be achieved by visualising the data in a way that relies on people's ability to easily perceive and understand visual patterns. In addition, as more data sets are inter-linked in Linked Data RDF (Resource Description Framework) format [5], the diverse information increases the effort required to understand the data, increasing the necessity of visualisation techniques even for technical experts [15]. Linked Data visualisation research has over time a set of requirements to create visualisations that are suitable for users of different levels of experience with Linked Data [15].

A significant portion of web data is spatial in nature [9], including Linked Data resources, meaning that the information can be related to geographic locations. This information is often analysed through data mining techniques [9], however understanding the results of this analysis is difficult [9]. Visualisation of geographical information on a map is recognised as the easiest method to help users understand it [9].

There has been significantly less research on mobile Linked Data applications than desktop or laptop applications interacting with Linked Data resources[17]. However given that mobile devices, such as smartphones and tablets, are increasing in popularity both in the developed and the developing worlds, it is important to study mobile Linked Data applications [39, 16].

Information visualisation on mobile devices introduces difficulties that are not present on desktop computers [14]. Mobile devices have limited screen space and have different input methods [14, 9]. Simply putting a desktop user interface on a mobile device will result in a bad user experience. There are general guidelines for user interface design for mobile devices [14, 15, 9], however most research currently focuses on specific parts of the interface, such as map navigation, interface navigation, etc. While formal research is scarce, there are user experience guidelines available for developers from the creators of the two most common mobile operating systems [41], such as iOS Human Interface Guidelines [24] or the Android Design Guideline [25], which give developers suggestions on how to design application interfaces that feel natural to use.

Most mobile Linked Data applications have been created for a specific use-case, and there are only a select few which allow the generic use of Linked Data information [15]. DBpedia mobile [4] and Stevie [6], presented in Chapter 3, are some of the very few mobile Linked Data applications, to the author's knowledge, which allow visualisation of geographical Linked Data without presenting the structure of the underlying data.

Determining which visualisation approach is better for a specific purpose may be achieved through analysing its ease of use [32]. Investigating the usability gives insight into how optimal a visualisation is for users, in terms of how well they understand the information being presented. In addition, testing an application's usability offers additional insight into the potential issues that users may encounter during their interactions [32].

1.1 Research Question

To what extent can mobile applications achieve equivalent usability to existing desktop applications when visualising Linked Data geographical information?

Usability is defined here as the attribute of measuring how easy it is to use and learn something based on the definition of Nielsen [32]. Based on the survey of Linked Data in mobile, geographical and Linked Data visualisation, presented in Chapter 3, there has been little to no previous research in trying to determine how mobile geographical Linked Data visualisation compares to desktop approaches.

1.2 Research Objectives

To answer the research question, the following steps and objectives are necessary.

RO1: To survey the state of the art of Linked Data and best practices in visualisation techniques for mobile devices, including the visualisation of Linked Data and geographical information.

RO2: To design an approach for visualising geographical information obtained from Linked Data end points, focusing on the usability and the user experience on mobile devices.

RO3: To create a prototype of the proposed design through an iterative development approach, which follows the guidelines concluded from the state of the art.

RO4: To evaluate the different stages of the prototype through user trials focusing on the usability of the application.

RO5: To compare the final prototype of the mobile application to a desktop counterpart through a usability study.

1.3 Technical Approach

This study involved the creation of a new Linked Data mobile application (*PVGeoVisualisation mobile*) and a baseline application (*PVGeoVisualisation*) based on the state of the art for Linked Data geographical information visualisations. These applications were used for the exploration and querying of an existing geographical Linked Data set, the United States Political Violence (USPV) data set. The applications allowed users to investigate data points and filter the information by building visual queries. Due to the smaller number of mobile Linked Data applications available to serve as a basis for the design, the mobile application has been designed and evaluated through four iterations with the help of usability studies. The usability differences between the desktop and the mobile applications were investigated in the final usability study.

The the United States Political Violence (USPV) data set was created by Prof. Peter Turchin and published in Linked Data format by the School of Computer Science and Statistics (SCSS) at Trinity College Dublin [7]. The data set describes political violence events that have occurred in the United States between 1795 and 2010. These events all have a location where they have occurred, which may be used to represent them on a map. The USPV data set is accessible through a Fuseki[1] SPARQL end point[2], allowing the applications to use and present the information on a map.

The desktop application was modelled based on the Map4RDF [27] application and tries to emulate as much of the functionality that can be applied to the USPV data set. Map4RDF is one of the few geographical Linked Data visualisation applications, to the author's knowledge, that enables objects to be placed on a map and allowing extra information to be added in callouts without the need for users to be aware of the structure of the presented information. Unfortunately, the researcher was not able to apply Map4RDF directly to the data set, hence the sgvizler [38] Linked Data visualisation library was used to create a web-based prototype of an application that mimics the design of Map4RDF but uses the USPV data set.

Due to performance issues of the sgvizler library on mobile, instead of using the library to create a web based mobile client, a native mobile application has been developed for the iPad. The iPad has been selected as the target device for the mobile prototype as it is the most popular tablet available [42]. The application accesses the USPV data set from the Dacura end point and presented the information on a map. Users were able to filter the presented information using a visual query builder, which is presented in Chapter 4.

The USPV data set in its original state did not possess geographical coordinates for each event, only a textual location property. Therefore, the data set has been enriched to contain the coordinates of the geographical locations that have been present to allow the applications use this information and place markers for each event on a map. Upon tapping or clicking the marker an information box, otherwise known as a callout, appears next to the marker showing the represented event's information. The details of the visualisation methods and how the data set has been modified are presented in Chapter 4.

[1] Apache Fuseki is a general purpose SPARQL server. `https://jena.apache.org/documentation/serving_data/index.html`

[2] `http://dacura.cs.tcd.ie`

1.4 Evaluation Methodology

Based on the author's findings, there has not been significant research in the domain of mobile geographical Linked Data visualisation for newer mobile devices such as smartphones and tablets. It is uncertain how research from ten years ago applies to newer platforms[3] and devices as the interaction methods[4] have changed over time along with users' familiarity with mobile devices due to their increased popularity [42].

In order to create a mobile application that is suitable for the proposed use-case, the initial prototype of *PVGeoVisualisation mobile* has been evaluated through a series of usability studies. This process was done in order to evaluate how suitable the application was for users, how they interacted with the application and how they interpreted the presented information. In addition, the researcher was able to identify a set of issues that made it difficult for participants to use the application or create frustration. This information, along with comments from the participants, guided the implementation plans for future iterations of the mobile application. The third version of *PVGeoVisualisation mobile* was compared to the desktop version through a usability study where participants were requested to interact with both devices, fill out a usability questionnaire for both approaches and then give their final comments along with stating their preference for one or the other.

Each usability test followed the same general scenario and methodology with some variations, depending on the focus of the experiment. Each experiment focused on investigating a specific area. This is described below:

- Experiment 1: Investigating the differences between two alternative designs for the visual query building interface
- Experiment 2: Investigating the differences between data set experts and novice users
- Experiment 3: Investigating the differences between the mobile and the desktop applications

The usability tests involved inviting in participants to interact with the applications and completing a set of prescribed tasks. During the users' interaction with the applications, the investigator was recording different measures that have been used to either assess the usability or issues in the application. In addition, the investigator recorded any comments or thoughts participants voiced to gain further insight into their experience. Once participants completed their tasks, they were asked to complete a usability questionnaire, detailed in Section 5.2.4, which provided quantitative measurements of the usability of the tested application.

1.5 Area of Proposed Contributions

The study aims to provide the following benefits to the fields of Linked Data, Linked Data visualisation, mobile interface design, and geographical information visualisation:

C1: An approach to visualising geographical Linked Data sets on mobile devices.

[3]Platforms such as iOS and Android, introduced in 2007
[4]Interaction methods have changed from a stylus based input to multi-touch gestures

C2: A tool to visualise geographical Linked Data sets on mobile based on the developed approach

C3: An approach to consume Linked Data from the WWW without the need for a designated server to manage and visualise the gathered information

C4: Experimental results on optimal visualisations and interface design for Linked Data visualisation and exploration on mobile.

1.6 Overview

This dissertation presents the study that has been carried out in order to investigate the differences in the usability of Linked Data geographical information visualisation between mobile and desktop approaches, and to create a mobile application that may be on par with a desktop alternative. The remainder of this document is structured the following way.

Chapter 2 serves to present fundamental concepts and support material that is required to understand the review of the state of the art in the areas of inked Data, Liked Data visualisation, Geographical Information Visualisation, Mobile human-computer interaction, and information visualisation.

Chapter 3 presents the state of the art in the domain of Linked Data on mobile, Linked Data visualisation, geographical information visualisation, and mobile human-computer interaction (HCI) and interface design. This section reviews recent research that has been used as the basis of the design for the PVGeoVisualisation applications. Some usability evaluation techniques will be presented that were applied in the usability evaluation of both applications.

Chapter 4 presents the design of the two created applications, focusing mostly on *PVGeo-Visualisation mobile* as the desktop application was strongly modelled on previous applications. The chapter details how the state of the art research has affected design decisions and showing the process from prototyping to implementation. It shows how different challenges have been addressed, and presents the end result that has been used in studying usability.

Chapter 5 presents the evolution of the mobile application through several usability studies to create an application that is well usable and is useful for the exploratory study of the USPV data set. Finally, the mobile and the desktop applications have been evaluated together to determine how well the mobile application compared to the other in terms of usability.

Chapter 6 concludes this report by summarising the findings of this study and evaluates how well the research objectives have been met. In addition, the answer for the research question will be presented and evaluated, and some possible focus for future work will be presented.

Chapter 2

Background

This dissertation investigates mobile geographic Linked Data visualisation. The main areas touched by the study are Linked Data applied on mobile devices, the visualiation of Linked Data, and how geographical information can be represented as Linked Data. However prior to be able to investigate and review these three areas, the building blocks that they are based on need to be understood by the readers of this dissertation.

This chapter aims to address the knowledge requirements of the State of the Art and Design sections by aiming to establish a basic understanding of the fundamental technologies and domains of the research areas of this dissertation. These areas include Linked Data and its technologies, the basics of information visualisation, and geographical information and its visualisation through the use of maps. In addition, mobile interface design is presented in order to help the readers understand some of the user interface (UI) and user experience (UX) design decisions of *PVGeoVisualisation mobile*.

2.1 Linked Data

Linked data is a set of principles that build on top of the technologies created for the Semantic Web in order to bring together different structured, machine readable data sets on the World Wide Web. The meaning of the data published on the Web should be well defined and have connections to other data and data sets, which in turn may have links pointing backwards [5].

Tim Berners-Lee has highlighted four basic 'rules' for publishing data on the web, otherwise known as the Linked Data principles [5], which serve as a guideline for the linked data project:

1. Things should have Universal Resource Indicators (URIs) as names
2. Hypertext Transport Protocol (HTTP) URIs should be used to facilitate the look up of data by both people and machines
3. When a thing is looked up, useful information should be provided using standards
4. Things should have links to other things to allow information discovery

These four rules enable users of Linked Data to much more easily consume and understand published information on the Linked Data cloud, as there is a sense of uniformity in how the data is being managed. This simplifies the usage of liked data resources and benefits its adoption by more data providers as more people follow the presented standard.

2.1.1 Linked Data Technologies

On top of HTTP and URIs, Linked Data uses the Resource Description Framework (RDF) to structure the data that is being presented in a graph based model. A graph based data model is a structure were objects and values correspond to nodes in the graph and the relationships are represented by the edges. RDF models the data using triples in the form of *subject*, *predicate*, and *object*, where the relationship between the subject and the object is expressed by the predicate. In RDF, the subject and the predicate are both URIs, while the object may be a URI or a literal value [5].

To describe and give meaning to things and the relationships between them, one can use vocabularies and ontologies. These may be defined by the RDF Vocabulary Definition Language (RDFS) and the Web Ontology Language (OWL). However to reduce diversity and the difficulty in understanding the data, the Linked Data momentum focuses more on the data and the interconnection of data entities instead of defining custom vocabularies and ontologies, and hence tries to reuse as much of the most common vocabularies as possible such as FOAF[1], Dublin Core[2], and Yago[3]. The Linked Open Data project is an example of an effort to convert existing data sources and publish it as RDF on the Web, while following the the principles highlighted above [15, 5].

To query and retrieve RDF resources, one may use the SPARQL Protocol And RDF Query Language (SPARQL) [43], which is able to manipulate and query RDF content over the web or in an RDF store. Queries may look for required or optional graph patterns, and the language allows for additional operations such as aggregation, sub-queries, value comparison, filtering, and constraining the source of the results. The results returned by a SPARQL query may be of the form of a set or an RDF graph [5].

2.1.2 Challenges in Linked Data Consumption

Due to the nature of Linked Data, information presented in applications may be taken from different data sources. While linked data applications provide controls to browse and navigate the information encoded, they are not intuitive or use-case and data specific based on the review of the applications in the domain [Section 3.3].

The amount of information encoded within the web of linked data is enormous (188 million triples in August 2014 [33]), and due to the size of this cloud and the fact that the publishing of the data is uncontrolled, there may be false links in the information [5, 15]. Determining whether a link can be trusted or not, whether it is valid or not, is difficult [44]. The problem may be made visible through the use of good visualisation techniques, which rely on human perception to identify outliers [15, 26].

Linked Data is heterogeneous in nature with different data types and content [5]. Heterogeneous data cause issues when the two data sources are merged. It is information in mostly unknown or in wide ranging format, possibly originating from multiple data sources. When

[1]Friend of a Friend: http://www.foaf-project.org
[2]Dublin Core: http://dublincore.org
[3]Yago: http://www.mpi-inf.mpg.de/yago-naga/yago/

merging data, the two or more sources need to be joined meaningfully before it is used by the application. The application has to be aware of the meaning of each relationship and value, and need to process knowledge of any mapping and similarities that may be present. It is unclear how heterogeneous data may be displayed and visualised to the user without the specification of the visualisation by an expert user of the data sets [12]. Previous applications focused on the data that is has been provided to the application and not the user interface and user experience, as highlighted by Nayebi et al. in [31]. Applications should provide interfaces which handle the underlying data without any issues.

Another possible key aspect of linked data user interfaces is the ability to add data sources to the application to further enhance the user experience. In addition to the application being required to be able to work independently of any known data source structure, the application would need to be able to handle and communicate issues that may arise such as the quality, trust, and relevance of the data [5, 15].

There have been examples of applications which have provided good examples for the navigation of linked data and the handing of different data sources [4, 44, 34, 20, 27], however there have been no examples to the author's knowledge which aim to handle the highlighted issues, summarised in Table 2.1, on a bigger scale [15].

Table 2.1: Summary of the challenges in linked data with respect to UI, data fusion, and data consumption

	Challenge description
1	Identify false links in the data set
2	Good visualisation methods to determine trustability of a link
3	Meaningful join of heterogeneous Linked Data
4	Generic visualisation of heterogeneous Linked Data
5	Provide a good user experience focusing on the meaning of the data and not its structure
6	Extension of the source data set
7	Data source independent operation
8	Handle the quality, trust, and relevance of the data

2.1.3 Summary

This section presented the basic of Linked Data along with the fundamental technologies that it is based on. It was shown how data stored according to the Linked Data principles is managed and can be accessed. Some challenges regarding the consumption of Linked Data have been presented and have been summarised in Table 2.1. These challenges can guide the design of a Linked Data application, helping to highlight areas of focus where additional considerations are required.

In the next section, the different information visualisation methods are presented along with a set of challenges that application designers need to consider during the planning phase of the application development.

2.2 Information Visualisation

This section introduces information visualisation. Initially, the two types of users of information visualisation tools will be introduced. Some basic information visualisation methods will be presented along with areas where designers of visualisation should pay attention to.

Information visualisation is the process of understanding data that has been represented graphically in a way that is more meaningful and easier to grasp by humans. The creation of such visualisations is facilitated through the use of computers. This is different from information presentation, which is the illustration or representation of information without any application or transformation to ease its understanding to the user [30, 15, 12, 14].

Information visualisation eases data processing by offloading the effort required to create a model of the information onto an external representation, and by taking advantage of the perceptive power of the human vision, the presented information can be understood by a person much more efficiently. This reduction in the cognitive load is achieved by reducing the effort in finding information and using visual patterns to detect information among other things [15].

The type of information visualisation method is dependent on the data that is to be presented, however in general, visualisations should present the data, avoid any distortions, make the presented data coherent, allow the comparison of the data to a reference, and allow different views on the data such as an overview or a view for fine details [30, 14].

Information visualisation enabled the use of the data that is being presented to a wider user base, as the information is encoded in a more human friendly manner.

2.2.1 Visualisation for Different Types of Users

There are two main categories of users for any visualisation scenarios. The *expert users* are a group of users who are familiar with the information that is being presented and have an in-depth knowledge of its underlying structure. *Non-expert users* are the users who have no knowledge of the underlying data and have no experience in using the data and with the involved technologies [15].

These two types of user have different requirements for what the visualisation of the information needs to achieve. Expert users require visualisations to take out the issues arising from the amount of data that is being handled. Visualisations allow for a good overview of the information and allows for easily determining problematic areas through outliers or unexpected trends. Non expert users require that the visualisation present them with a view of the data that allows them to easily understand and manipulate it without requiring any knowledge of the structure of the data [15, 1, 26, 26].

The main challenge is to find a good visualisation technique for the presented data that is useful for both user groups while fulfilling all the requirements for all users [15].

2.2.2 Information Visualisation Techniques

There are several different information visualisation techniques that may be used. Approaches, such as bar and pie charts, line and scatter plots, and histograms are used to visualise, summarise, and analyse statistical data, while tree and graph based structures are used to indicate hierarchies and relationships. To present geospatial information, visualisation techniques often place markers, bars, or charts on maps to indicate the geographical properties. Table 2.2 presents the set of most common visualisation methods that are being used [30, 26, 9].

2.2.3 Key Areas of Focus in Information Visualisation

The following points are some areas of focus that provide some guidelines in delivering a good user experience in information visualisation [14].

- *Mapping:* The mapping of information and the relationships between the objects being represented need to be clearly defined and kept consistent throughout the user experience.

- *Selection:* The relevant set of information should be selected for the visualisation, however sufficient data should be provided for the best experience possible. Unnecessary data will only confuse users as it adds extra cognitive load.

- *Presentation:* The information should be presented in such a way that it can easily seen. Visualisations need to be both attractive and have to display all the required data. If one of these aspects is missing, the visualisation may be ineffective.

- *Interactivity:* The device should provide some tools to interact with the data being presented. Higher usability of the visualisation may be achieved with better and more flexible interactions.

- *Human factors:* Human cognitive capabilities should be taken into when designing visualisations. Visualisations serve as a medium to facilitate the understanding of the data being presented, and visualisations should reflect this idea.

- *Evaluation:* Visualisation techniques should be tested by users in order to determine whether they are effective and useful.

2.2.4 Summary

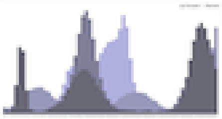
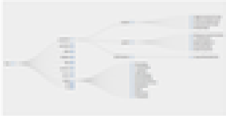
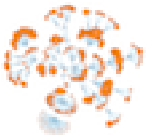
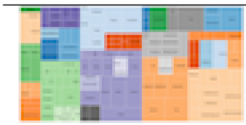
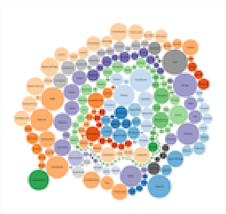
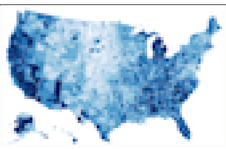
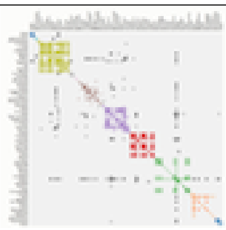
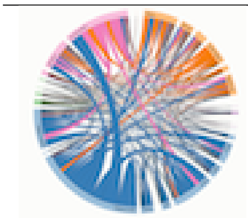
This section presented the basics of information visualisation. The two types of users, experts and novices, was introduced along with some of the basic requirements for each type. One of the main challenges in information visualisation is to fulfil the requirements for both user groups with one visualisation approach. Some example visualisation methods were presented in Table 2.2, however it should be noted that there significantly more visualisation techniques available. Finally, some areas of focus have been highlighted which should be considered when designing for information visualisation applications.

In the next section, the visualisation techniques specific to geographic information are presented.

2.3 Geographic Information

Geographic information, also referred to as geospatial information, spatial information, or location-based information, is a kind of information that has a reference to a geographical

Table 2.2: Table presenting a set of common information visualisation methods

	Name	Description
	Bar chart	This visualisation is the values of different nodes (placed o the x axis) proportionally to each other. The chart is capable to represent different relationships on one diagram for comparison.
	Tree	This is a graph based visualisation method where the structure corresponds to the hierarchical ordering of the data
	Hyperbolic Tree	Otherwise known as a force directed tree, this visualisation is a tree but the nodes have been spaced out to reduce information cluttering
	Treemap	This visualisation method is used to display hierarchical data using nested rectangles
	Sunburst partition	This visualisation is the tree-map method but is in a radial format instead
	Bubble chart	Bubble charts represent data in circle in a similar way than bar charts, however the information is less accurate. They are useful to present a lot of information in a small area
	Cloropeth map	This is a visualisation of geographic areas where the shading corresponds to the proportion of the variable being presented. An example visualisation would be the population density of a country.
	Adjacency matrix	This visualisation represents a graph where each node is represented along the x and y axes. The value of the relationship is shown by the colours in each cell.
	Chord diagram	This visualisation is representing a graph similarly to an adjacency matrix, however the nodes are arranged radially and the relationship is represented between the nodes by an edge.

entity. This reference may be a set of coordinates, but may only be a reference to a geographical location, which in turn has it's exact location defined. A majority of the data that is generated today is geographical in nature [9].

2.3.1 Geographic Information Visualisation on Mobile

Visual exploration of geographical information is essential, as the type of information that is displayed is complex in nature and without a graphical representation, creating a mental model of the presented information is very difficult. In addition, Burigat and Chittaro [9] have found that desktop geographical visualisation techniques may not necessarily apply directly to mobile applications due to additional restrictions and typically different use cases. Overlaying bar charts on the map is a technique which is not well applicable on mobile as the displayed chart occupies too much of the screen space available hiding most of the remaining information.

Users require visualisations that allow the direct manipulation of the presented data, allowing for a fast and intuitive user experience. Direct interaction also increased the confidence users have in the results that applications return. This approach is more preferable over automatic analysis tools, which tend to be often too complex and difficult to understand. By adding additional cognitive load, the usability of the application decreases [1]. Interactive maps, discussed below [Section 2.3.2], are an approach to solving this.

Modern geographic visualisation are capable to present interactive and manipulable maps showing different kinds of geographic information. These visualisations are created by incorporating generic cartographic knowledge into the data that is being presented. Maps allow the interaction with geographic information in such a way that the effectiveness of visual thinking is increased [1].

Andrienko and Andrienko [1] present a good overview of the different information visualisation methods of spatial data on interactive maps. They explain the maps should be able to adapt to the requirements of the user and they should be able to change the visualisation of information as they desire. Different types of information have different visualisation techniques that suit them. These are presented in Table 2.3. However users often wish to visualise information that are heterogeneous in nature. A possible approach to overcome the incompatibility issue is to visualise the two type of information separately, then merge the visualisations together.

Table 2.3: Visualisation methods for different types information (based on the work presented in [1])

Variable count	Type(s) of variable(s)	Relationship	Visual variables	Presentation methods
=1	nominal		colour or shape	area colouring or coloured signs
=1	ordinal		value	area shading, shaded signs
=1	numeric		value or size	area shading, standalone bars
⟩1	logical		colour	structured signs
⟩1	numeric	comparable	size	parallel bars
⟩1	numeric	included in common total	size	pies, segmented or parallel bars
⟩1	numeric	ordered inclusion	size	nested squares, parallel bars

Geo visualisation on mobile devices is very much desirable due to the availability of the exploration and analysis of data whenever and wherever the user wishes. However mobile devices add some restriction to visualisations, such as the small screen size, limited processing power, storage, and battery life. In addition, mobile devices have different input methods than desktops. Next the mobile techniques for interactive maps and off-screen element visualisation are discussed.

2.3.2 Interactive Maps

Dynamic queries allow users to analyse the data that is presented much more efficiently as feedback of the query that is being constructed is presented on the interface, allowing users to explore more of the application and the underlying data [9, 1]. To achieve this, the visual query building interface is directly linked with the graphical representation.

There are different user interface elements which may be used in visual query builders, some of which are presented below:

- *Range-slider*: These elements allow for the selection of continuous values. A subset of these are the discrete range-sliders, which have their range split into a series of discrete values.
- *Checkbox:* These elements allow features to be turned on or off
- *Dropdown:* These user interface elements allow the selection of a discrete value out of a set of possible values

Whenever a user alters a setting in the visual query builder, the application should reflect the changes on the map by displaying the results of the new query. Users should be able to gain more in-dept information of entries by selecting the corresponding markers. The detailed information should be presented in a contextual popup. This behaviour allows for a rapid and intuitive exploration of the underlying data.

It has been shown that on/off visualisations, the situation where only the elements that satisfy the requirements are presented, make it difficult for the users to understand relationships and gain a good insight into the presented information, as users tend to make significantly more errors and interface interactions. It has been suggested in [9], that alternative visualisations for the query results should be presented instead of hiding the elements that do not qualify. An example for such visualisation may be the addition of a bar next to the markers of elements which indicates the confidence value of the entry for the current query.

In interactive maps, it is often the case that there are data points that are not presented on the current view of the application, however it is often desired that users are given hints regarding these data points. In the next section, different off-screen visualisation techniques will be discussed.

2.3.3 Off-screen Element Visualisation

As the user interacts with the application on the map, some information may get outside the field of view due to actions such as zoom. It is important to communicate the existence of off-screen data elements to the user, as these information element may be important for the understanding of the data. There are several alternatives for the visualisation of off-screen elements, some of which are Overview&Detail, Focus&Context, Halo, and Wedge [10, 11].

2.3.3.1 Overview&Detail

The Overview&Detail method involves the presentation of the information in two separate panels. The main panel which takes up the majority of the user interface presents the detailed view of the information, and allows the user to move the map around, zoom, select entries, etc. The overview panel is a smaller UI element, which presents an overview of the information, displaying all data entries. In addition, some approaches have placed indicators in the overview to show users the scale of the detail view and help identify the users what section of the map they are viewing. A drawback of this approach is that the overview panel takes up precious screen estate, in addition the difficulty in understanding the data is increased as the users have to deal with two different scales of presentation.

2.3.3.2 Focus&Context

Focus&Context user interfaces allow users to see the information they are interested in in detail, while the other information are presented as an overview. While the method does not use a second UI element, it distorts the main view, potentially confusing users. In addition, the same issue of different scales remains from Overview&Detail.

2.3.3.3 Halo

The Halo off-screen visualisation method involves the surrounding of markers with a circle. These circles are wide enough to reach the edge of the current view of the map, giving some indication to the user about the presence of off-screen locations. A drawback of the method is that in case of a high number of off-screen elements, the circles get overlapped, and will not give hints to the user regarding the type of amount of information available in different directions.

2.3.3.4 Wedge

This method is an alteration of the Halo method, however instead of circles, the user interface presents arrows at the edge of the screen in the direction of the off-screen element. Some previous applications have adapted this method by altering either the colour, width, or length of the arrows to show more information.

In addition to support more efficient browsing through dynamic queries, most applications only present these arrows for off-screen elements that fully satisfy the current query.

2.3.4 Summary

Maps on mobile possess the same limitations that are present on traditional desktop environments, however there is the addition of the limitations of mobile applications, such as the small screen size, which affects the map based visualisation. Due to the limited screen size, the need for the visualisation of off-screen elements is much more required as the interface is not able to communicate the same amount of information [11]. These limitations are discussed in the following section.

Given that interactive mapping applications use a visual query interface, the limitations resulting from a small screen when applied to mobile become more apparent as the interface is taking up much desired screen space. In addition, controls that are being used on the desktop may not be applicable or natural on mobile [9].

One of the main challenges in creating a geographic information visualisation application is to balance the amount of information being displayed while avoiding information overload. This includes the balancing of the visualisation feedback and control elements of the user interface. Other potential issues arising from mobile is the performance of the application to be able to provide real-time feedback for the changes in the filters, as it has been previously shown that the usability of a map based visualisation is enhanced by immediate feedback [9, 11].

In order to understand how geographic information visualisation is affected on mobile devices, the main challenges that mobile application designs need to overcome have to be considered. These challenges along with design guidelines for mobile applications are presented in the following section.

2.4 Mobile Interface Design

Users of mobile devices tend to find themselves in different scenarios than desktop users, and hence the interaction with the device and the application become a utility, a secondary task, as the main focus of the user is on the environment, hence the amount of attention applications receive from their users is much lower than otherwise [14]. In addition the environment the device may be used in is variable, with different lighting conditions, which affects the perception of the visualisation displayed on the device.

2.4.1 Mobile Versus Traditional Environments

Mobile devices have several additional restrictions that make interface design and information visualisation more difficult.

2.4.1.1 Screen size

Mobile devices have much smaller screens than traditional desktop environments. This limits the amount of space that may be used for interface elements and the display of information and visualisations, hence desktop visualisations cannot be directly translated to mobile environments purely due to their size, as the data overview would not be entirely visible on the small screen [14].

2.4.1.2 Input

Mobile devices have different types of input from desktops. users interact with the user interface directly through the use of a touch screen or a pen, while in traditional environments users would be interacting with the interface through a mouse or a trackpad. In addition while designing the user interface and especially the visualisations, one must consider that part of the available screen is used for input and the space available when the user is using an on-screen keyboard is further reduced, limiting the amount of space for visual feedback [14, 4].

Traditional solutions of information presentations such as Overview + Detail or Focus + Content, presented in Section 2.3.3, do not apply well to mobile devices due to the limited screen space or the amount of non-essential information presented. However other methods such as off-screen information referencing, such as Wedge [Section 2.3.3], and scroll&zoom have showed to be good alternatives for reducing the amount of information presented to the user at any one time but still informing them of the availability of other information [14, 9].

2.4.2 Other Limitations on Mobile

In addition to the limitations highlighted above, some other considerations need to be taken into account when designing visualisations for mobile [14].

- *Connectivity:* mobile devices tend to communicate through wireless network connections, and the availability of this connection is not reliable. While the issue becomes less common as wireless network availability and reliability are improved, one must keep in mind the potential issues that may arise due to this limitation.
- *Frameworks and libraries:* Due to the nature of mobile devices and the limited amount of resources available to them, some traditional visualisation environments may not be available.

2.4.3 Basics of iOS Interface Design

The design language of Apple's iOS relies strongly on three concepts: deference, clarity, and depth [24]. This section overviews the main ideas and areas of iOS design with respect to the aforementioned concepts.

2.4.3.1 Deferring to Content

The design of iOS encourages the presentation of the content over the interface and controls, it is the information that should be the main focus of the application using as much of the screen space as possible. An application should be aesthetically pleasing, combining appearance and functionality in a coherent way. Therefore the design layout should not affect the usability of the controls and the interface, buttons and informative text should be well spaced and easy to interact with.

Different areas of the screen have different level of importance and focus in users' eyes and therefore information should be placed in such a way that it reflects their importance in the interface. More important elements should be placed towards the top left corder of the screen, while less important elements should be placed towards the bottom right corner.

2.4.3.2 Providing Clarity

An application should allow users to directly manipulate any information that is being presented on the screen, may it be an image, text, or any other resource. A set of multi touch gestures have been defined by Apple that are familiar to users and how these gestures are used should be easily understood by users and fit in the context of both the application and the iOS platform.

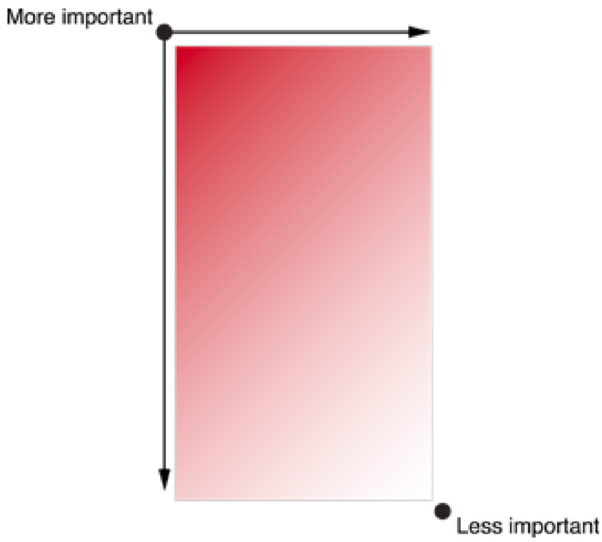


Figure 2.1: Image indicating the level of importance of parts the screen in an iOS application as shown in [24]

Navigating the application should not be an issue to users, and the hierarchy of screens and elements should be well understandable to users. The interface should give visual hints to users regarding their position in the navigation hierarchy. The use of colours and contrast should highlight important information and make it feel better in context.

2.4.3.3 Use of Depth

Applications should use different visual layers to present information and navigation between these different layers should be achieved through realistic animations that reflect the motions of everyday objects. Animations should be used in the application to provide visual feedback on actions and to enhance the users' feel of manipulating information directly. In addition animations may help users to see the results of their actions easier. However the use of animations should be consistent and not exaggerated, as it may degrade usability.

Information that is not required by users at all times should be shown on a different layer of the interface and shown only when it is required, however the transitions between these layers should be natural and help users in keeping track of the application context and navigation hierarchy.

Chapter 3

State of the Art

The state of the art will focus on presenting the forefront of the research areas covered by this dissertation. The chapter focuses on four themes: mobile Linked Data applications, Linked Data visualisation, the design of query building interfaces and interface evaluation techniques.

Initially, the use of linked data within the mobile domain will be investigated. Use cases and mobile applications using linked data will be presented along with their approaches and solutions. The storage and replication of linked data on the mobile devices will be briefly described. A short discussion of geographical information within the domain of linked data will follow. Concepts such as GeoNames and LinkedGeoData will be presented along the methods of how spatial information contained within these data stores may be used.

Linked data visualisation will be introduced, highlighting common techniques and a set of requirements for visualisation of information suitable for the different user types. Some examples of previous approaches and applications will be presented and analysed.

The state of the art will also investigate the design of visual query building interfaces, which are capable to reflect the state of the query being built on the displayed data, giving users real-time feedback. Finally, the evaluation methods for the quality and user experience and usability of visual data analysis tools will be investigated, including the use of the System Usability Scale.

Prior to concluding with an evaluation of the chapter, all presented examples will be compared and common trends within the approaches will be highlighted. In addition, it will further analyse the presented approaches and identify a set of best practices and guidelines that will be used in the design chapter.

3.1 Linked Data Visualisation

"Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods."[1]

The visualisation of Linked Data is essential to expand the usage to a wider group of users, as non-expert users do not necessarily have knowledge of the format and structure of the information that is encoded in the RDF format [15]. As previously explained [Section 2.2], visualisation breaks down the knowledge barrier that exists in the usage of Linked Data

[1]linkeddata.org

and allows the interpretation of information by the use of concepts that are more familiar to users [15].

Like any other information domain, linked data has two clearly identifiable user bases: expert and non-expert users. Expert users understand the foundations of Linked Data and the involved technologies, they have experience with the information that is being presented, both in terms of the structure of the data and the possible use cases, and hence are able to use the data in its raw format that is RDF or a SPARQL result set. Non-expert users on the other hand, while they may be aware of the existence of Linked Data, are not comfortable or familiar with the technology and hence are unable to use and interpret the data without help [15, 12].

3.1.1 Requirements for Linked Data visualisation

Dadzie and Rowe [15] have stated a set of requirements for the consumption and visualisation of linked data that are essential to deliver a good user experience with a high usability rating for both expert and non-expert users. A summary of their findings, which have been established based on a review of previous research, is stated below.

Most visualisations share a set of essential features which are required in order to help users interpret complex or large amounts of data. All visualisation tools are required to be able to

A1: create and present an overview of the data,

A2: filter the information to eliminate less important information, and

A3: be able to drill down into, and show more detailed information of areas of interest.

In addition, more complex visualisations, such as the ones required for Linked data, must be able to

B1: visualise relationships between entries,

B2: support the use and display of multidimensional data, and

B3: give the ability to export the visualised data, to allow the user to transfer it to other applications.

Linked data visualisation tools should be able to visualise the underlying data in a meaningful way that is applicable to data and helps the user to better understand it. The visualisation should ease the tasks that the user has to execute using the application, and help the user achieve their overall goal much more efficiently [15].

In addition to the requirements above, linked data visualisation tools should possess the following features in order to reach the full potential that linked data has to offer. These features are mostly required by expert users, however they benefit non-expert users equally [15, 1]. Linked data visualisation applications should offer

C1: clear and intuitive navigation through the web of data, with the option to go back and list the previously visited nodes;

C2: the possibility to explore the data, without restrictions, starting from any node;

C3: users the ability to access the underlying data to inspect links, and identify errors and noise;

C4: offer the option to execute custom SPARQL queries;

C5: extract the underlying RDF data to reuse it in other applications.

To enable the use of the application by a wider user base, visualisation tools should possess features to

D1: navigate easily through large data sets;

D2: allow exploration of the data to gain understanding of it;

D3: offer the creation of queries through helper methods, such as visual elements;

D4: allow analysis of regions of focus;

D5: present the results of queries and usage to others.

3.1.2 Mobile Applications with Geographic Linked Data Visualisation

This section introduces three mobile Linked Data applications that feature geographical information visualisations. these applications are DBpedia mobile, Stevie, and mSpace mobile. These applications are compared and analysed in detail in Section 3.1.5.

3.1.2.1 DBpedia Mobile

DBpedia mobile [4] is a mobile application created for tourists which allows them to discover real-world objects around their current location. The application uses resources from the semantic web to populate its map with information.

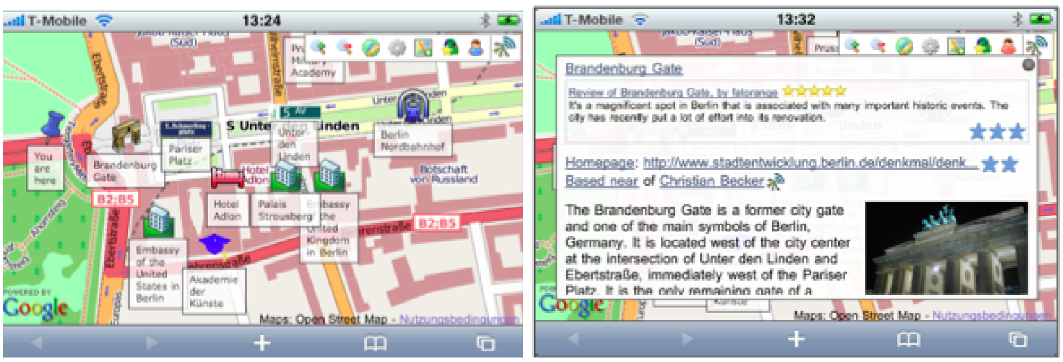


Figure 3.1: Images of the user interface of *DBpedia mobile* showing information on the map, displaying an overview information panel. [4]

The application filters information based on the user's location, and for each entity, the related information from the main linked data sources are loaded. Users are able to see this information on a summary screen, which is presented whenever the user clicks on an icon on the map. The application allows the users to see more information on the detailed information view, and the user may navigate links that are originating from the viewed entity.

The user is allowed to browse any data sets, as the application is not restricted to a specific one, instead it uses a semantic web search engine to gather links and build a semantic graph for the user, which in turn may be extended as the user is navigating and discovering the links between objects in the graph.

DBpedia mobile is designed to be user friendly, and does not require its users to be aware of the data format that is being used, however it allows domain experts to query data using a SPARQL interface.

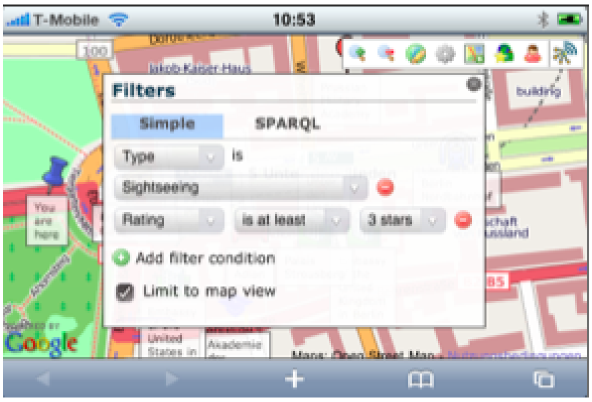


Figure 3.2: Image of the user interface of *DBpedia mobile* displaying a SPARQL query interface. [4]

3.1.2.2 Stevie

Stevie [6] is a mobile application for the collaborative sharing and creation of points of interests. The application allows its users to create, modify and delete POI information from the system.

The application displays all known information on a map, indicating each data entry with a marker. When a user taps on a marker a dialog is presented with detailed information of the data entry. The application works with one set data-set and does not offer unrestricted navigation of the semantic web as the underlying linked data to resources on LinkedGeoData and GeoNames is not exposed to the user.

3.1.2.3 mSpace Mobile

mSpace [44] mobile is a context aware mobile semantic web application. The application is able to maintain the context of the information currently being presented while the user is exploring related information from other data sources that are linked from the current known information. The application has been specifically designed to support smaller screened devices with the ability to manipulate the information through visual elements and to be able to handle data sets of large scale.

The application's user interface has been designed for small screened devices allowing users to run complex queries with little effort. The semantic web queries are enhanced based on the context of the application. In case the user is focused on location related searches, the application would use the device's location to tailor results further. Search results are

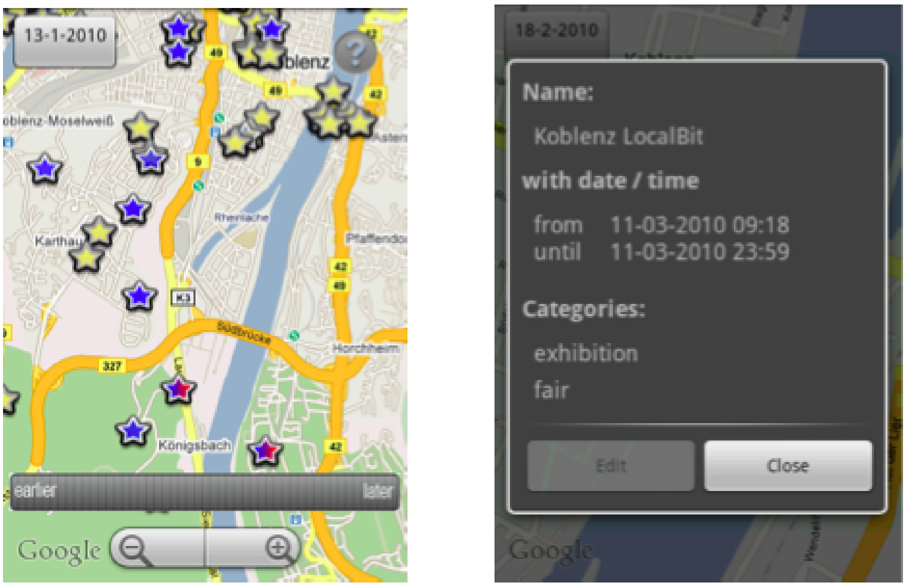


Figure 3.3: Images of the user interface of *Stevie* showing the map and the individual entry views. [6]

displayed on a map in order to give additional information that is visually encoded to reduce the cognitive load on the user and enhance the user experience.

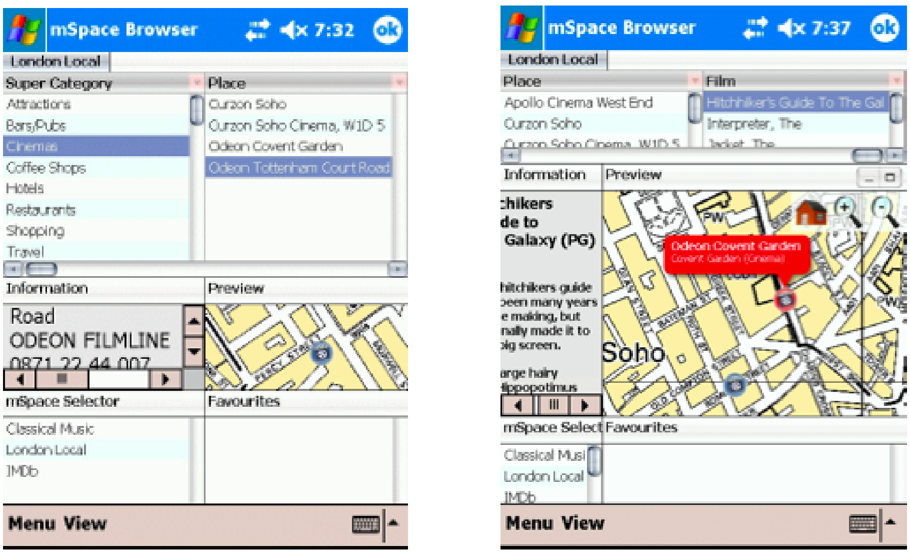


Figure 3.4: Images of the user interface of *mSpace mobile.* [44]

The application offers different use cases, not just location related searches. This is achieved by the unrestricted sources of information. A user may define any semantic web end point as a resource, and the application would use it in order to provide the best use case and information to the user. The different modes of interactions offered are unrestricted data browsing, filtering information based on user constraints, information overview of selected entries, information preview, and saving information of current selection.

3.1.3 Mobile Applications with Generic Linked Data Visualisation

This section introduces mobile Linked Data applications which feature generic visualisations of the underlying information. These applications are Ontowiki mobile, Qpedia, More!, Who's Who, and wayOU. These applications are compared and analysed in detail in Section 3.1.5.

3.1.3.1 Ontowiki mobile

OntoWiki Mobile [20] is a mobile semantic web application which provides a generic approach to the usage of linked data information. The application allows its users to browse data sets, filter information and present it in views that adapt to the data, and in addition, the application also allows its users to edit and add data to the data set.

Within its browsing functionality, Ontowiki Mobile presents the underlying data in a list format to facilitate navigation within the data set. The data entries are ordered in classes, and the UI reflects the class hierarchy.

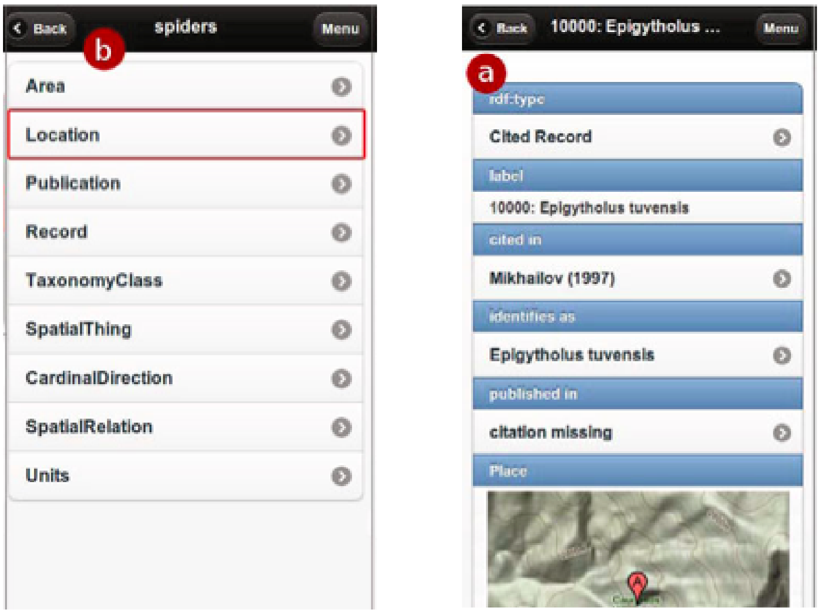


Figure 3.5: Images of the user interface of *Ontowiki mobile.* [20]

The browsing functionality of the application allows the user to browse for data types which contain a specified property. The user initially specified the property that they are looking for, then the application presents the results set in a list view. Each entry in the list is a property, based on which the result set may be further filtered. The user may view individual entries in a list-based view where each entry corresponds to an attribute-value pair. In addition to viewing entries, users are able to edit and create entries using the authoring interface.

3.1.3.2 Qpedia

Qpedia [17] is a Semantic Web application which allows the dynamic use of linked data without requiring the user to know about the structure of the information or SPARQL. The application allows the users to search for information on DBpedia through the search by example approach. The users are able to search for entities based on keywords, properties, or geographical location. The searched terms are constrained to the values of the RDF properties.

The application offers three main views. The initial view of the application allows the specification of search terms. When the user submits a query, the application presents the results in a list view. The application allows the view of individual entries, where the list of the property-value keys are presented. In the case where and entry possesses a GPS coordinate, the user is able to view the information on the map.

In addition to the previous interfaces, the application allows the users to view and edit the SPARQL queries being used, and the properties of entries.

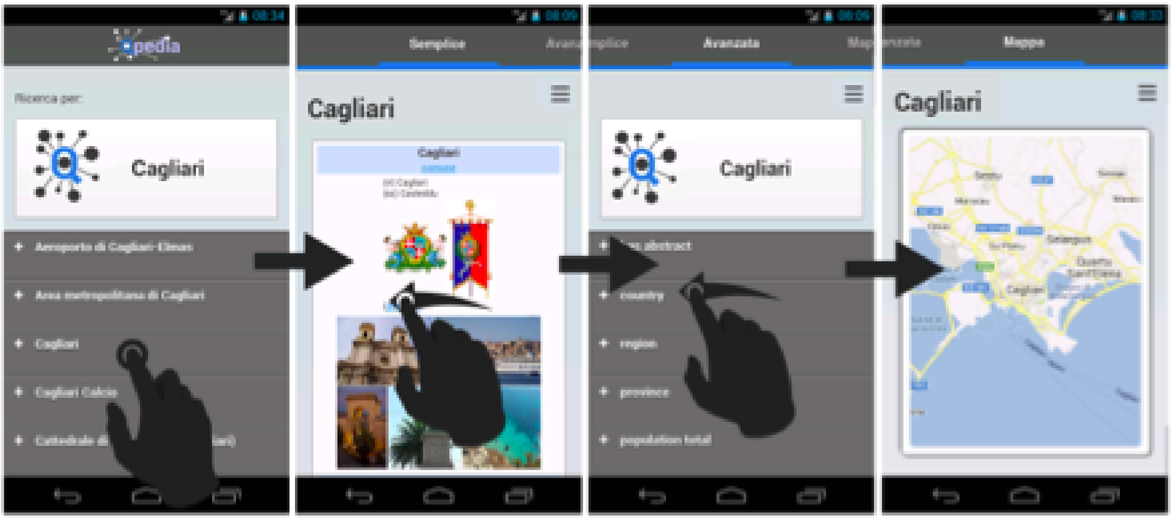


Figure 3.6: Images of the user interface of *Qpedia*. [17]

3.1.3.3 More!

More! [34] is a semantic web application that allows its users to find information regarding researchers at conferences. Information such as the current and previous works, and contact information can be viewed through the application.

The application uses information from the Research.fm data set in order to create the view interface. Users are able to view a person's information upon scanning a QR code. The application does not allow the users to explore further within the data set, and does not expose the underlying FoaF data, which could be used to offer more insight to the user.

3.1.3.4 Who's Who

Who's Who [13] is a linked data application which allows users to gain information of publications of researchers. Initially, the user is presented with a list of researchers. By selecting a researcher, the user is brought to the individual researcher view where publications are grouped together based on year and presented in a card deck visualisation, each bubble corresponding to a publication. Upon tapping on a bubble, the user is presented with a popup which displays the list of collaborators.

The application works agains a single specific data set which is not interlinked with other data sets, hence the application does not support browsing outside of the known domain.

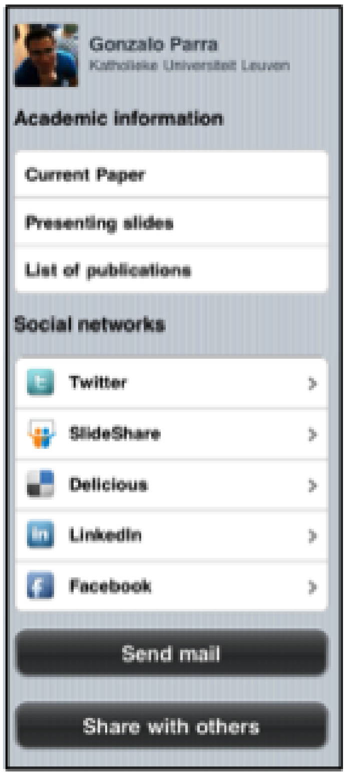


Figure 3.7: Image of the user interface of *More!*. [34]

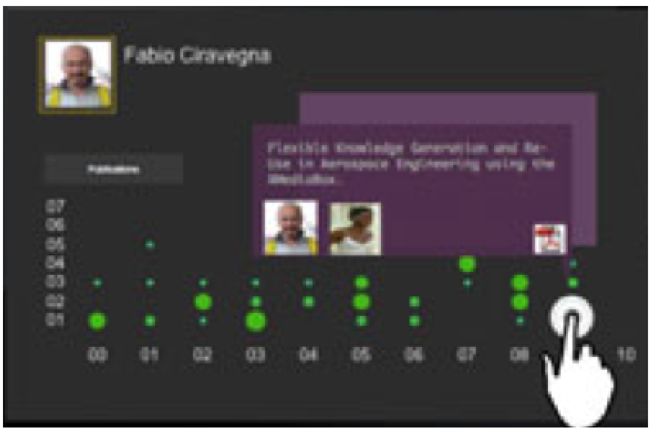


Figure 3.8: Image of the user interface of *Who's Who*. [13]

3.1.3.5 wayOU

wayOU [19] is a mobile application which is able to provide social and location based information to students and staff of the Open University. Users are able to browse information about other users and edit their own information.

The application offers some basic interfaces. Users are able to browse the list of other users and view the details of each individual users. In addition, they are able to view their own profile and edit the information associated with it.

Figure 3.9: Images of the user interface of *wayOU*. [19]

The application does not allow the exploration of the web of linked data as it is restricted to a single data set. Users are not able to query, view, or export the underlying RDF data to enable other use-cases.

3.1.4 Desktop-based Linked Data Visualisation Applications

In this section, map4RDF and the GeoNames desktop-based Linked Data visualisation applications will be presented. These applications are compared and analysed in detail in Section 3.1.5. They are specific to visualising geographical information, however other generic Linked Data desktop-based visualisation applications such as Fenfire, IsaViz, LESS, OpenLink, RDF Gravity, and RelFinder and presented in the appendices [Chapter A].

3.1.4.1 map4RDF

Map4RDF [27] is a faceted Linked Data browser that enables geographical Linked Data information to be visualised on an OSM or Google Map. The information has to be encoded according to the LinkedGeoData ontology or the Basic Geo Vocabulary[2] in order for it to be displayed.

The application allows users to filter the geographical objects that are being visualised using a sidebar, where the list of OSM classes (such as airports, cities, routes) is presented.

[2]http://www.w3.org/2003/01/geo/

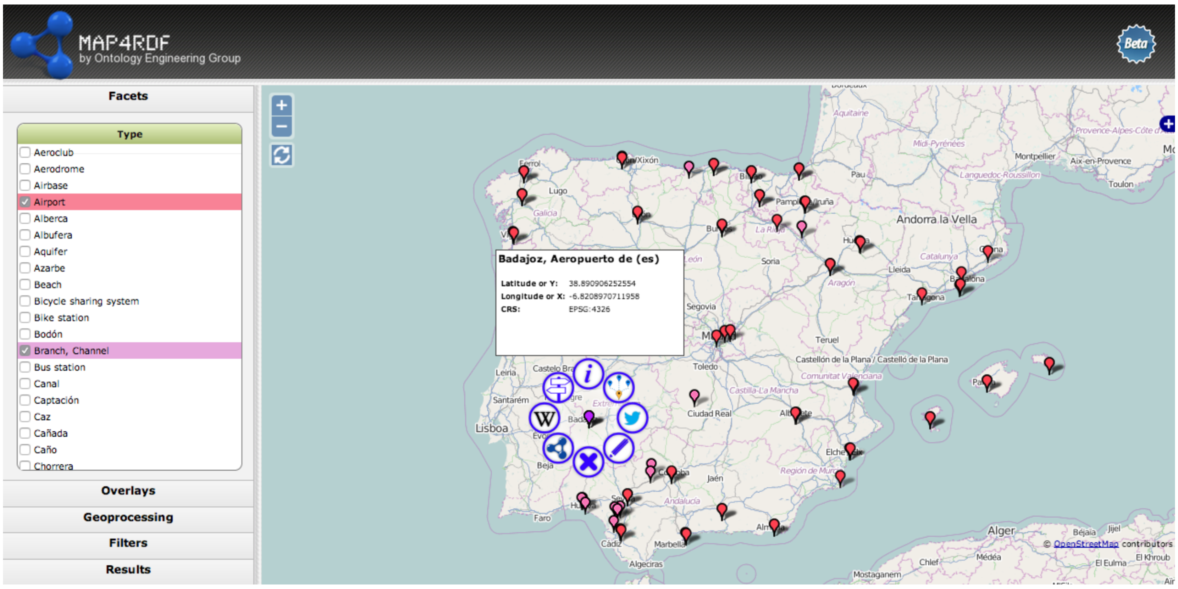


Figure 3.10: Interface of map4RDF as shown in [27]

The application allows for the definition of custom queries and the presentation of polygons instead of markers.

In addition to the presentation of the data, the application allows for the filtering of the information based on geographical features, such as "find all nearby features", allows users to inspect the RDF information, as well as see further information from Wikipedia or other linked resources.

3.1.4.2 GeoNames browser

The GeoNames browser[3] allows the visual exploration of the underlying data. Much like in map4RDF [Section 3.1.4], the information is presented on a map and each corresponding entry is displayed as a marker. The application however only loads the information for the current view, and whenever the view changes the data points contained in the visible rectangle are loaded.

The information can be filtered based on the types of the displayed information, which are defined in the GeoNames ontology. By clicking on the "Layers" button, a popup window is presented to the users where the different classes of information are presented. Users are able to select the classes of data points that they wish to display and the map will be refreshed with the new information according to the specified filters.

3.1.5 Analysis and Summary of Applications

Table 3.1 compares all applications with respect to the presented requirements for Linked Data visualisation tools [Section 3.1.1]. The applications had various use cases and some of them were even non use-case specific meaning that they presented Linked Data information and allowed its navigation without any assumptions put over it. While applications without any assumptions allow for a much more liberal exploration of the Linked Data space, their

[3] http://www.geonames.org/

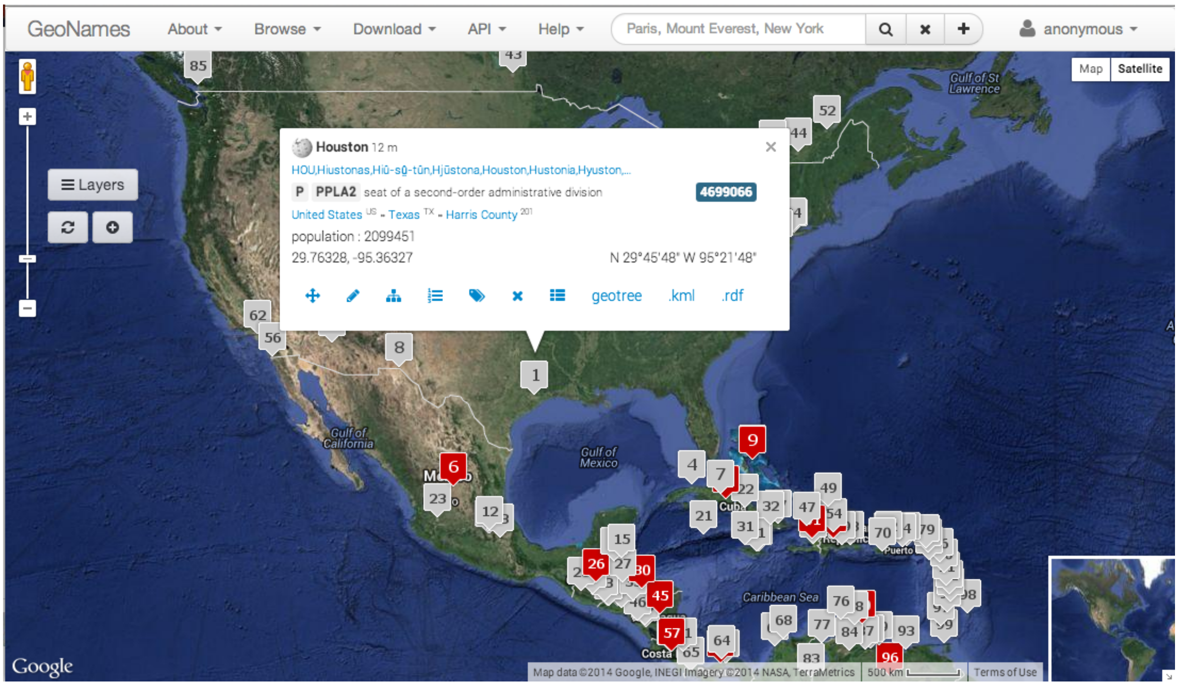


Figure 3.11: Interface of the GeoNames browser

usage for a specific purpose may be difficult such as for tourism, as the discovery of the sought for information is much more difficult without the specialised visualisation.

Visualisation allows users to quickly gain insight into the information that is presented and grasp the relationships that are presented, which is a fundamental requirement for a Linked Data visualisation tool. The mobile applications presented have been created to address a specific need, such as the presentation of nearby tourist attractions, and presenting information from LD sources. In general, desktop-based Linked data applications, presented in Appendix A aim to visualise the structure of the information and the relationship between entities.

The applications presented different navigation techniques that may be applicable for the proposed application that supported the exploration of the data set. However DBpedia mobile and map4RDF were the only applications which were designed with novice users in mind fulfilling the visualisation requirements D1, D2, D3, D4, and D5 to allow the application to be used by a wider user base. Users were able to use any Linked Data set, gain more information into the presented data, create custom queries to filter the information to the needs of users, focus the information to a limited area, and allow to share the information with others.

Table 3.1: A comparison of different mobile linked data visualisation applications with respect to the set of outlined requirements

	DBpedia Mobile	Ontowiki Mobile	mSpace mobile	Stevie	Qpedia	More!	Who's Who	wayOU	map4RDF	GeoNames browser
Overview data	yes	no	yes	yes	no	no	no	no	no	yes
Filter out data	yes	yes	yes	no	yes	no	no	no	yes	yes
Detail/Drill down view	yes	yes	yes	no	no	no	no	no	yes	yes
Information preview	yes	no	yes	no	no	no	no	no	yes	yes
Visualise relationships	partial	yes	no	no	no	no	yes	no	no	no
Multidimensional data	yes	yes	yes	no	yes	no	no	no	yes	yes
Export data-set	no	no	yes	no	no	no	no	no	no	yes
Intuitive navigation	yes	yes	yes	yes	yes	-	yes	?	yes	yes
Navigation history	yes	yes	no	no	yes	no	no	no	no	no
Ability to explore	yes	yes	yes	partial	partial	no	partial	partial	yes	yes
Raw RDF view	no	no	no	no	no	no	no	no	no	yes
Custom SPARQL	yes	no	no	no	yes	no	no	no	no	no
Large data set support	yes	no	yes	no	no	no	no	no	yes	yes
Unrestricted LD navigation	yes	no	yes	no	no	no	no	no	no	yes
Visual query building	yes	no	yes	no	no	no	no	no	yes	yes
Share current view	no	no	yes	no	no	no	no	no	no	yes

3.1.6 Linked Data Visualisation Libraries

In this section two Linked Data visualisation tools are presented which can be used to create non use-case specific visualisations, using a Linked Data source.

3.1.6.1 VisualBox

VisualBox [21] is a general Linked Open Data visualisation tool allowing users to define queries for any Linked Data end point using the SPARQL query language and visualise the results of the query. When creating visualisations, users have to define the data model they wish to use for the visualisation and create the corresponding query so that the data model is populated with results. The application creates the desired visualisation, may it be a chart or a map, based on the resulting model instances.

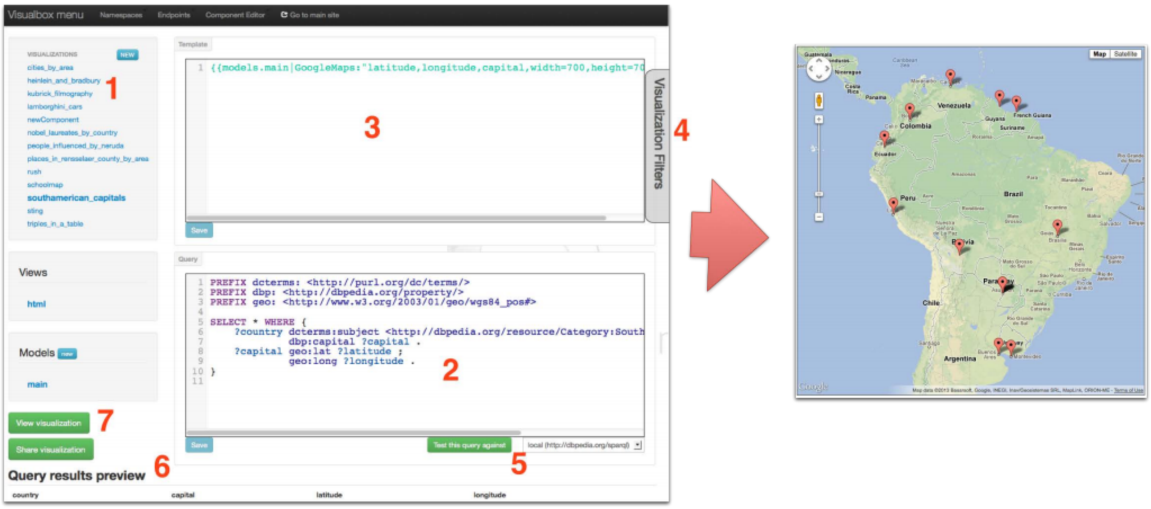


Figure 3.12: Query construction interface and resulting map based visualisation in VisualBox [21]

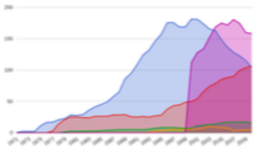
The application allows the creation of visualisations, however without additional development work, it does not allow the filtering of the information as the presented visualisation does not have controls to alter the query that has been applied to the source data set. This example is not a use-case specific application of Linked Data visualisation but a tool allowing others to create their own linked data visualisation for their purposes.

3.1.6.2 sgvizler

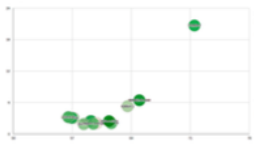
Sgvizler [38] is a JavaScript library for visualising the results of Linked Data queries. The library offers several visualisation options such as area charts, bubble charts, geo maps, tree maps, and force-directed graphs. Sgvizler also offers a map based visualisation, where individual entries in the results set correspond to markers. The application is similar to VisualBox, however the library does not offer the creation of queries and visualisations using an interface but required coding and development skills to create visualisations.

The library is able to execute SPARQL SELECT queries and the results of these queries are visualised through the Google Charts[4] visualisation library. This tool can be used to

[4]Google Charts: `https://developers.google.com/chart/`



(a) Area Chart



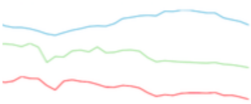
(b) Bubble Chart

(c) Pie Chart

(d) Geo Map

(e) Tree Map

(f) Timeline



(g) Sparkline

(h) Scatter Chart

(i) Force-directed Graph

Figure 3.13: Chart examples of sgvizler as presented in [38]

create web based data visualisation applications using any Linked Data end point as its source.

Figure 3.14: Map based visualisation example using sgvizler as presented in [38]

3.1.7 Summary

In this section the requirements for Linked Data visualisation have been presented based on the survey work by Dadzie et al [15]. Example mobile Linked Data applications have been reviewed and compared with respect to features such as filtering of data, data overview, intuitive navigation, and the ability to explore the data set. Map4RDF and the GeoNames browser have been presented that are desktop-based alternatives. In addition, some Linked Data visualisation libraries have been presented which ease the creation of different Linked Data visualisations.

In the next section, the different approaches to represent, consume, and manage geographic Linked Data are presented.

3.2 Geographical Information and Linked Data

A significant amount of everyday information is linked to spatial features such as places, roads, landmarks, etc. This kind of information has the same challenges as any other information, the increasing amount makes it difficult for users to understand it, hence its visualisation is required. However unlike other type of data, this is best visualised on a map.

3.2.1 The Basic Geo Vocabulary

The basic RDF geographic vocabulary[5] allows providers of Linked Data resources to represent geographic information according to the WGS84, a cartography and navigation standard. RDF entities may use the lat and long terms to define geographic coordinates for points and features amongst other. The example below demonstrates the use of the vocabulary in RDF/XML to store a point along with its coordinates.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#">
  <geo:Point>
    <geo:lat>55.701</geo:lat>
    <geo:long>12.552</geo:long>
  </geo:Point>
</rdf:RDF>
```

3.2.2 GeoNames

GeoNames is a Linked Data geographic names database which possesses entries for several million locations and points of interest on Earth[6] and stores corresponding geographical and elevation coordinates. GeoNames offers features such as textual address search and access to browse its data set. In addition visual exploration of the data set if possible through the GeoNames browser application [Section 3.1.4].

3.2.3 LinkedGeoData

LinkedGeoData [40] aims to publish geographic information on the Linked Data cloud by converting information from the OSM project to RDF. To enrich the published information, the data is interlinked with DBpedia and GeoNames amongst others.

LGD provides access to the information through several methods including data set downloads, live and static SPARQL end points[7], Linked data through URI dereferencing, and a REST API.

Links to other data sets are established based on LGD node information with a very high precision on a class by class basis using labels and spatial information to match entities. LGD allows the execution of spatial queries to limit the mount of information that is being requested, and example of such filter is the matching of entities which are within a radius or a bounding rectangle.

[5]http://www.w3.org/2003/01/geo/

[6]GeoNames: http://www.geonames.org/

[7]Live end points are synchronised with OSM continuously, where static end points are a copy of the OSM data on a specific date

OSM data is constructed based on two main types of data: nodes and ways. Nodes are entities which describe a point in space and therefore have a GPS coordinate. Any other entity, polygons, lines, etc are described using ways, which are a series of nodes. Entities such as roads, places, rivers, regions, etc are described using ways. LinkedGeoData reflects the OSM data structure by having two different label types, ldg:node⟨id⟩ and lgd:way⟨id⟩. Each data entry possesses one such label, where the ⟨*id*⟩ part corresponds to the entity's ID on OSM.

3.2.4 Applications Using Geographic Linked Data

DBpedia mobile, Ontowiki mobile, mSpace mobile, STEVIE, and map4RDF have been presented in Section 3.1.2 which have used geographical information in their visualisations. These applications use the encoding of the basic geo vocabulary to store geographic coordinates and present them on a map. Some of the applications use the LinkedGeoData ontology, reusing classes for geographic entities instead of defining custom ones.

3.2.5 Geographic Linked Data Within the USPV Data Set

The proposed application would use the United States Political Violence Linked Data set [7] and visualise the information on a map. Each event object in the data set has a corresponding location node, however this information only possesses an unstructured location name. In order to display this information on a map, the location name needs to be geocoded; a process where an address is converted into coordinates. The application would have to either use a second Linked Data set, such as LinkedGeoData, use a geocoding service such as GeoNames, or use an enhanced version of the USPV data set to receive the coordinates.

One approach would be for the application to look up the geographical entity in the target geographical dataset to retrieve coordinates. While this approach would ensure the most up to date information and allow users to access extra information along with the coordinates, the access of a secondary data set would slow the application down.

An alternative approach would be for the application to use an augmented version of the USPV data set, where the location nodes would possess the latitude and longitude coordinates in addition to the unstructured location information. These coordinates would be encoded using the basic geo vocabulary. However this would require to modify the original data set, as well as the data set's ontology, to inform the consumers of the USPV data set regarding the availability of this new information. This approach would reduce the amount of effort needed in visualising the information, hence speeding queries up, however it would not enrich the displayed information with links to other sources.

3.3 Linked Data on Mobile

While linked data has been heavily explored and used on traditional computer environments, there have not been many examples of mobile applications or research that focuses on exploring and improving the linked data user experience. Examples of mobile Linked Data applications have been introduced in Section 3.1, however focus of the presentation was on the visualisation techniques, the user experience of the application and the Linked Data. In this section, those mobile applications will be reviewed in relation to their use of Linked

Data, the technologies involved, and the application architecture used. Some trends will be highlighted, such as the common issues that have been encountered in the development of the application and how these have been overcome.

3.3.1 DBpedia Mobile

DBpedia Mobile [4] allows its users to discover, search and create Linked Data. Information that is relevant to their current geographic position, is presented to users through a map view. Users also have the option to browse Linked Data through a Fresnel [36] browser.

The initial view presents the user with a basic view of the surroundings showing some information from LinkedGeoData. All entries are marked with an icon, the type of icon that is shown depends on the Yago type of the entity. Upon clicking on an icon a summary view is presented, from which users may access the full view of the resource or other elements such as images or the description. The full view of a resource is presented in the Fresnel browser from which the user may navigate to other RDF resource is they have been linked.

DBpedia mobile uses a server-client approach to achieve its functionality. Each user has their own RDF graph which is used to generate the different views of the application. When a user requests an information, the mobile device submits this request to the server, which in turn will query the RDF graph and extend the results set using the search result from Sindice[8]. The results set is then used to generate the view on the server side that will be presented to the user on the mobile. The application does not restrict the sources of information that are being used, however the server only has a version of DBpedia, any other sources of information that are included in the visualisations have to come from Sindice. The server considers owl:sameAs links during the augmentation of the users' graph to enrich the information being presented.

Users are able to apply filters to the information that is being presented to them. Filtering may be based on the type of properties the application considers but may also consider the geographic coordinates. This filtering functionality us used to present the initial view to the user. The geographic area is restricted to whatever the on-screen map currently displays, then the application crawls this spatial area to discover geographical entities. The application offers a simple filter builder interface to facilitate the creation of these filters to non-expert users.

3.3.2 Ontowiki Mobile

The Ontowiki framework [20] can be used to present, authors, and manage RDF data sets. The framework follows the View Model Controller (VMC) architecture, and through this it has been adapted to mobile devices, where the views are web pages designed for the smaller screens, which are rendered through the use of templates. Within the architecture, the model corresponds to the RDF store, however the Ontowiki framework uses the Erfurt API in order to allow the use of not just the Virtuoso based end points.

The mobile client application stores the accessed RDF data in its local storage to allow offline usage and to increase the performance. However as users are able to edit the information graph, synchronisation of the data and the resolution of conflicts needs to be handled,

[8] http://sindice.com/

in addition issues may arise when multiple users edit the same resource while offline. The application relies on SPARQL's update mechanism, and to ensure concurrency the system uses a version control system that has been inspired by Git.

3.3.3 mSpace mobile

mSpace mobile [44] is a semantic web application without any restrictions on the data sets that it might use in order to enable a web scale application. The application possesses an architecture that consists of three layers: the mobile application, the query service, and the triple store. The mobile application layer is responsible for visualising the data and interpreting the user actions. This module is connected to the mSpace Query Service which is responsible for translating user requests into semantic web queries. These queries are run both locally against the triple store and the specified semantic web end point in case of missing data.

The application in addition to enriching information from the semantic web uses local information such as the geographic position of the device or the context of the data to further filter the information that is being presented.

There are several areas where the application is lacking and future research is needed such as determining whether the trustability of the used data sources, and the issue of mobility.

3.3.4 Stevie

Stevie [6] is a linked data application that allows its users to view, edit, and create POIs. The application follows the client-server architecture and the server communicates with an RDF store. The client application is responsible only for the visualisation of the information and does only a small amount of processing locally, all concurrency control is done on the server side. The server offers a REST API to the clients. The clients submit instructions to the server through this API, which are then translated to SPARQL queries and executed on the triple store.

3.3.5 Qpedia

Qpedia is a linked data application that allows the use of DBpedia without the users having to know about the underlying technology. The application has two main components, the UI module and the query manager.

The UI module displays the information on the UI and communicates with the Qpedia server by sending queries and receiving results for the submitted queries. The query manager is responsible for creating the SPARQl queries based on the interaction by the user. The application instead of communicating with the live DBpedia server, which is supported by the Virtuoso software, talks to a Qpedia server which is a server built on top of Jena[9] and contains a version of the DBpedia data.

3.3.6 More!

The More! [34] application allows its users to gain more information about a researcher. The application works against the live Research.fm data store, which is a Linked Data end

[9] https://jena.apache.org/

point. The retrieved information is converted into objects by the mobile application based on a mapping prior to being used and displayed to users.

3.3.7 Who's Who

The Who's Who [13] application allows users to view publications of a selected researcher. The application follows the client-side architecture model, where the client communicates to a Jena based server. The application uses a specific data set and is not interlinked with other linked data sets.

The client application creates SPARQl queries locally then submits it to the server in case the information it wishes to work with has not been already accessed. The server runs the requested query and returns the results to the client application. The client application stores the results in memory and renders the user interface based on the available information. When a user clicks on an item, the application runs a SPARQl query locally to filter out irrelevant information that is being displayed.

3.3.8 wayOU

The wayOU [19] application allows users to gain information regarding other users. The application works against two specific SPARQl end points to access information regarding locations and places, while the other one is used to retrieve information regarding users. The mobile application combines the information from the two local data stores and creates the user interfaces based on this data.

3.3.9 Summary

The technical approaches of the presented Linked Data mobile applications were very similar to each other. All previous approaches have used server side rendering of the interface and visualisation to overcome the performance limitations of the mobile devices, however this affected the latency in the application. Client side rendering of the visualisation would reduce this latency and improve the application's usability.

Only the Ontowiki mobile application has demonstrated the possibility to use it offline. However as highlighted in Section 4.2.1, a constant connection to the servers cannot be guaranteed, hence offline usage would be beneficial for any LD mobile application. Of course, this introduces additional issues with respect to the version of the data that one is using and how to update the local copy. None of the presented applications, except wayOU, have demonstrated approaches which do not rely on a server between the Linked Data end point and the mobile application.

Ontowiki mobile was the most use-case independent application presented, however it is unable to dereference RDF resources, thus only the information that is presented on the interface can be used by users, which makes it more difficult to browse the LOD cloud and follow links. In addition, ontowiki mobile is not a geographical application, as in it is not able to present geographical information on a map.

Table 3.2 compares the presented mobile applications to each other with respect to the presented approaches in overcoming the challenges of mobile Linked Data highlighted in Section 4.2.1.

Table 3.2: A comparison of the different mobile linked data applications with respect to the Linked Data challenges on mobile

Common feature	DBpedia Mobile	Ontowiki Mobile	mSpace mobile	Stevie	Qpedia	More!	Who's Who	wayOU
Offline usage	no	yes	no	no	no	no	no	no
Server independent	no	no	no	no	no	yes	no	yes
General purpose	no	yes	yes	no	yes	no	no	no
Local filter	no	yes	no	no	no	no	yes	no
Any source	yes	yes	yes	no	no	no	no	no
SPARQL	yes	yes	yes	yes	yes	yes	yes	yes
Dereferencing	yes	no	yes	no	no	no	no	no

3.4 Usability Evaluation Methods

Usability is the ease of use of a software application [37, 18]. This reflects how easy it is to learn to use the application, and gives insight into the perceived efficiency of the application. Usability testing aims to investigate the usability property of applications in order to determine how suitable an application is given for its use-case [37]. The application is tested in a controlled environment, such as a laboratiry, with a set of participants who are asked to complete certain tasks using the tested application [37]. These experiments aim to assess whether the application is usable by the target user base, and does not aim to evaluate the users [37, 18].

Overall when measuring the usability of an application, there are three aspects of the usability that one wants to investigate:

- *Effectiveness:* How well can users complete the tasks they were given
- *Efficiency:* How difficult is it to complete the assigned tasks?
- *Satisfaction:* How did the users find the application? Were they satisfied with their experience? [18]

During the usability testing it is essential that participants complete tasks that reflect the real use-case of the application in order to gain representative results [18]. The findings from the results, which indicate issues, can be used to alter the application in order to increase its usability.

This section aims to introduce the basics of usability evaluation. The different metrics that are applicable to the usability testing of the proposed applications will be presented along with the possible usability questionnaire methods that give good insight into the usability and user experience (UX) of applications.

3.4.1 Usability Metrics

This subsection presents different quantitative and qualitative metrics that investigators of usability evaluation experiments may record.

Quantitative metrics

There are two main metrics that are collected in a usability: the time it takes for users to complete the tasks and the number of errors (e.g. navigating to a unwanted or unrelated part of the application) committed [37]. These two measured give the best insight into how easy it is to use the application, however there are additional quantitative measures such as the number of times a source of help has been consulted (i.e. documentation). The list of the main usability metrics, along with their purpose, is presented in Table 3.3.

Qualitative metrics

During usability testing, it is often the case thet the interaction with the articipants is monitored by an investigator who is facilitating and directing the experiment [37, 18]. This person may take noted throughout the experiment where different qualitative metrics can be recorded, such as any interesting actions the user took, or issues that they have encountered. The **Think-Aloud Protocol** [28, 18] is a method to gather such data by requiring participants to say aloud what they are doing and why they have done something. In addition they may voice concerns or suggestions, which hint may areas that should be investigated further.

An alternative approach to gain insight into the participant's thought process is use the **Co-discovery** technique [18]. This technique requires that two participants work together to complete the assignned tasks. This method feels more natural to people and may allow the investigator to gain more insight as participants may share more information with someone else than just saying it [18]. In addition, this method required users to explain their strategy to solve the task to the other person, allowing the investigator to gain the seeked information with just listening to the conversation of the two participants [18].

3.4.2 Usability Questionnaires

This subsection presents three different quantitative usability measurement techniques which use a questionnaire form to gather data on users' perception of the usability of the application and their opinions on their experience.

System Usability Scale

The System Usability Scale (SUS) measures the overall usability of the application [8]. The scale is independent of technology and can be applied to both software and hardware systems [8]. The questionnaire is made up of 10 statements, shown in Table 3.4, which aim to investigate the overall usability of an application [8]. Users rate each statements on a scale of 1 to 5. The score of the SUS, is calculated by substracting 1 from the odd responses (1, 3, 5, 7, 9), subtracting the value of even responses (2, 4, 6, 8, 10), then sum the results. This score (between 0 and 100) gives a high level overview of the usability of an application which may be used for simple comparison purposes. As the questionnaire returns but a single number, it is difficult to determine how usable individual features or properties of the application are [32, 37].

Table 3.3: Usability metrics and their purpose

	Metric	Purpose
M1	Task completion time	This metric can be used to compare the performance of users. It also gives insight into how long it takes users to get used to the interface or how well the interface communicated information.This may indicate how suitable the application is to a given task.
M2	Number of errors	This metric can indicate how well users understand the task or the information presented. Depending on the error, it can also indicate whether a certain UI element is natural to use or whether it is confusing to the user.
M3	Number of issues	This metric reflects the usability of the application, as the higher the number the more difficult the user has found it to use the application and understand the interface and the information presented.
M4	Type of issue	This gives insight into the issues with the application and can suggest future work or fixes that are needed
M5	Context of issue	This gives additional insight into an encountered issue and may help researchers understand how it has been encountered and how severe the issue is
M6	User comments	This qualitative information may give insight into several areas and may suggest future work possibilities. In addition it also helps understand researchers the quantitative data.
M7	Suggestions	Suggests areas of future work that may be more beneficial to the usability of the application
M8	Usability questionnaire	Quantitative measures of usability. The questionnaire makes it easy to assess and gain an approximate understanding of the usability of the application and how the different app features perform

Table 3.4: Statements of the System Usability Scale

	Statement (Question)
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to be able to use this system.
5	I found the various functions in this system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

Practical Heuristics for Usability Evaluation

The Practical Heuristics for Usability (PHU) questionnaire aims to serve as a low cost usability questionnaire which investigates key aspects of the application [35]. The questionnaire allows investigators to investigate how individual aspects of the application are performing with respect to usability. Unlike the SUS, the PHU is a set of guidelines to create a questionnaire and not a set of specific questions. Designers of the usability tests need to formulate appropriate questions for the provided area of focus. The list of areas is presented in Table 3.5.

Table 3.5: The Practical Heuristics for Usability Questionnaire

	Statement
	LEARNING
1	Help and Documentation
2	Adopt the User's Viewpoint
3	Simple and Natural Dialogue
4	Design for Advancement
	ADAPTING TO THE USER
5	Provide Maps and a Trail
6	Show the User What is (Not) Possible
7	Intuitive Mappings
8	Minimize Memory Load
9	Consistency in the System and to Standards
	FEEDBACK AND ERRORS
10	Feedback
11	Prevent Errors
12	Error Messages
13	Clearly Marked Exits and Error Recovery

After-Scenario Questionnaire

The After-Scenario Questionnaire [29] is a very short usability evaluation method, 3 questions, focusing on the perceived usability of an application. The questionnaire aims to assess user satisfaction after the interaction with an application has taken place. The questionnaire aims to gain insight users' perception of three different aspects of the investigated application's usability: how easy it was to complete the task, how long it took to complete the tasks, and how helpful was the application in completing the tasks. The questionnaire is presented in Table 3.6.

Table 3.6: Questions of the After-Scenario usability questionnaire

	Question
1	Overall, I am satisfied with the ease of completing this task.
2	Overall, I am satisfied with the amount of time it took to complete this task.
3	Overall, I am satisfied with the support information (on-line help, messages, documentation) when completing this task.

3.5 Summary

The State of the Art initially reviewed the visualisation of Linked Data. This review was mainly based on a survey by Dadzie et al. from 2011 [15]. A set of requirements for good Linked Data visualisations [Section 3.1.1] was estabilished, and 10 Linked Data aplications have been compared according to these requirements [Table 3.1].

In addition to the visualisation requirements, a set of Linked Data challenges on mobile devices was presented [Section 3.3], and the eight mobile Linked Data applications were compared with respect to these highlighted challenges [Table 3.2].

Methods to represent geographical information information such as the basic geo vocabulary [Section 3.2.1] and LinkedGeoData [Section 3.2.3] were presented.

Finally, the concept of usability evaluation and some methods, such as metrics collection and questionnaires, were presented [Section 3.4] to help readers understand how usability was measured in the evaluation of the *PVGeoVisualisation* applications.

Chapter 4

Design

This chapter presents the design of the desktop and mobile applications that have been created to investigate the differences in the usability of the two visualisation approaches. The desktop application has been designed to reflect the use case of map4RDF [27] and to act as a baseline for the experiments. The mobile application has been designed considering the requirements for the visualisation of Linked Data [Section 3.1.1] and the previous approaches presented which have proposed solutions for the limitations of mobile visualisations [Sections 3.1.2 and 3.3]. These two applications were used to investigate the research question of this dissertation.

Initially, the use case of the *PVGeoVisualisation* applications is presented along with a set of requirements that it would need to fit. The general approach of the visualisations will be presented along with how geographical information is visualised. Some challenges regarding the use of Linked Data will be highlighted and the design of the desktop application will be presented briefly prior to discussing the mobile application in detail.

4.1 Application Use-Case

It is desired that the *PVGeoVisualisation* applications are to be used for exploring the USPV data set. In addition, users of the applications should be able to:

U1: Browse the USPV data set by using a map

U2: Inspect individual USPV events and be able to access the details of the event without requiring to know RDF

U3: Be able to find one more more events by defining a limit for USPV event properties such as the category or the date

U4: Be able to filter USOV events without requiring to know or use SPARQL

4.2 Application Requirements

The requirements of mobile Linked Data applications [Sections 2.1 and 4.2.1] and Linked Data visualisations [Section 3.1.1] have been applied to the use-case of the *PVGeoVisualisation* applications and the basic user needs. The following are the requirements that the design of the two applications would need to satisfy.

REQ1: Visualise the USPV event instances on a map: Display markers for each event on the corresponding location on the map

REQ2: Access the USPV data through the Linked Data end point maintained by SCSS

REQ3: Use Linked Data technologies: SPARQL and RDF to manage the data

REQ4: Provide method to users to be able to limit the range of events to be visualised

REQ5: Provide a visual method to alter the SPARQL query used to generate the result

REQ6: Provide a method to select an event and display its properties

REQ7: Be usable by novice users, not just data set experts

REQ8: Hide the underlying technology: users should not need to write SPARQL queries or inspect raw RDF data

In addition to the above, the mobile application would need to also satisfy the requirements found below.

REQ9: Store and create visualisations locally without the need for a custom server

REQ10: Be able to work offline once the initial data has been loaded

REQ11: Provide the most up-to-date version of the data

REQ12: Follow the mobile UI design guidelines

The following sections describe the challenges in using Linked Data on mobile, then design of the *PVGeoVisualisation* applications in response to the above requirements and use-case.

4.2.1 Linked Data Challenges on Mobile

The nature of Linked Data that it is online and any application that wishes to use Linked Data resources have to be online in order to do so [5]. However mobile devices do not possess always on Internet connectivity [20]. To provide a good user experience, the application should be able to offer some functionality even in case of a connection loss.

This required the caching or saving of information onto the local storage of the devices. There are two possible ways, either caching data that has been received during use or downloading a copy of the data that is on a server. Both methods have its advantages and disadvantages. Downloading the full information provides full access, however may take up a significant disk space on the device. On the other side, caching information gives access to previously encountered data, in case of disconnection the amount of information usable is limited. A challenge is to find an optimum balance between these two approaches in order to provide good usability with respect to offline availability of functionality and performance [20].

In case information is being cached locally on the device, an issue arises with respect to the synchronisation of this information. Linked Data is constantly changing, hence an application will have to make sure that information that is stored locally is up to date and reflects that of the server. A challenge is to control the frequency and the method of updates. In case the application allows the editing of this information, the issue is further complicated as offline editing makes concurrency control much more difficult [20, 44].

Linked data is available through multiple end points. Is applications are to provide the best possible user experience they have to enable the use of any linked data source, hence information has to be merged from multiple sources. Applications doing so will have to make sure that entities that are equivalent are merged together, though labels such as owl:sameAs aim at easing the difficulty, there may still be implicit statements in the data that the application should be able to understand [12, 15].

In addition, Linked Data is heterogeneous in nature, meaning that data may be of different types. Generic Linked Data applications must be able to support the use of different types of information [15]. The visualisation of such information is a main challenge that has been highlighted previously in Section 3.1.

Most current linked data end points are slow. Applications have to be able to provide full functionality and avoid presenting users with long processing times.

4.3 Technical Approach

The application was using the United States political violence data set. The application focused on visualising the Event (pv:Event) objects in the data set on a map. These objects had several properties based on which the set of events could be altered using a visual query builder. The full structure of the Event class is presented in Figure 4.2. The data set was hosted by the School of Computer Science and Statistics on a Fuseki SPARQL server.

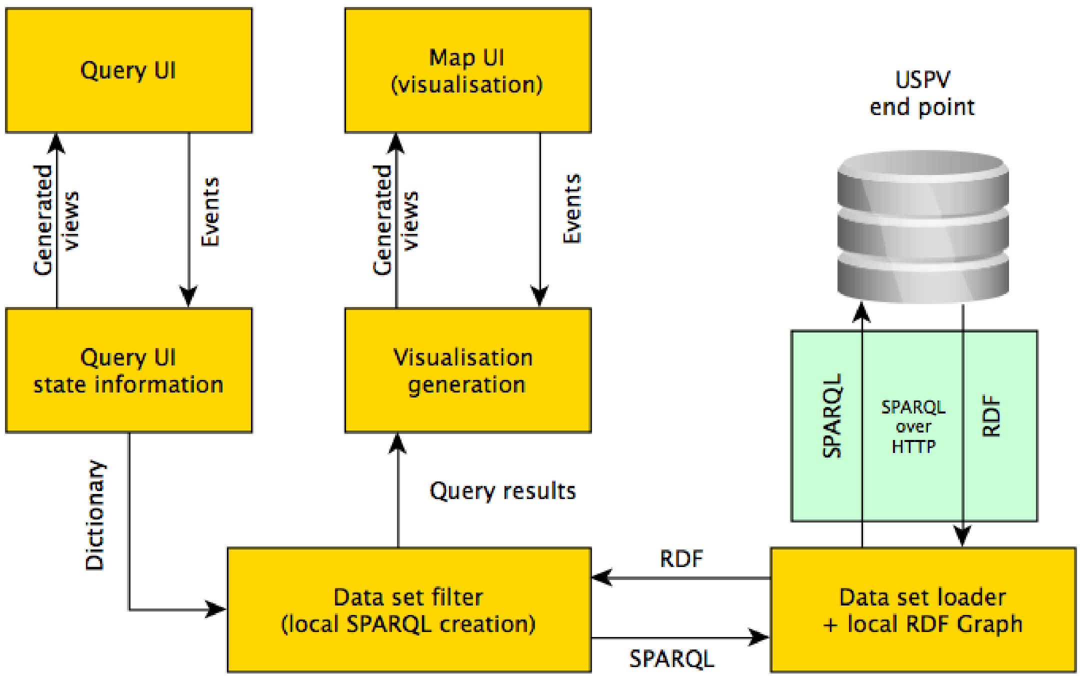


Figure 4.1: Overview of the general technical architecture and the application components

4.3.1 Data Retrieval and Storage

The applications were using the SPARQL over HTTP protocol to run a SPARQL query and receive a set of events satisfying the input query. The query that was being submitted depended on the set of filters that have been applied by the user. The desktop application through the use of the sgvizler library created a visualisation using the Google Charts API with the results received from the Fuseki server containing the USPV data.

The mobile application possessed a local RDF graph where it stored the results of each query, allowing for offline functionality. When a user was requesting a new visualisation, the mobile application ran the query against its local and the remote graphs and displayed the results of the local query on a Google map. Whenever the results of the remote query have been received, the local graph was updated, then the query ran again and the visualisation updated. This allowed for a faster display of information without waiting on the slow remote, however the the issue of outdated information was less noticeable as the visualisation was seamlessly updated upon refresh.

4.3.2 Data Structure

Additional information was added to the data structure to allow faster filtering on local RDF graphs by reducing the number of links queries needed to consider. Some information in events were stored in secondary level nodes, such as the fatality count or the date of the events, which was desired to be an attribute that one could alter in the query easily to limit the number of results displayed. It was discovered that the LinkedGeoData end points were very slow for real-time high throughput usage in looking up the coordinates for location names, hence the geographical information had to be encoded in the source data set. Without this change, the applications would need to potentially look up hundreds of location information every time a query is run, which would have been wasteful and slower.

The original data set has been modified so that the pv:Location nodes possessed the latitude and longitude coordinates of their unstructured location attribute. A Node.js[1] script was written which used the Google Geocoding API to look up the coordinates for the unstructured location field for each pv:Location node in the USPV data set. The Google service has been chosen for the purpose of this geocoding process as it is able to give the most accurate result for each entry and it is capable to handle typos and errors in the values, which was an issue in the data set, as these information were added by humans. There was the option to use another solution, such as looking up locations using LinkedGeoData, however the researcher would have had to handle the typos and other geographical errors himself. As the focus of this study was on the visualisation, the easier and faster solution was selected so that the rest of the study could follow on.

4.3.3 SPARQL Query Creation

All SPARQL queries that have been used for the querying of the information were based on the same template. The applications were using a moustache[2] templating engine to define the sought values in the template and create the resulting query text, which ten

[1] Node.js is a browser independent JavaScript runtime. `http://nodejs.org/`

[2] https://mustache.github.io/

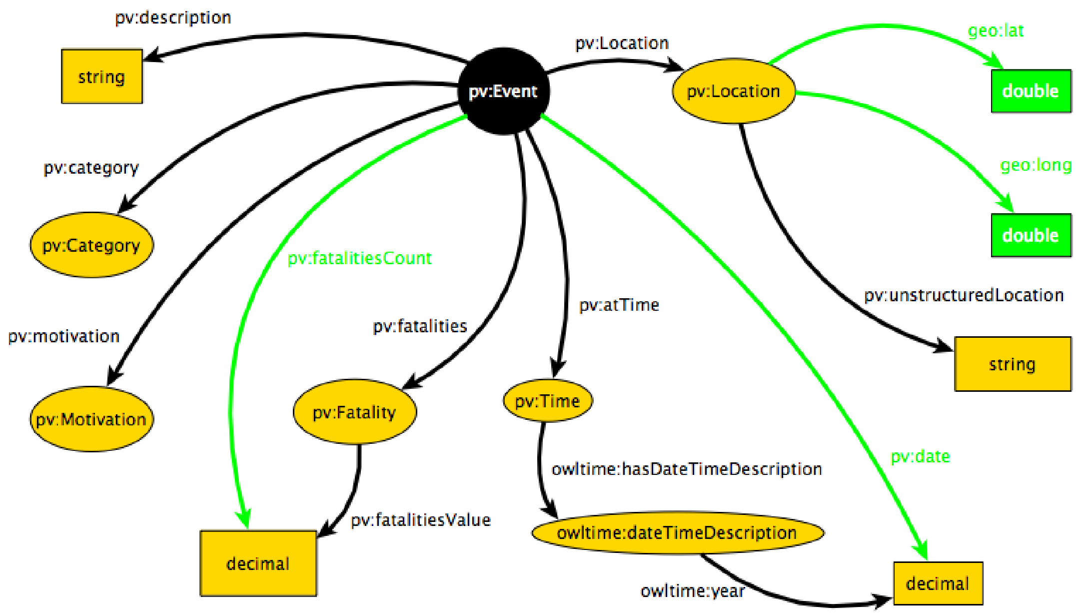


Figure 4.2: Structure of the visualised information. Black edges represent existing relationships while green edges and nodes represent the added information.

was submitted to the end point. Consider the following example: if you had the template "$\{\{value\}\} = property$" and your value that you are altering is currently *5*, then the result of the template and the value combined is "$5 = property$". Anything included in the input text by double curly braces represents a key in the input dictionary and is replaced by the corresponding value.

Users were able to limit the results based on the Category, Motivation, fatality count, the year, and the location of the events. The selected values or range of values were determined based on the state of the visual query interface, detailed in their respective sections, and a dictionary based on the UI state was created, which in turn was used to generate the SPARQL query. The details of the queries used in the desktop and mobile applications are presented in the corresponding sections.

Unfortunately, due to the limitations in the Fuseki platform, it was not possible to execute geographical queries on the remote using SPARQL and its extensions, hence each application has to use different methods to address the filtering based on the location. The applications allowed users to limit the presented events based on US states. The two approaches are presented in the corresponding sections.

4.4 Desktop Application

The desktop application has been designed to reflect the interface of map4RDF [Section 3.12], which allowed users to filter the presented information using a side bar. The information was presented on a map, where the objects of interest corresponded to markers. The information pane of the events reflected the design of map4RDF, however the information displayed was changes to be relevant to the data set, hence when a user clicked on a marker, the application showed the details of the selected event such as its description, date, category and motivation.

The desktop application used the sgvizler Linked Data visualisation library to run queries against the USPV end point and create a map based visualisation for the results. Samples of the desktop interface are presented in Figure B.1 in the Appendices.

4.5 Mobile Application

None of the mobile applications reviewed in the State of the Art, Chapter 3, have been designed in tablet devices in mind as they have focused on smaller screens, such as smartphones and PDAs, however the researcher believes that users would prefer to use a bigger screen to explore the USPV data set, as it would give a bigger overview of the visualised data and it would be more suitable for its use case.

This section considers the challenges and requirements of mobile Linked Data visualisation and how they were applied to the design of the iPad application while keeping in line with the design guidelines of the iOS platform [24]. A summary of the design guidelines can be found in Section 2.4, while the visualisation requirements of Linked Data applications is presented in Section 3.1.

4.5.1 Interface Design

The mobile application has been designed based on the desktop application while considering the limitations and issues of mobile highlighted in the State of the Art and Background chapters.

The application needs to have a visual query interface where users able to set how the USPV Events are filtered and which ones are displayed on the main map. The properties which users should be able to define are the Category, Motivation, Location, fatality count, and the year. The interface should present options according to the following:

- The Category and Motivation properties should be limited to a set of predefined options as according to the USPV ontology.This list should be shown to users from where they are able to select the options they want.

- The locations should be the list of US states displayed in a similar fashion to the Category and the Motivations.

- Users should be able to define a range for the fatality count and the date, or set individual values in case users are looking for specific dates.

The map based visualisation should reflect the desktop application however the information box on the desktop would be too crowded on mobile, and hence it should be designed in line with the iOS design guideline where the elements are clearly spaces and well legible. In addition while investigating the design with the paper prototypes it should be decided whether the query UI should be places on the same view as the map or if the two interfaces should be present on separate views.

4.5.1.1 Data access and storage

The application would have a local RDF store in order to allow for offline functionality and faster visualisations. To query this graph the SPARQL query language would need to be used. The process for creating the SPARQL query would be identical to the desktop application and the general overview was presented in Section 4.3.3. In order to visualise the information, the local graph would need to return a set of Event objects which can be achieved through a SPARQL SELECT query, which can return the results formatted in a table, where each row would correspond to an Event. By iterating through the rows, the information can be places on the map based on the corresponding *geo:lat* and *geo:long* information.

However if the application would query the remote end point using a SPARQL SELECT query, the application would have to convert the results to RDF in order to be able to store it locally, hence the need for a second query that is siilar to the local query in returning Event objects however it would have to in RDF format, which is possible using a SPARQL DESCRIBE query. In such a query, it is possible to return a custom graph, and the remote query was requesting the relevant information, as presented in Figure 4.2. This returned RDF graph graph was used to update the local graph.

Query UI

The query interface should have an entry for each *pv:Event* property that users would be able to restrict. These properties are the Category, the Motivation, the fatality count, the date (the year the event occurred in), and the location of the event (restricted to US states). For each property the list of allowed values should be presented in the interface, and if selected the corresponding view should be highlighted. Whenever a user alters an item, the interface should be updated to reflect the changes that have been made.

4.5.1.2 Paper prototype

This section discusses the paper based prototyping and design of different user interface elements.

Information box

The information callout box of events on the desktop application is not optimal for mobile as the text is not well well spaced and the control buttons are too small for touch. The iOS design guideline recommends the information and controls to be well spaced, so a vertical layout approach was taken by placing the elements in a one columned table. The initial prototype for the info box as well as the second version correcting the highlighted issues are presented in Figure 4.4.

Query UI

Two designs were created for the query UI. The initial approach reflected the desktop design where the map and the visual query builder were present on one screen. The query UI was a list view with all the entries on the left of the window and the map based visualisation on the right side. Unfortunately it seemed that the resulting query interface would not be

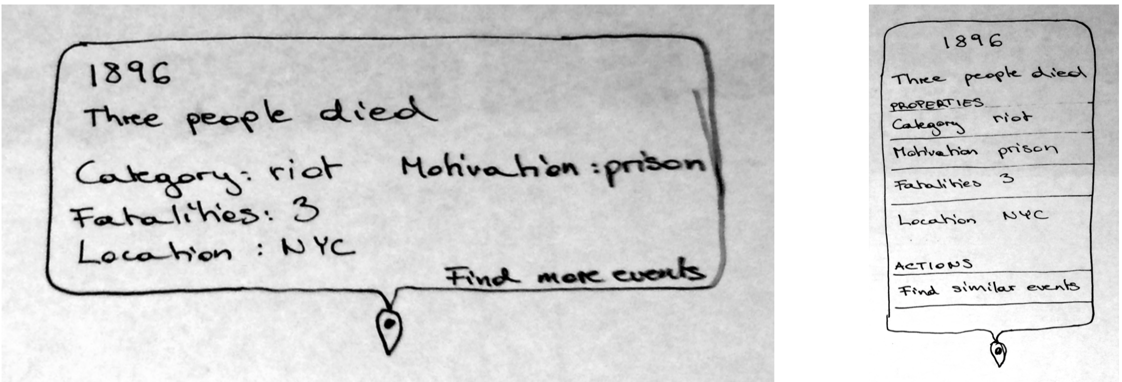


Figure 4.3: Paper prototypes of the information box. On the left the direct application from desktop application, on the right the new design.

able to present all options on the screen and users would need to scroll a significant amount to find the options they would be looking for. In addition due to the limited space the entries in the list of options were squashed together making it difficult to parse the information.

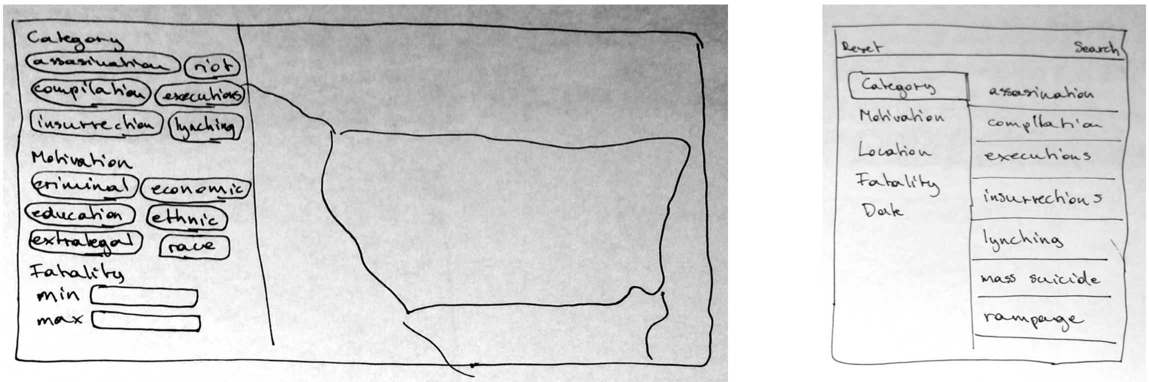


Figure 4.4: Paper prototypes of the information box. On the left the direct application from desktop application, on the right the new design.

The second approach had the query UI on different views, where the main view contained only the map on the entire screen. By having the query interface on a separate screen the elements could be made much bigger and more legible. Users would be able to bring up the query UI with a button press. To reduce the amount of scrolling required, the query UI split the screen into two columns. On the left the list of properties was presented, while on the right the list of options was shown for the currently selected property. Upon tapping on a property, the correct list of options would be presented. Two alternatives for the list of options was created, one with a vertical one columned table and the other with a tag cloud based approach where elements span the space horizontally and vertically aiming to occupy the least amount of space possible.

4.5.2 Implementation

This section details the implementation of the application detail and the challenges that have been encountered during the implementation. In addition any changes to the design of the application will be explained whenever it was required to solve an issue. An overview of the architecture of the mobile application is presented in Figure 4.5.

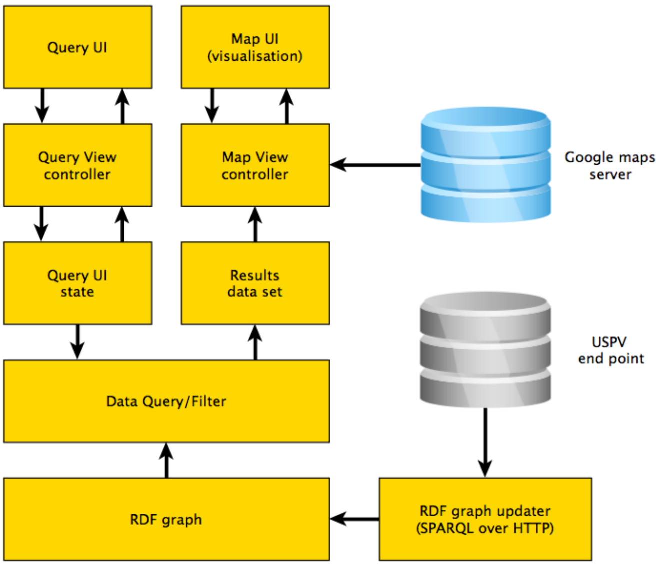


Figure 4.5: Overview of the architecture of *PVGeoVisualisation mobile*

4.5.2.1 Data access and storage

Initially, it was proposed that the application would update the local RDF graph for each visualisation, however due to performance issues it was decided that the application would run a general query against the remote LD end point whenever it was launched and updating the local graph only once per session. In addition, continuous update was not necessary, as the data in the end point is not live, as in it does not get updated every day. The resulting process looked the following:

- When the application launched, or when explicitly requested by the user, the application would run a SPARQL DESCRIBE query against the remote requesting all Event objects in the format that has been presented in Figure 4.2.
- When receiving the RDF graph, the application would UPDATE the local RDF graph.
- Whenever the query interface changes, the dictionary representing its state has been updated.
- Whenever a new visualisation was requested, a SPARQL SELECT query was created based on the query UI dictionary and ran against the local RDF graph. The map was cleared from markers and the Events in the results set were placed on it.

4.5.2.2 Interface

The application would use the View-Model-Controller approach for its architecture, as any standard iOS application would [24]. The two main interface elements would receive their own controllers and data models based on which the views are generated. The controllers were isolated entities which were not able to talk to each other and would only use the information that was present in their data structure to create the views on the interface. Therefore the interface that the users looked at were adapting to the information that was being presented.

The state of the interface was stored in a dictionary. Whenever interacted with the UI, the state dictionary was updated and regenerated based on the new information. A visual demonstration of this process is shown in Figure 4.6.

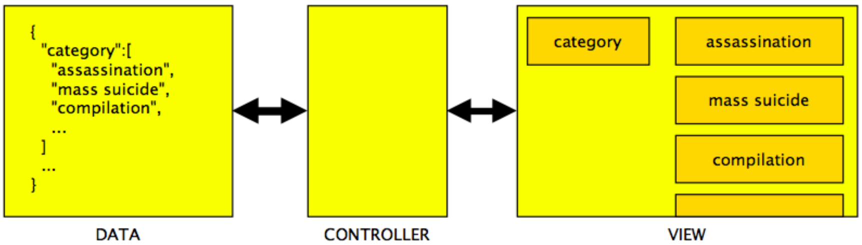


Figure 4.6: A visual presentation of the relationships between the view, data and controller of the query UI.

Map/Visualisation UI

The map interface used the same interface representation and creation method as the Query UI with some differences in UI layout. The data, instead of being a representation of the event properties and possible values, was the set of events that the local query has returned. Whenever the controller detected that the data was altered, it cleared the map and updated it with the new information. Whenever a user requested detailed information about an event (i.e. when a marker was tapped) the controller would create and present an information box with the details.

4.5.2.3 Prototype 1

The design of the initial prototype was based on the design of the second paper prototype presented in the previous section. The application had the query interface and the visualisation on separate screens, increasing the amount of screen estate available for both interfaces. However the separation of the two interfaces eliminated the real-time feedback in the visualisation whenever the user altered a property in the query UI. Instead users were required to explicitly switch between views through button presses.

The map

The main interface consisted entirely of a map. Events were presented on the map using red markers, the default options in the Google Maps iOS library[3]. The design of the information panel, otherwise known as callout, of the markers was based on the design of the paper prototype with the information well spaced out and well legible. The callout also allowed users to take some actions such as searching for similar events. A screenshot of the final callout design is shown in Figure 4.7. Callouts were shown to users whenever they tapped on an Event and were dismissed whenever they tapped outside of it.

[3]https://developers.google.com/maps/documentation/ios/

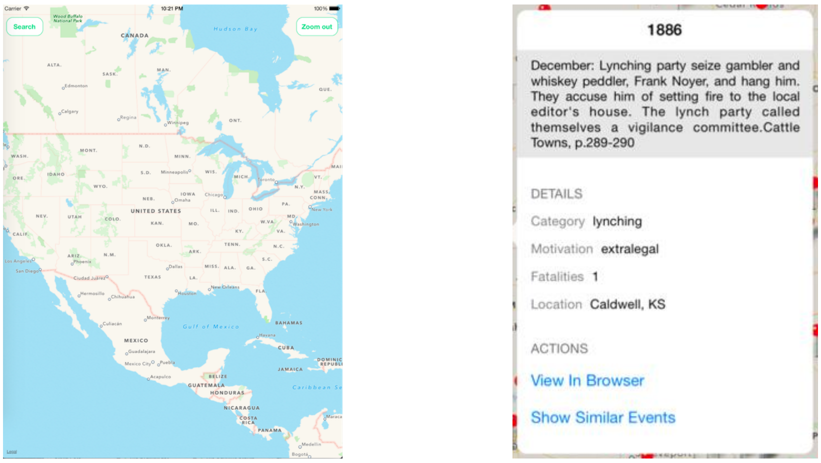


Figure 4.7: Screenshots of the map UI and the Event callout (Bigger versions of these images are found in Section B.2)

Query Building Interface

The query UI allowed users to set values for different properties based on which the Events in the USPV data set were filtered. The design of the interface followed the two column approach where on the left side, the list of properties was shown. Whenever a user tapped on one of the properties, the corresponding list of predefined values was shown in the right column. In case the property was numerical, the options to set the minimum, maximum, or other values was shown instead of the list.

There were two design for the options lists. The first design, called UI1, presented the values in a tag-cloud based approach, where the values were ordered in such a was so that the tag cloud takes up the least amount of space possible. The second design, UI2, was a vertical list based approach, where the entries were made to be well legible and were ordered in alphabetical order, however not all values were initially visible on the screen. The two interface design are presented in Figure 4.8.

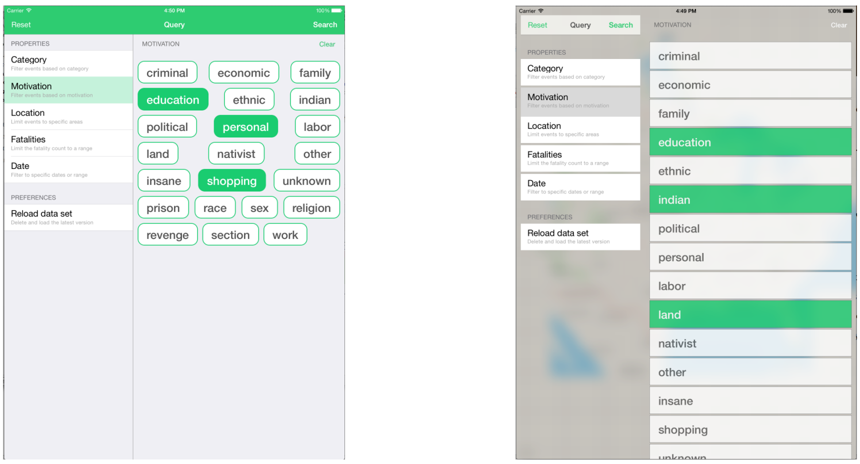


Figure 4.8: Screenshots of UI1 and UI2 (Bigger versions of these images are found in Section B.2)

There were some additional visual differences between the two interfaces, such as the semi transparent background or the positioning of the search and reset buttons.

The numerical selection panes were identical in both approaches, however instead of entering numerical values through a number pad or the keyboard, the use of a numerical spinner was considered as an experiment to investigate whether the input method would be welcomed by users. This is presented in Figure 4.9.

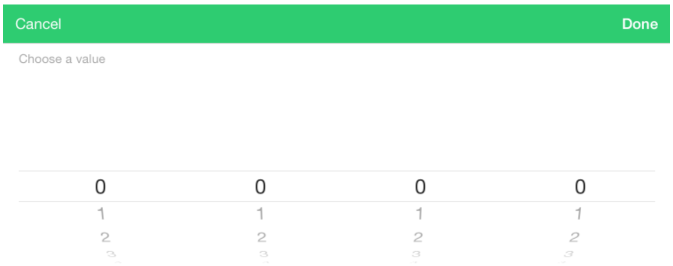


Figure 4.9: Screenshot of the number input spinner

To present the map and apply the currently query to the data set, the ”Search” button on the top of the query UI was needed to be pressed. Users had the option to clear the selection for each property individually or reset the entire query to its default state.

Chapter 5

Evaluation

This chapter discusses the evaluation of the mobile application for the visualisation of the USPV data set with a high focus on usability. The evaluation has been carried out iteratively in three stages, where a stage consisted of an initial development phase, followed by the experiment, and concluded by the evaluation of the findings.

5.1 Introduction

This dissertation involved the evaluation of a mobile geographical linked data visualisation application as it has evolved over time. The application was tested by several participants who have been recruited based on different criteria depending on the main focus of the specific experiment, which are described in the respective sections. The evaluation involved three stages, where the focus and areas of improvement depended on the findings of the previous experiment, or in case of the first experiment, the initial design (see the Design chapter). Based on the findings, changes have been added to the mobile application and evaluated in the following experiment. The final stage involved the comparison of the mobile application with the desktop application previously presented [Section 4.4].

Aiming to build on the design presented in the previous chapter and evaluate the application that has been built, this chapter is structured the following way: Section 5.2 presents the generic experimental design including the structure of an experiment and the tasks users had to complete. This is followed by Experiments 1, 2, and 3 in Sections 5.3, 5.4, and 5.5 respectively, which present each individual experiment and the findings in detail.

5.2 Generic Experimental Design

As shown by the State of the Art chapter [Chapter 3], there has not been as much research in mobile approaches to geographical linked data visualisation as on the desktop, hence prior to comparing the two applications together, ensuring that the most common issues with such a visualisation tool were not present in the application was important. In order to do so, an initial experiment was held to determine the issues users had with the mobile application, allowing for the design of changes potentially fixing the issues. A second experiment, prior to the final comparative one, served to verify whether these changes have resolved the listed issues.

Measuring the usability of the application throughout the experiments also allowed to highlight how changes in the application affected the users' perception of the presented information and their performance in completing the tasks. The measurement of usability also allowed to investigate how the application evolved over time as more features have been added and how suitable it has become for different user groups.

In addition each experiment had a secondary purpose, where different attributes of the mobile application were investigated. The initial experiment focused on investigating the difference between two interface designs, the second experiment looked at how different types of users use the mobile application, while the final experiment contrasted the mobile application with the desktop approach.

The focus of the experiments were determined based on the information at hand. By looking at the research questions and objectives, and at the current set of results, different investigation options and paths may have been apparent. The researchers have prepared several objectives and a small set of them, which were deemed the best suitable for the study, were selected for the next experiment.

All experiments follow the outlined scenario in subsection 5.2.1 with some minor changes if it was necessary, which are detailed in the respective Sections. Each experiment focused on comparing two different entities which corresponded to one factor being varied. The initial experiment compared tho UI designs, the second experiment compared two groups of participants, while the final experiment compared two applications. The aim was to try to keep all other variables static in order to be able to compare the varied factor better.

5.2.1 The Experiment Scenario

The participants were asked to execute a series of tasks on one or two applications, using a laptop or a tablet depending on the type of the application being examined. These tasks asked the users to determine the location of an event or to find a relationship within the data. Participants were encouraged to follow the think aloud protocol, where they sat out loud that they are thinking or explain why they have done a certain action. They were able take notes and record their answer on a worksheet [Section D.3] that was given to them at the beginning of the experiment, which presented the list of the tasks for the current experiment [Section C].

Once the participants have completed the assigned tasks, they were asked to fill out the Usability Evaluation Questionnaire [Section D.2]. If the experiment investigated the interaction with more than one application, the users were asked to complete the above process again but on the other device, otherwise the experiment carries onto the final interview stage where participants were asked about their experience. They were encouraged to comment on the application(s), usability, and the user experience.

During the experiment process, the investigator may have sometimes asked questions like "What are you thinking?" or "Why have you done that?" in case some clarification was needed to understand an action, or in case an interesting scenario has occurred. The investigator was recording the comments from the participants, as this information was used to gain further insight into the issues with the mobile application.

5.2.2 The Metrics

In addition to recording qualitative measures, such as the comments from the users, the investigator was recording other metrics. A set of common usability metrics have been described as part of the usability studies review in the state of the art and detailed in Table 3.3 [Section 3.4.1]. The experiments, in general, will be collecting these metrics, and any additional measures that have been collected in an experiment are detailed in the corresponding section.

5.2.3 The Tasks and the Experiment Worksheet

The tasks that the subjects have been asked to complete changed between investigation sessions, as to avoid the possibility of someone completing the same task twice. Some examples of the involved tasks are found below. The full list of the tasks used in the experiments can be found in the appendices [Chapter C].

- In the 20th century, were there more assassinations in the eastern or the western United States?
- The American Civil War was fought between 1861 and 1865. In which states where there fatalities due to the conflicts of this war?
- In the second half of the 19th century, in New Mexico, three people died due to lynching. Who were these people? In which city and in which year did they die?
- Following on from the previous question: Which neighbouring state had the highest fatality rate due to similar events?
- Take Northern Colorado in the 20th century. Where and when did the event with the highest death count occur? What was the cause of these deaths?

The tasks used in the experiments have been created based on a targeted exploration of the data set using the mobile application. Depending on the focus of the experiment, whether it was the query building experience or the exploration of the dataset, the investigator started to explore and build different queries, altering values and trying out several possibilities. Whenever an interesting result set has become apparent, a question was formulated around it. This allowed for the creation of tasks that would require similar steps from participants, allowing the investigator to observe the thought process of other users and identify issues that have not been apparent before the experiment. Some examples of such issues are the problem of illegible UI elements, unexpected or unnatural application behaviour, or the lack of feedback on actions.

There was a big focus on creating tasks that required users to use multiple features of the application. The answers could not be found by simply looking at the data. The aim of the tasks were to engage users with the geographic visualisation, requiring their perception of patterns and common knowledge to solve the tasks in addition to using the applications.

5.2.4 The Combined Usability Evaluation Questionnaire

Presented in Section 3.4, the usability of the application was measured to determine its suitability for the generic purposes of data exploration and task oriented use by different types of users. By measuring usability of the overall system, one can get a good insight on how well the application suits the user group, while measuring the usability of different components or areas of focus allow for the possible identification and determination of severity of problems.

The usability questionnaire is made up of three sections. Section one is an adaptation of the standard System Usability Scale, with 10 questions [Section 3.4]. This aims to determine a quantitative score for the usability of the application. The word system was replaced with the word application, while the word cumbersome in statement 8 was replaced with awkward as suggested by [3].

Section 2 is a questionnaire based on Perlmans Practical Usability Evaluation questionnaire [Section 3.4]. It aims to evaluate specific components and features of an application that are essential for good usability. The original questionnaire consists of 13 question, however the adapted version features less question, as any areas that have been subject previously in the questionnaire have been omitted. This can be achieved as the score obtained Practical Usability Evaluation does not depend on the structure of the questionnaire.

Section 3 is an adaptation of Lewis After-Scenario Questionnaire [Section 3.4], which aims to score the application's usability with respect to how well the users feel regarding their completion of the tasks that they have completed, and what their opinion is regarding how suitable the application is for the type of tasks that have been proposed.

The aggregated data obtained from the questionnaires allows the investigator to determine the usability of an application and help highlight areas of issue. These results were used to evaluate the design changes of the mobile application and to compare the mobile and desktop applications.

5.2.5 Analysis Methods

In the initial experiment, the values that were being collected were used to determine which of the two query interfaces is more preferable and how usable the overall application was. Some examples of these values were the amount of time a question has been completed in, the ratio of correct and incorrect answers, and the number of errors a participant has made per question. In the second experiment, these metrics were used to investigate the difference between novice and data set expert users. The third experiment involved the comparison of the mobile and desktop-based applications.

The usability scores were calculated from the questionnaire responses according to the method described in Section 5.2.4. Section 1 uses the standard SUS scoring system, while section 2 and 3 are a sum of the entry value (or the value subtracted from 5 in case of a negative question). The average of the scores can be used to compare the usability and effectiveness of the two interfaces. Higher score in the questionnaire means better usability, while a lower score in completion time and the number of errors committed implies a more effective approach.

The data can be analysed differently by splitting the data into multiple sets based on a criteria. One approach is to separate values based on a user property such as the experience level, or whether they have used the application beforehand. By investigating the data from different angles, additional findings have been discovered. Table 3.3 [Section 3.4] presents what different metrics aim to investigate and how they may be used in the analysis. Each metric gives a quantitative figure for the corresponding entity (as described in Table 3.3), which are then compared to the comments and actions of the users. Based on the value and the user feedback, the usability of a feature is indicated.

5.3 Experiment 1: Initial Usability Evaluation and the Comparison of Two Designs for Information Presentation

This section describes the first experiment, including its design and hypotheses. The collected data will be presented and analysed, and finally some changes to the mobile application will be suggested.

5.3.1 Design

The aim of this experiment was to determine the usability of two alternative designs for information layout in the query interface of the tablet application and to set a baseline set of usability scores for geographical visualisations and searches. As presented in Chapter 4, the application possessed a query interface, which was used to define what information is being presented on the main map.

The query interface allowed the selection of predefined values for Motivation, Category, and Location of Events. Two different layout designs were created, however it was difficult to determine which approach would be more appropriate or more usable by users, therefore the initial experiment focused on determining the difference between the two possible layout designs. In addition the investigator was observing the interaction of participants with the application and was looking for any issues that the participants may have encountered, may it have been due to the general design of the application or a specific components, to layout issues such as illegible text or non apparent buttons, etc. These additional measures allowed to determine areas of focus for the next iteration in the development of the application, such as the issues to correct or additional features required.

The first layout approach, UI1, showed the list of the predefined options in a collection view, which yielded a more compact presentation of the list of predefined terms, similar to a tag cloud. A tag cloud is a collection of words [22] spanning from left to right and from up to down. In UI1 it allowed users to see all possible values on the screen without having to scroll away. Unlike in most tag cloud based visualisation, the size and colour of the letters have not been altered based on the frequency or other property of the tags. However the tags were not in alphabetical order, instead they have been ordered so that the cloud takes up a minimal space.

The second approach, UI2, displayed a list of predefined values in a vertical table-like list with one column and multiple rows. Entries were presented in an alphabetical order. The tags were clearly visible with a large font and room between two rows in the table.

These two approaches are presented side by side in Figure 4.8. The general behaviour of the interfaces was similar: users were able to select one or more items in the list. The selected items were highlighted to indicate their state.

The participants were encouraged to follow the think aloud protocol to help the investigator to identify issues and common trends by users. Sometimes, the investigator may have asked questions such as "Why have you done that?" and "What are you thinking?" as a method of gaining more insight into a certain action or inaction. The investigator aimed to help the participant as little as possible. This was to help examine how a user would troubleshoot an issue without any help other than what the application had to offer.

5.3.2 Hypothesis

The cloud tag based design for information layout on the the visual query interface is more effective than a vertical list based design.

5.3.3 Scenario

In order to evaluate the application and its usability, participants of the experiments were be required to interact with the application. To control the interaction behaviour and the tasks that users executed, they were required to complete a set of tasks using either UI1 or UI2. These tasks were data set search tasks, where one had to find one or multiple events using only the application presented on the tablet. The set of tasks were be common for all participants and were presented on an experiment work sheet, where participants were able to record their answers and drop down any rough work that they needed. The tasks used in the experiment are presented below.

T1: In a prison riot in 1959, 2 inmates, 3 guards, and the deputy warden have been killed. Where did this event occur?

T2: How many people have been killed due to events with religious motivation in California after 1990?

T3: Who was assassinated in the second part of the 20th century around the Washington Metropolitan Area? (The Washington Metropolitan Area consists of the District of Columbia and the nearby cities from the neighbouring states.)

Each question has been designed to serve a specific purpose and to help investigate different areas of the application. T1 required users to get familiar with the application and how the visual query interface works. In addition, they are required to interact with the events of the results set by tapping on the markers on the map and inspecting the details of each event to find the right answer. T2 investigates how users are able to process information which have multiple results, and how well they are able to create a summary of the results. T3 investigates geographical reasoning using the application.

Participants were encouraged to record their answers on the experiment work sheet. Upon completion of the three tasks, the participants are asked to fill out the usability questionnaire, which has been described in Section 5.2.4.

5.3.4 Analysis Methods

To determine the difference between UI1 and UI2, the analysis method outlines in Section 5.2.5 has been used.

5.3.5 Types of Users

Participants were either friends or relatives of the investigator. They were recruited through social media or through direct contact. In total there were 10 participants, and all of them were unfamiliar with the data set. 5 of them had previous experience with similar mobile applications.

5.3.6 Experimental Methodology

This phase of the research presented each participant with the same experimental setup. Each participant has received the same set of tasks and questionnaire to complete. The experiment split the participants into two groups, where the first group worked with the application using UI1 while the second group was working with UI2. No other conditions was varied in the experiments apart from the design of the application's interface. All instances of the experiment took place in a quiet distraction free environment to reduce the interference by and distraction from outside factors.

The experiment consisted of two stages. In the initial stage the participants were asked to use the tablet application to complete the three tasks of the experiment in order. The time it took for the participants to complete each task was measured from the time the tablet is given to the user or has completed the previous question. Participants were asked to follow the think-aloud protocol helping the investigator to identify potential issues and common trends. The investigator was recording any errors the participant made, areas of confusion, or issues. Help was only be provided to participants if they could not proceed with a task without outside interference.

The second stage of the experiment consisted of a questionnaire, where the participants were asked to rate different aspects of the application on a scale of 5. The scored answers from this questionnaire was used to compare the usability and effectiveness of the two approaches. Based on the acquired results, the difference between the two designs, the set of lacking features, and the areas of issues was determined.

To gain even more insight into the user's experience with the application, the investigator interviewed the participants regarding their experience with the application, asking about any frustrations or problems they have encountered.

5.3.7 Raw Data

The raw data collected in the experiment is presented in Tables 5.1 and 5.2.

5.3.8 Analysis of Data

This section discusses in detail how the two different designs have performed with respect to usability and users' performance.

Table 5.1: Experiment 1: Usability Questionnaire raw data

Participant number	Tech familiar	SUS scores per question										PHU									After-Scenario			
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q1	Q2	Q3	Q4
UI1																								
1	no	3	3	3	3	4	1	5	1	3	3	5	3	5	5	5	3	5	4	2	4	2	3	5
3	no	4	1	4	2	4	1	5	1	3	2	4	3	5	4	5	5	5	3	4	4	4	4	5
5	yes	4	1	4	1	4	2	5	1	4	4	4	3	5	3	5	4	4	2	3	4	5	5	5
6	no	3	2	3	4	3	3	2	2	3	3	4	2	5	2	3	4	2	2	1	2	2	2	2
8	yes	3	2	4	1	4	2	2	3	4	1	4	4	5	3	5	4	5	2	2	5	4	4	3
UI2																								
2	no	3	2	4	1	4	1	5	1	4	1	5	5	4	4	4	5	4	2	2	4	4	4	4
4	yes	5	2	4	1	4	1	5	2	5	1	5	4	4	4	5	4	3	4	2	4	5	5	5
7	no	3	3	4	2	3	1	4	3	1	2	4	3	5	4	3	5	4	3	1	4	3	3	4
9	yes	5	2	5	1	5	1	4	1	4	2	5	4	5	5	5	5	5	2	4	5	4	5	5
10	yes	4	2	5	2	5	1	4	1	4	2	3	4	5	4	5	5	5	3	2	4	5	4	5

Table 5.2: Experiment 1: User performance raw data

Participant number	Tech familiar	Completion time (seconds)			Correct answers			Errors		
		Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
UI1										
1	no	285	210	220	1	0	1	0	4	1
3	no	495	690	570	1	0	1	2	1	1
5	yes	86	69	112	1	0	1	2	3	1
6	no	142	371	201	1	1	1	2	2	2
8	yes	200	260	170	1	1	1	3	4	2
UI2										
2	no	310	148	348	0	0	1	1	3	2
4	yes	300	675	588	1	1	1	1	1	3
7	no	87	179	251	1	1	1	1	2	2
9	yes	93	300	109	1	1	1	1	2	1
10	yes	118	436	96	1	1	1	0	1	0

5.3.8.1 Usability of the Application

The scoring method of the SUS is presented in Section 5.2.4, based on which the results of this experiment, shown in Table 5.3, have been obtained.

Table 5.3: Average usability scores for UI1 and UI2

	SUS	Practical Usability	After Scenario
UI1	70	33.2	14.8
UI2	82	35	17.2
Familiar	83	35.4	18.2
Not familiar	69	32.8	13.8
Overall	76	31.1	16

F1: UI2 has achieved a higher usability score than UI1.

The difference shown by Finding 1 (F1) was not statistically significant as shown by the results of the paired t-test [Table 5.4]. The t value was lower than the required when consulting a t-table regardless of the selected p value. The difference in usability between users based on skill level was statistically insignificant as well. This was probably due to the low sample size of 10 participants altogether.

Table 5.4: Average usability scores for UI1 and UI2

	t value	DF	p-value
UI1 vs UI2	-1.2418	4	0.2821
Familiar vs non-familiar users	1.5739	4	0.1906

F2: Users, especially the ones with technical skills, felt more confident using the application

The Practical Usability questionnaire demonstrated F2. User satisfaction with respect their performance in completing the tasks also showed that users were 12% more confident in their responses while using UI2. In addition to investigating how the two interfaces compare, another source of guide for further development is the investigation how the application suits people with different skill levels overall. Users who are familiar with similar applications have found the tested application to be well usable, unlike others. This difference is shown in Table 5.3.

Considering Jeff Sauro's interpretation of SUS scores[1], where he assigns a letter grade to SUS scores based on a survey of more than 600 usability studies. The SUS score from UI1 corresponded to a mean value according to the normal distribution of the SUS scores from the study. UI2 would have received an A for usability, hinting that the usability of the design was exceptional.

F3: UI1 scored a C in overall usability

[1] http://www.measuringusability.com/sus.php

F4: UI2 scored an A in overall usability

In addition, the Practical Usability questionnaire was broken down into three sections. When investigating the scores in the individual sections, one was able to see that in both designs, the application received low scores in the "Feedback ad Errors" section. The data shows that participants were missing some feedback from the application in case they have made an error which may be a possible cause for the decrease in usability and users' performance. It may have been possible that the addition of such feedback would resolve such issues.

F5: The lack of feedback on actions is the most probable cause for a drop in usability

As with UI1, the low usability scores indicate that minimal feedback to users make the application difficult to use by people who have no previous experience whatsoever. In order to boost the usability of the application, this area should be the main focus of future development.

5.3.8.2 User Performance

F6: Participants completed tasks much quicker using UI2

Participants who used UI1 had issues initially in finding the category and motivation terms they were looking for. This is reflected in the completion time of T1, shown in Table 5.5, which took significantly shorter for participants in UI2. The alphabetical list based view allowed users to more quickly identify terms by scanning the list also allowing them to anticipate the location of the item in the list. This was not possible in the tag based view, as items were not ordered alphabetically and the spacing between tags were uneven as the ordering was optimised for minimal space usage in the interface. This difference is even more noticeable in users who are not familiar with similar applications, as hinted by table 5.5.

Table 5.5: Comparing average completion time (in seconds)

	Task 1	**Task 2**	**Task 3**	**Average**
UI1 - Overall	241.6	320	254.6	272.07
UI2 - Overall	181.6	347.6	378.4	269.2
UI1 - No experience users	307.33	423.66	330.78	353.78
UI2 - No experience users	198.5	163.5	299.5	220.5
Overall	211.6	333.8	316.5	270.635

F7: Not having all elements presented on one screen initially does not affect performance significantly

Finding 7 is backed up by the completion time of T2 and T3, which only differ by a maximum of 8%.

Most errors in the use of the application were caused by participants not selecting multiple options in the query interface, or they were setting values for the category, which was

Table 5.6: Average number of correct answers and errors

	Correct answers	**Errors**
UI1	2.4	2
UI2	2.6	1.4
Novice users	2.4	2
Expert users	2.6	1.4
Overall	2.5	1.7

not required. Participants in both group were able to complete the tasks with similar results. Some users have had issues trying to find some category or motivation values in UI1's collection view, which is somewhat reflected in the error rates shown in Table 5.6.

As a tablet application would mostly benefit novice users, it may be more preferable to use a list based view representation as this user base would not use such application for an extended period of time.

5.3.8.3 Common and Reoccurring Trends

During the experiment, the investigator was taking notes on the behaviour and actions of participants, as well as noting any notable thoughts or reasons for actions that the participants would say. These are presented in Table 5.7 below as well as their occurrence ratio. Given these trends, one is able to assign importance to issues and order them accordingly. The more occurring an issue is, the higher its privilege would be, as it is probably one of the more pressing factors that affect usability of the application and users' performance.

Table 5.7: Percentage of users where the stated trend or issue presented itself

	Trend or Issue	**UI1**	**UI2**
1	Wrong value is selected using the number picker	80%	60%
2	Confused how to select min and max	100%	100%
3	Did not notice reset button	80%	60%
4	Multiple value selection not apparent	100%	40%
5	Optional nature of properties no apparent	60%	40%
6	Optional min/max selection not apparent	60%	40%
7	Unsure how to select neighbouring states	100%	100%
8	Unsure if query is correct	60%	60%
9	Error in query not apparent	100%	40%
10	0 results vs error in query not apparent	100%	60%
11	Query UI navigation bar confusing	60%	0%
12	Tapping on pin for callout not apparent	40%	20%
13	Does not use map as a help	100%	80%
14	Clears sections individually	60%	60%
15	Explores using the map	60%	60%
16	Forgot what values were selected if any	100%	100%
17	Got lost in query UI when trying to find an error	100%	100%

F8: Users were confused regarding the selection of numerical values

The interface allowed users to set a minimum or a maximum, or limit the selection to specific values. This interface is shown in the first image of Figure 5.1. Participants did not understand the difference between the different options and all users have made mistakes initially in defining these conditions. This issue has to be addressed with a change in the methodology.

F9: 70% of the users initially had issues with the number input method

The interface uses a set of number spinners, the second image of Figure 5.1, which required users to rotate the four spinners to set a value. Users were frustrated with the initial mistakes in selecting the values and would noticeably prefer to enter a value using the keyboard or a numpad.

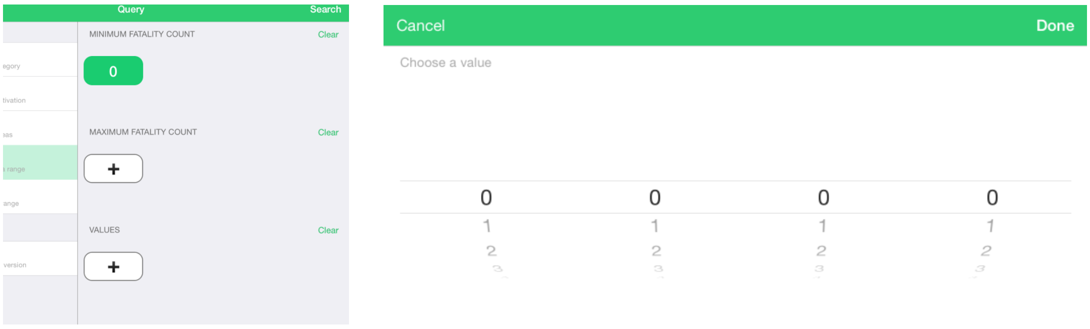


Figure 5.1: Screenshots of the number value selection and the number spinner.

F10: The ability to select multiple or no values for the Category, Motivation, and Location was not apparent to participants

Finding 10 was a major source of errors in the application. The multiple option behaviour is shown in Figure 4.8. This finding is reflected by issues 4, 5, 16, and 17 in Table 5.7. In case one did not select any values for a property such as Category, the application would have considered all options. The multiple option behaviour only became apparent to users when they have accidentally clicked on a second value while one was already selected. It was even less apparent in the tag cloud based approach of UI1. This issue may be addressed with feedback on users' action and possibly by the addition of some tooltips or instructions.

F11: Participants did not realise when there was an error in the query until much later

F12: Participants forgot the values they selected while constructing a query

It was not apparent to users, in general, that they have made an error in the creation of a query. They have noticed the errors late, and when they tried to correct their mistakes, it was difficult to identify the source of the error. In addition the fact that users forgot their selection did not help the resolution of issues. providing instant feedback on users' actions on the results set would possibly avoid the issues and the amount of time that has been wasted in trying to resolve an error. Additional feedback may be provided in helping users remember their selection in the query interface, avoiding the requirement to double check each query element prior to viewing the results on the map. Users should not be presented with the full information regarding the results for the current selection as it would cause information overload, however feedback regarding the number of elements that match the current query would benefit them and give insight into how well the query fits the information they are looking for. In addition, whenever a mistake is made in the query, users would be able to immediately notice their mistake.

F13: Participants were confused regarding the navigation of the application.

While users were able to locate the search button on the navigation bar in UI1, they did not notice the reset button of the query on the other side of the screen. On the other hand, this issue was not present in UI2, however users were having troubles locating the navigation bar. In both cases, the buttons and user elements need to be more noticeable.

F14: Participants did not use the map as a help for the construction of their query

F15: It was difficult to users not familiar with US geography to complete the tasks

The aim of the application is to get users to explore the underlying data set visualised on the main map, and allow they to execute more complex geographical queries, however the participants were not using the map as a helping tool. In addition 80% of the users were not familiar with US geography. They had significant issues in trying to identify neighbouring states. The use of the map as a helper was not apparent to users mostly due to the change in views, causing the map to be unreachable from the query interface during query construction. Most users, especially the ones unfamiliar with the geography, would possibly benefit significantly if the geographical aspects of the query construction would be done through the use of a map.

5.3.8.4 Comments by Users

Participants were shortly interviewed at the end of the experiment to gain some insight into their experience with the application and get their opinions, additional information that could greatly benefit the development of the application and the next steps in the research. The most common of the comments that have been mentioned are presented in Table 5.8.

Table 5.8: Percentage of users where the stated comment has been mentioned

Comment	**UI1**	**UI2**
List entries should be alphabetised	40%	0%
Value picker initially confusing	60%	40%
Pins on the edge of the screen not visible	60%	40%
Selecting neighbouring states is difficult	80%	60%
Search button not noticeable	20%	0%
Help or tooltips would be beneficial	40%	40%
Nice and simple UI	60%	60%

Most users complained regarding the difficulty in finding and identifying geographical locations such as cities or states on the map and felt that they have lost a significant amount of time finding the location they were looking for. Strengthening the need to address the issues of Findings 14 and 15. In addition they have confirmed the observed confusion regarding the number selection and the value picker (Findings 8 and 9) and have expressed a preference for textual input instead.

F16: Presence of off screen markers was not apparent

Participants have said that they sometimes have not realised the presence of off screen markers. This is a confirmation of the phenomenon from previous research, described in Sections 2.4.1 and 2.3.3, and shows that feedback for off screen information is still required on modern mobile platforms. A phenomenon that may be addressed with the highlight of the currently selected areas, this would possibly resolve other issues such as the search for states and avoiding the loss of the location of interest when browsing around the map, allowing users to quickly return.

5.3.9 Conclusion

The analysis of the data from the experiment shows that the states hypothesis is not true and that users preferred the list based interface. Overall, UI2 has performed better in all measures, however there are elements in UI1 that have been preferred better over the other version. UI1 and UI2 are shown in Figure 4.8.

Based on the findings, it was possible to identify areas of focus that needed attention and rework in the application. The design changes and actions, A[1-6], that have been taken in the second development stage after this experiment are summarised below.

A1*: Used UI2 as the baseline for the second prototype*

Based on the findings of the experiment, namely Findings 1, 4, 6, and 7, it was determined that UI2 was more suitable for the baseline for the second prototype, hence the work described below has been done on UI2.

A2*: Added visual feedback and cues*

Many of the findings showed that there was a lack of visual feedback in the application which caused a drop in usability and user satisfaction. This drop is caused by issues such as the unperceived ability to select multiple values [F10], forgetting previous choices [F12], or the inability to identify an error in the query [F11].

In order to address F10 and F12, a selection counter has been added to each property, which indicated how many items, if any, have been selected in the query interface. This hopefully gives hints to users regarding the possibility to select multiple values, in addition it may serve as a memory aid helping users remember their choices faster. These changes are shown in Figure 5.2.

A results counter has been added to both the query and the map views, which show to users how many events satisfy the current selection. This number was altered whenever a change in the query has been done, hinting at the effect of the users' action. If users monitor

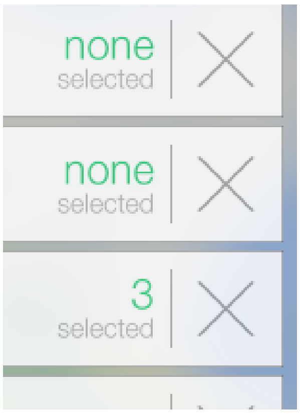


Figure 5.2: Screenshots of the property indicators along with the modified clear buttons

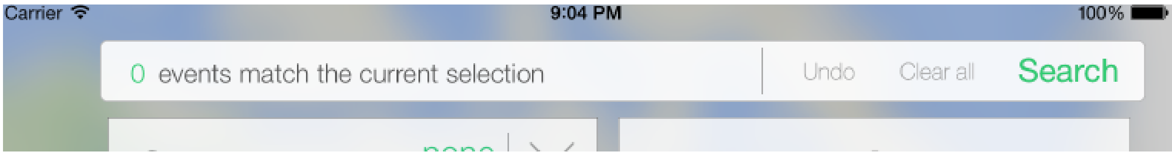


Figure 5.3: Screenshot of the results counter.

this value, they would be able to recognise which action have undesirable results. The results counter is shown in Figure 5.3.

The addition of these visual cues have the potential of improving usability, as highlighted by F5.

***A3**: UI elements have been made more apparent*

Many elements in the application were not apparent enough and confused users in the navigation of the application and the understanding of the presented information. To address this issue, many UI elements have received more legible text and may have been moved to more noticeable areas, and in some cases they have been made larger.

In about 50% of the cases, the clear button in the query UI was not apparent to participants, hence it was moved to a more used location on the screen (next to the property selection) and has been made larger, as shown in Figure 5.2.

***A4**: Reworked numerical number selection*

Findings 8 and 9 shows that users were confused regarding the selection of numerical properties. During the design of this feature, it was desired that one would be able to define a range of values, however the possibility to limit the numerical values to a non continuous set of numbers was also desired, hence the options to be able to select a minimum, maximum, and multiple individual values.

To address this issue, while keeping the flexibility in the selection of the values, the feature has been redesigned and the problem approached from a different angle. Instead of defining values, one would define conditions for the property, such as "Date is greater than 1990" or "Fatalities equal to 10". This yields a common interface for the selection of any value, while keeping the flexibility for defining any condition for the property. This selection interface is shown in Figure 5.4.

Figure 5.4: Screenshots of the condition list and the condition picker views.

A5: Improved map functionality

Finding 15 shows that users had issues with the geographical aspects of the tasks, especially since most of them were not familiar with US geography. Participants have not used the map as a helping tool as desired by the investigator [F14], suggesting that this functionality was not apparent; in addition the map may also be cumbersome to use in such a way.

To address this issue and to encourage people to use the map during query construction, a smaller map has been added to the query UI, where the states have been drawn onto the map. The selected states have been highlighted like in the list view. Tapping on a polygon selects the corresponding state, allowing users to select location using only the map. This functionality would hopefully resolve issues around the selection of areas such as neighbouring states.

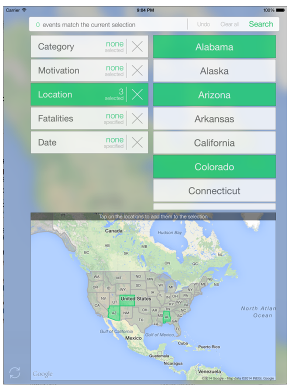


Figure 5.5: Screenshots of the mini map on the query interface (otherwise known as the "query map").

In addition, the area highlight feature has been added to the main map hoping to help users identify what are the areas of interest and where markers should be expected to be found.

A6: *Added map overview*

An issue in the first prototype was that users sometimes did not realise the presence of off-screen markers [F16]. To address this issue and to show users where markers are and how many there are an overview of the map has been added. Whenever a user displays the main map interface, it shows an overview of the data on the map, fitting all markers onto the screen. In addition, the previously presented area highlight would hopefully allow users to contain their exploration to an area of interest.

5.4 Experiment 2: Evaluating Usability of the Mobile Application Between Expert and Novice Users

This section describes the experiment that has been carried out in order to evaluate the changes, stated in Section 5.3.9, that have been applied to the mobile application in order to create the second prototype (P2). The experimental design, hypotheses, and the analysis of results along with the set of findings and actions to be taken are described in the below sections.

In addition, the difference between experts and novice users was investigated. Experts are users who who are familiar with the data set, it's structure and what information contained in the data set. Novice users are the ones who do not possess this knowledge and may have only little experience with the data set.

5.4.1 Design

The aim of this experiment was to evaluate the improvements that have been made since first experiment and to determine the how well the application suited both novice and expert users. These changes are outlined in the list below and are visually presented in Figure 5.6, however a detailed explanation is found in Section 5.3.9.

C1: Both the query interface and the map view have received indicators regarding the amount of events that satisfy the current query.

C2: More prominent 'Search' button and the relocation of the other query control buttons such as the 'Clear All', renamed from 'Reset', and the 'Undo' button.

C3: Indicators for each Event property, showing users how many values have been selected. In addition, the individual clear button for each property has been moved next to the indicator for easier access and discovery and the text replaced with an 'X' icon.

C4: Addition of a map to the query UI, presenting all US states with a polygon overlay. Tapping on a polygon will select the respective state in the list and vice versa.

C5: Indication of the selected states on the map map by a polygon overlay.

C6: Rework of the numerical value selection. Instead of selecting the minimum, maximum, or an individual value, users will be required to define certain conditions on the property, such as whether *the date is greater than x*. The number picker interface has been replaced with a condition editor.

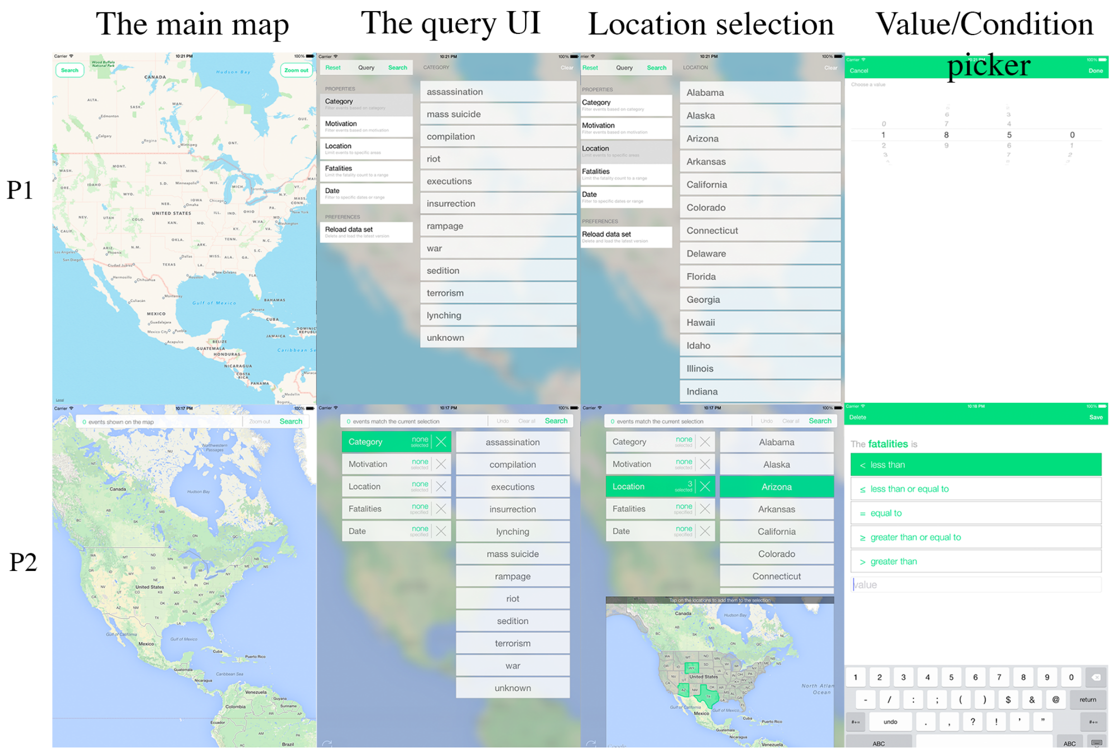


Figure 5.6: Visual demonstration of the changes applied to UI2 (winner design of P1) and the difference between UI2 and P2 (used in Experiment 2)

To investigate the difference between the two user types, the experiment separated the participants based on their level of familiarity with the data set. Similarly to the Experiment 1, participants were be asked to complete three tasks using the application, however everyone received the same version of the application. These two sets of participants were hopefully representing data experts and novice users, and reflecting the average skill levels of such users.

As in the previous experiment, the participants were asked to complete a set of tasks [Table 5.9]. Each question has been designed to serve a specific purpose and to help investigate different areas of the application or to explore possible use cases that have not presented themselves in the previous experiment.

5.4.2 Hypothesis

There should be an improvement in usability, compared to the usability baseline created in Experiment 1, given the addition of visual feedback to the application.

5.4.3 Scenario

The scenario outlined in Section 5.2.1 will be followed in this experiment.

Table 5.9: Task list of Experiment 2 and their purpose

	Task	Purpose
T1	Find all 20th century events that have occurred in a coastal city of California. How many such events were there? And how many of them had a political motivation?	Use of geographical reasoning to limit the results set to a coastal area on the map. This query is not possible to achieve without spacial reasoning and it is a difficult task for a computer, as the definition of a coastal depends on the person. In addition, participants are required to interact with the events on the map to find the right answer.
T2	The Kansas City Metropolitan Area includes two the Kansas Cities, one from the state of Kansas and the other from a neighbouring state, as well as the nearby cities. Which motivation for these events was the most common?	Investigates how users are determining what is the area of focus of the task and how they are modifying the query according to their discovery, and how they discover geographical information.
T3	Route 65 is a north-to-south US highway from Gary, Indiana to Mobile, Alabama. Count the number of fatalities along the route which have occurred either in the 19th or the 20th century.	Investigates how the application can be used for exploration and how suitable it is to determine summary information without the use of built in functions.

5.4.4 Analysis Methods

The experiment uses the general analysis method outlined in Section 5.2.5, which has been also used in previous experiment (Experiment 1). The analytical process will remain the same.

5.4.5 Types of Users

There were two groups of participants, with 9 people overall. The novice users (4 participants), those who have no knowledge of the structure of the data set, did not have experience with similar applications in the past. The expert users (5 participants), the ones who are aware of the sort of information that is contained in the data set and know its structure, all had previous experience with familiar apps.

The lack of novice users with no previous experience of the application may have had a significant effect on the completion time and the number of mistakes users made. However this issue may not have been present, as the interface has changed since the previous experiment, and the participants may not have gained all the benefits of familiarity with the application from the previous experience. However, during analysis, this unfair difference between the two user groups was considered.

Table 5.10: Experiment 2: Usability Questionnaire raw data

Participant number	Tech familiar	SUS scores per question										PHU									After-Scenario			
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q1	Q2	Q3	Q4
Previous user																								
1	no	4	2	4	3	4	2	5	1	4	2	5	4	5	5	5	5	5	2	5	5	3	3	5
2	no	5	3	4	2	4	1	4	2	3	3	4	3	5	4	3	4	4	3	3	3	3	4	4
3	yes	3	2	4	2	4	3	3	1	3	3	4	3	5	3	4	4	3	4	2	3	2	2	3
4	no	4	1	5	2	4	1	4	2	5	1	5	4	4	4	4	5	4	2	3	5	5	5	4
New user																								
5	no	4	1	4	1	5	1	5	1	5	1	3	4	4	5	5	5	4	1	3	5	5	5	5
6	yes	4	2	4	2	4	2	3	2	4	2	5	5	4	4	4	4	4	2	4	4	4	4	4
7	no	4	1	4	1	5	1	5	1	4	1	5	4	3	3	5	4	3	2	3	4	4	4	5
8	yes	3	1	5	1	5	1	5	1	4	1	5	5	3	3	5	4	5	2	1	5	5	4	5
9	yes	4	2	4	2	4	2	4	2	4	2	4	3	3	3	5	5	2	2	2	4	4	3	4

Table 5.11: Experiment 2: User performance raw data

Participant number	Tech familiar	Completion time (seconds)			Correct answers			Errors		
		Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
Previous user										
1	no	251	573	855	1	1	0	2	4	4
2	no	469	725	483	0	1	1	1	4	1
3	yes	634	343	640	1	0	1	2	1	3
4	no	74	180	137	1	1	1	1	2	1
New user										
5	no	258	178	199	1	1	0	0	1	2
6	yes	253	227	152	1	0	1	2	3	2
7	no	137	212	243	0	1	1	2	2	1
8	yes	86	125	202	1	1	1	1	0	0
9	yes	183	185	199	1	1	1	0	3	4

5.4.6 Experimental Methodology

Participants were presented with the same experimental setup. Each participant received the same set of tasks and questionnaire to complete. The experiment split the participants into two groups, according their expertise with the data set. No other conditions was varied in the experiments. Each participant completed the tasks in a quiet distraction free environment to reduce the interference outside factors.

The remaining of the experimental methodology is identical to the first experiment, which is described in Section 5.3.6.

5.4.7 Raw Data

The raw data collected in the experiment is presented in Tables 5.10 and 5.11.

5.4.8 Analysis of Data

This section discusses the data collected throughout the experiment with respect to the hypothesis.

The initial aim of the experiment was to compare the usability between users who had previous experience with the application versus the ones who were new to it, but as it turned out, all data set experts were new to the application and all non-experts have taken part in Experiment 1. This introduced a large bias towards new users and the data below would show that new participants have performed better then users who participated previously. Therefore the main comparison in the analysis was between data set experts and novices.

5.4.8.1 Usability of the Application

The average scores for SUS, Practical Usability, and After Scenario questionnaires is presented in Table 5.12 along with an average of the scores for the two groups to gain an overall view of the usability. The comparison of the usability score with the baseline score established in Experiment 1 is shown in Table 5.13.

Table 5.12: Average usability scores per expertise level in Experiment 2

	SUS	**Practical Usability**	**After Scenario**
Novice users	75.625	34.75	14.75
Expert users	85.5	34.2	17.4
Combined	80.5625	34.475	16.075

F1: There was a 10% difference in the usability scores between novice and expert users.

As it can be seen from the above data, there was a well noticeable difference in the usability rating of the application between novice and expert users. The corresponding letter grades for these SUS scores are B for 75% and A for 85%.

F2: The application was well suited for data set experts

F3: Novice users were often confused regarding what actions were permissible

These results suggest that the application is well suited for expert users for general discovery purposes of the data set [F2], the lower scores by novice users may be due to the was the information is presented. Usability if often descreased due to too much or too little information being presented. The user interface may suffer from information overload or a lack thereof for novice users. This is also suggested in the After Scenario scores which show a 13% difference in user satisfaction [F3]. A possible solution to this problem is the explanation of the available features, and the introduction of methods to make these features more apparent to users.

The combined usability scores of the application has dropped by 2% at most from the scores obtained by UI2 in the first experiment. However as stated there, all participants were unfamiliar with the data set, therefore comparing the scores of UI2 with the scores from novices obtained in this experiment reflect the effect of applied changes more accurately as it represents the same type of user base and will better serve as a comparison to determine the effects of changes.

Table 5.13: Comparison of average usability scores between Experiment 1 and Experiment 2

	SUS	**Practical Usability**	**After Scenario**
EX1: UI2	82	35	17.2
EX2: Novice users	75.625	34.75	17.75
% difference	-6.375%	-0.72%	+3.2%

F4: The usability of the application has decreased from the previous version, but user satisfaction has increased

Table 5.13 shows that the SUS score of the application has dropped with respect to the previous version, showing that the addition of new features have negatively impacted the usability of the application for novice users. However both the practical usability and the user satisfaction ratings have remained at the same level. This shows that users either did not fully appreciate the set of new features or they have not been apparent. A possible reason for the lower SUS scores may be that since all participants in the novice group have previously used the application, they may have had higher expectations for it than new users, causing a more strict rating and lower results. Another possible reason for the lower usability score is the fact that the addition of features caused an increased information overload on users, increasing confusion which negatively affected their experience.

F5: Additional features to the application may have increased the information overload of users, causing the drop in usability rating

5.4.8.2 User Performance

Tables 5.14 and 5.15 present the average of the recorded values for the user performance metrics.

F6: Data expert users were much more effective at using the application than others.

Table 5.14: Average completion time (in seconds) per task per data set experience level in Experiment 2

	Task 1	**Task 2**	**Task 3**	**Mean**
Novice users	357	455.25	528.75	447
Expert users	183.5	185.5	199	189.33

Finding 6 is shown by completion times that are less than half of novices'. This data reflects the observations that have been made during the experiment regarding the levels of confusion of users while interacting with the application and solving the tasks. Expert users when encountering an issue or if become uncertain, quickly resolved to a trial and error approach, while others have spent time going through each individual section and thinking about what to do prior to taking any corrective action.

F7: Data expert users tended to resort to a trial and error approach when uncertain regarding what approach to take, instead of thinking about that to do

Table 5.15: Average number of correct answers and errors per question per participant experience level in Experiment 2

	Correct answers	**Errors**
Novice users	2.25	2.167
Expert users	2.4	1.6

F8: Data expert and novice users got the same number of correct answer in average.

Finding 8 shows that any user was able to use the application for exploratory purposes and possibly gaining the same insight into the presented information as a data expert, however the amount of time required was different depending on user and the level of their skills.

5.4.8.3 Key issues

This section discusses key issues that have been noticed while participants were interacting with the application. Table 5.16 presents the main encountered issues along with their occurrence ratio.

F9: The functionality of the map in the query interface is not apparent.

Users were unsure how to use the map or what functionality it had, or how it could be used to add neighbouring states. While the map already possesses a tool-tip, it appears to be not apparent enough and the map does not become apparent to users but until much later in the usage of the application. A method should be developed which would help bring users' attention to the query map at the first time users are using the location selection interface, in addition the instruction tool-tip should be made more visible. If users become aware of the functionality of the query map, tasks requiring detailed spacial reasoning would require much less effort from users.

Table 5.16: Percentage of users where the stated issue presented itself in Experiment 2, broken down by data set experience level

	Issue	**Novices**	**Experts**
1	Initially confused how condition picker works	80%	60%
2	Unsure how to specify a range of values	75%	60%
3	Clear all functionality not apparent	50%	80%
4	Query map functionality not apparent	75%	80%
5	Difficulty in finding cities	75%	80%
6	Unsure how to save a condition	100%	40%
7	Multiple value selection not apparent	50%	0%
8	Unsure what cities qualify as coastal	50%	60%
9	Difficulty in reducing data set to a city	75%	80%

F10: A majority of the users have cleared event properties individually instead of using the clear all button when resetting the query interface.

Finding 10 is a possible cause of the 'Clear all' button not being apparent enough. To resolve this issue and allowing users to begin a new task much more easily, and to avoid potential errors caused from an accidentally retained selection from a previous task, the button and its functionality should be made more apparent.

F11: Participants had issues in locating the 'Save' and 'Delete' buttons on the condition editor interface.

F12: Novice users did not understand the concept of defining conditions on numeric properties.

Instead of looking at the navigation bar, participants were looking for the controls on the bottom of the screen or expecting the application to automatically add them [F11]. In addition, novice users were confused regarding how to define ranges on the interface as they did not understand the concept of defining two conditions separately [F12]. A possible resolution to these issues is the relocation of the control buttons to the bottom of the screen and to make them more prevalent. In addition an 'Add' button should be introduced, which indicates to users that another condition may be defined. These changes would possibly reduce the confusion of users regarding the interface and avoid time lost due to the inability of locating the buttons and the effort of trying to define a range.

F13: Participants unfamiliar with US geography had major issues locating cities and other geographical features.

The majority of participants were not familiar with the geography of the United States and hence were having issues in trying to locate cities on the map, as reflected by Issue 5. A significant portion of the time spent on tasks 2 and 3 have been used for the exploration of the map in the aim of looking for the states cities in the tasks. Once users have found the area of interest, they were able to quickly complete the task. In order to reduce or eliminate this time waste, the application should offer tools to help users locate a geographical location, which may increase the user satisfaction rating of the application as well as its usability.

F14: The multiple option behaviour in the selection of values was still not apparent to users unfamiliar with similar applications.

While the addition of visual feedback to the query interface has helped some users to understand that the selection of multiple values is allowed, half of the novice users have made errors due to only selecting a single value. When asked regarding this issue they commented that they have not realised the possibility to select more than one entry. To overcome this issue, the application should provide hints with respect to the multiple selection nature of the interface to reduce errors caused by not selecting all the required options.

F15: Participants hesitated using their own reasoning and interpretation of the displayed information.

Many users tried to search for or limit the results to a certain area, such as the cost of California or Kansas City, as the number of markers shown on the map was significant. Participants did not wish to individually check these markers, and the great number of elements on the map seemed to interfere with their ability to reason in a geographical context. Initially users would find the location they were looking for but instead of inspecting only the events that are nearby the requested location, they have started to explore the map and check other events they are not relevant to the task. A possibility is to give users the ability to define custom areas of interest, such as all events near point X, which would remove all events and objects from the map that do not fit in it. This would decrease the load of information on users, however it may have a negative impact as well. During the experiment, it was noticed that some users have not realised that when a location has been selected, only events in the highlighted areas are presented. If the application were to allow for custom locations, user may not realise, after defining such location, that they have set a filter and the application only shows events that satisfy the current limitations.

F16: Participants, especially novice users, has issues in executing counting and summation based tasks.

Participants in general had issues in trying to execute functions such as determining the number of events with property P, which are simple tasks for computers. In order to complete such task, users have to individually inspect each event and record the details on an external source, such as the work sheet. Users who were familiar with applications that query data sets, were expecting functionalities that allow to execute special functions such as 'count the number of events with property X'. While there is no specific interface to give such insight, it was not apparent to users that they may be able to execute such queries by only using the query interface, and observing the results counter. The discovery of such "hidden" features of the application should be helped so that the productivity of users is improved.

5.4.8.4 Comments by Users

Like in the first experiment, the participants were interviews at the end in order to gain more insight into their experience with the application and discover any issues that the investigator may have missed. Table 5.17 shows comments that participants have said at the

Table 5.17: Percentage of users where the stated comment has been mentioned in Experiment 2, broken sown by data set experience level

Comment	**Novices**	**Experts**
Difficult to count fatalities	75%	80%
Confusing range selection	75%	20%
Difficult to look for cities	75%	60%
Better than previous version	100%	NA
One-by-one inspection of events is cumbersome	50%	20%
Simple to use	50%	60%
Clear and nice UI	75%	20%
Results counter is helpful	75%	0%
Fast responses	25%	60%
Tutorial would be helpful	50%	20%

end. It is important to note that these are not statements that users agree or disagree with but comments that participants have explicitly said and their occurrence ratio.

It is interesting to see that all previous users agree that the application has improved since the previous version, hence the stricter grading from previous users may be the cause for lower usability scores.

F17: Comments by users reflect the observed trends and issues.

Observations regarding the cumbersome nature of the information inspection show that users agree with their actions and that the application in its current state is not suitable for the investigation of a large result set, instead it is more suitable for exploratory tasks such as Task 3 where users had to follow a route and inspect the events along it.

A majority of users have complained regarding the difficulty in finding geographical features, especially if they were unfamiliar with the geography of the US. Adding a search feature for locations would significantly help such users in completing their task faster.

F18: A majority of the participants found the application simple to use and informative.

Many of the participants agreed in that the application is simple to use and the information is clearly presented, however the scores from the questionnaires and their performance show that they do not understand all aspects of the information. In addition, users have agreed that they would require some initial help with using the application to be more comfortable with completing their tasks.

5.4.9 Conclusions

While overall usability of the application remained around the same levels, the SUS scores for novice users has dropped by 6%, this shows that the addition of extra features may not benefit novice users positively.

There is a significant different between the performance of data expert and non experts as shown by the completion time, which is less than half for exerts than others.

In order to increase the comfortability of novice users with the application and get the more familiar with the application in less time, the following actions have been taken in preparation for Experiment 3 in the hope of resolving the highlighted issues.

***A1**: Introduction of an on-screen tutorial*

Findings 3, 4, 5, and 12 all highlight the issue that the participants were unsure of the feature set of the application and what certain interface elements served for. It was determined in Finding 8, that one users understood the set of functionalities, they were able to complete the tasks with limited amount of errors and issues encountered. This shows that if there was a way to train users in the application quickly, they would be able to perform significantly better.

An option for training users in the use of the application is the explanation of features by pop-up dialogs, explaining what it is for and how to use it. Such a system has been added for the third prototype in preparation for the third experiment. When the user starts the application for the first time, the tutorial interface walks users through each key interface element and explains what information is being presented and what it is used for. This change would hopefully have an effect on both the usability of the application, user satisfaction, and the performance of the users in completing tasks.

***A2**: UI elements have been made more apparent*

Participants did not notice certain UI elements as highlighted by Findings 9, 10, and 14. The text on elements which were not clearly visible such as on the clear all button or on the instruction on the query map was increased and the shading was altered to make the text stand out better. In addition, instructions have been added to some user interface sections.

***A3**: Condition picker buttons have been made more apparent*

Participants had trouble identifying where the buttons were on the condition picker interface as they were looking for them on the bottom of the screen. To resolve this issue, the buttons have been moved to the bottom of the screen in line to users' expectations. In addition to help users remember that there is the possibility to add another condition, an add button has been added along to the save and delete buttons.

***A4**: Location search has been added*

To help users who are not familiar with the US geography locate cities and other geographical features, location search has been added. When tapping on the location search button, a text field is presented to users where they may enter their location name or address. The application uses the Google Geocoding API[2] to geocode the input and receive corresponding coordinates which are presented on the main map with blue markers.

[2]The Google Geocoding API: https://developers.google.com/maps/documentation/geocoding/

5.5 Experiment 3: Comparative Usability Evaluation of the Desktop and Mobile Applications

This section describes the experiment that has been carried out to investigate the difference between the mobile and desktop application. The experiment used Prototype 3 of the mobile application, which has been created by applying the changes outlined in Section 5.4.9. In addition the effects of the changes on the mobile application have been investigated. The experimental design, hypotheses, and the analysis of results along with the set of findings and actions to be taken are described in the below sections.

5.5.1 Design

The aim of this experiment, in addition to evaluating the effects of the changes applied since the last experiment, was to determine the difference in both the usability of the desktop and the tablet application and how the two applications affect the efficiency and the performance of users in completing a set of tasks.

In order to increase the usability of the application for novice users, and to resolve some of the issues that have been highlighted, the changes and features that have been added to the application are detailed below and in Figure 5.7.

C1: Location search: Users are able to search for a location on the tablet application where the results will be presented on a map with a blue marker

C2: Buttons across the application have been made more prevalent

C3: Hints and instructions have been added to the query interface

C4: A tutorial has been added, which inform users of the major functionalities of the application and the different UI elements

While previous experiments focused on creating and resolving issues around the tablet application, this experiment compares the tablet application with a desktop version. Participants were asked to complete 2 tasks on both devices and a usability questionnaire per device. To ensure there was no bias introduced with respect to the order of devices, participants started with a randomly chosen application out of the two. In previous experiments it has been observed that participants became much more effective at using the application after the first task. To avoid a bias due to a participant being unfamiliar with the data set, a simple practice question was given to them.

In addition to the standard set of measurements that have been taken during the experiment, the investigator recorded the number of times a participant needed help and the nature of the help. Comparing how well participants work with the application without using help is a good comparator for usability. During the interview stage of the experiment, the participants were asked for their preferred application along with a reason for their choice.

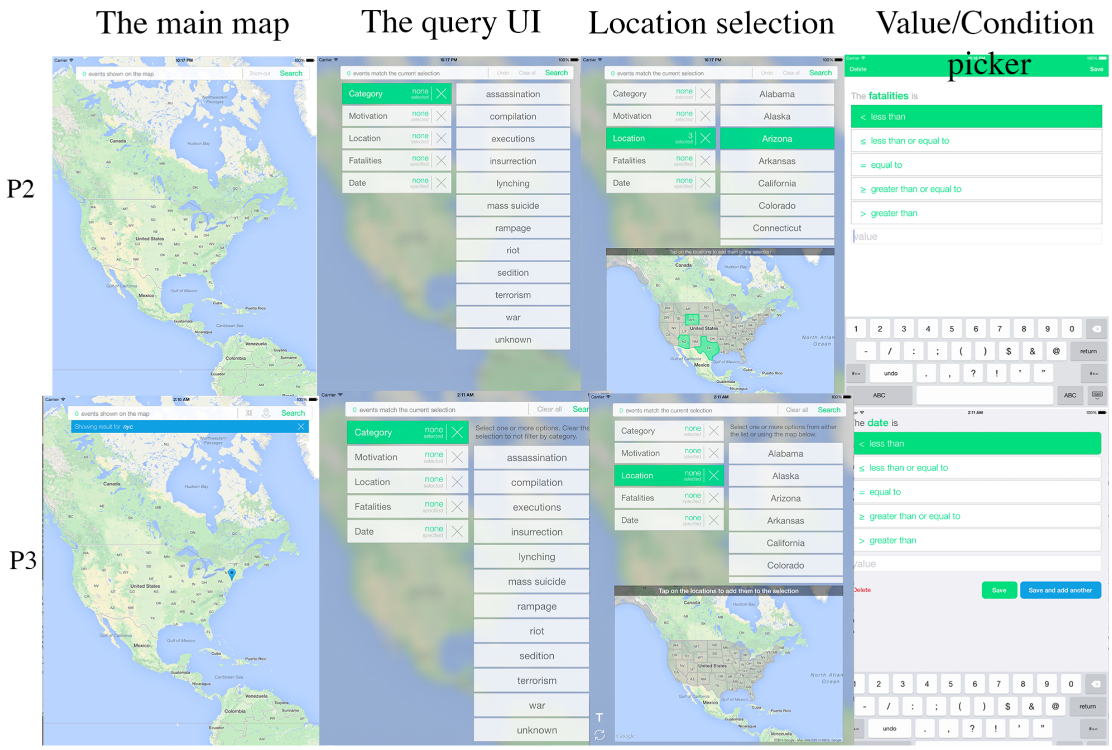


Figure 5.7: Visual demonstration of the changes applied to P2 to create P3

5.5.2 Hypotheses

H1: Data experts should prefer the desktop application due to its one view nature and the experts' experience with similar desktop tools

H2: The usability of the mobile application depends on whether the user prefers using it over the desktop application

H3: Task completion time should be lower in the desktop application as users have the query controls and the data visualisation on one view

5.5.3 Scenario

This experiment followed the same experimental scenario that has been outlined in Section 5.2.1, however given that participants were interacting with two applications, the scenario was executed twice. Participants completed two tasks using the first device, then completed the usability questionnaire, after they have been asked to complete another two tasks on the other application and fill out yet again the usability questionnaire but for the other device. The tasks that the participants were asked to complete are presented in Table 5.18.

5.5.4 Metrics

In addition to the general metrics outlined in Section 5.2.2, the following measures have been recorded in this experiment.

M8: The number of times help was given to participants and the nature of the help

Table 5.18: Task list of Experiment 3 and their purpose

	Task	Purpose
P1	Find all assassinations with extralegal motivations. When did they occur?	The practice task was designed to be simple and encourage light exploration of the data set on the map, allowing users to get familiar with the structure of the information presented on the map. In addition, to complete the task, some light-weight usage of the query interface was required, which would encourage users to get familiar with it.
MT1	Consider the states on the west coast of the US in the 21st century. How many events were there where more than one person has died? What was the most common motivation of these events?	The first questions required users to select a set of states, this selection required the use of the map in order to identify them based on some geographical relationship. In addition a certain date range limitation had to be added to the query to limit the number of results to a manageable set. Once the correct query has been created, users would be presented with 4 or 5 events on the map, which they would require to inspect. This question encouraged the use of the query UI and the map. In addition, it required some exploration of the data set.
DT1	Consider the states that share a border with Mexico in the 20st century. How many terrorism related events were there where more than 2 people have died? What was the total fatality count?	
MT2	Route 5 is a North to South US highway from the Canadian border in Washington state to San Diego in California. How many riots were there along the route in the frst part of the 20th century? How many people have died due to these events?	The second question focused on users' geographical reasoning. It required them to determine which events qualified for the requirements of the task. In addition, the task focused heavily on map based exploration.
DT2	Route 25 is a North to South US highway from Buffalo, Wyoming to Las Cruces, New Mexico. How many riots were there along the route in the first part of the 20th century? How many people have died due to these events?	

M9: The preference of participants between the desktop and mobile applications along with a reason for the choice

5.5.5 Analysis Methods

The experiment followed the analysis method outlined in Section 5.2.5, however there has been additional data as there were two sets of usability questionnaire data per participant.

5.5.6 Types of Users

The participants in this experiment were either friends of colleges of the investigator and have been recruited through either social media, email, or phone. More details regarding the age distribution, familiarity with the data set, and whether they have previously taken part in an experiment are shown in Tables E.5 and E.6. The age group based breakdown may be usable for future research, however the skill based breakdown is useful for the analysis of the recorded data.

Table 5.19: Preference of users based on their experience with the data set

	Tablet preference	**Desktop preference**	**TOTAL**
Novice users	6	6	12
Expert users	3	1	4
TOTAL	9	7	16

5.5.7 Experimental Methodology

Participants were presented with the same experimental setup. Each participant received the same set of tasks and questionnaire that was needed to be completed. The experiment did not split the participants into groups, however instead they were asked to complete two sets of tasks using two different applications. No other conditions but the order of the devices was varied. Each participant completed the tasks in a quiet distraction free environment reducing the interference outside factors.

The remaining of the experimental methodology was identical to the previous experiments, which is described in Section 5.3.6.

5.5.8 Raw Data

The raw data collected in the experiment is presented in Tables 5.20, 5.21, and 5.22.

Table 5.20: Experiment 3: Usability Questionnaire raw data

Participant number	SUS scores per question										PHU									After-Scenario			
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q1	Q2	Q3	Q4
Mobile																							
1	4	3	3	1	3	2	4	3	4	1	5	3	5	4	4	5	2	3	2	4	3	4	3
2	5	1	5	1	5	1	4	1	5	2	5	5	5	5	5	5	5	2	4	5	5	5	5
3	4	1	5	2	4	2	4	1	5	2	4	4	4	4	4	4	3	3	2	4	3	3	4
4	4	2	5	1	4	1	5	1	4	2	5	5	5	5	4	4	5	1	3	5	5	4	5
5	5	1	4	1	5	2	5	1	5	1	5	5	5	5	5	5	5	1	5	5	5	5	5
6	4	1	5	1	5	1	5	1	4	1	5	5	5	5	5	3	5	4	5	4	4	4	5
7	4	2	4	2	5	2	4	2	4	2	4	5	5	4	5	4	4	2	5	4	4	4	5
8	4	4	5	2	3	2	4	2	4	5	5	4	4	3	4	4	5	2	3	4	4	4	5
9	4	1	4	1	4	2	5	1	4	1	5	4	4	4	5	5	4	1	3	4	5	4	4
10	3	2	2	4	4	2	4	5	1	2	4	2	5	5	4	2	4	5	5	3	2	2	4
11	4	2	4	3	4	2	4	2	4	3	4	4	4	3	4	4	4	3	3	4	4	4	4
12	4	1	5	2	4	2	5	1	4	2	5	3	2	4	5	5	4	2	2	4	3	4	4
13	3	3	3	3	4	4	2	4	2	1	5	5	4	4	3	2	5	4	4	3	2	3	4
14	4	2	4	1	4	2	5	1	4	1	4	5	4	4	4	4	4	1	4	5	4	4	5
15	2	3	3	4	5	2	4	4	2	5	5	2	5	2	3	2	2	5	3	4	1	2	5
16	3	3	3	1	4	1	4	1	3	1	5	5	3	5	4	5	4	1	3	3	3	3	4
Desktop																							
1	4	2	5	1	3	2	5	2	5	1	5	4	4	4	5	4	2	2	2	5	5	5	4
2	5	2	5	2	5	1	3	2	4	2	5	5	5	5	4	5	5	3	4	4	4	4	5
3	3	1	4	2	4	1	3	1	3	1	4	4	5	5	5	4	4	3	3	3	3	3	4
4	3	3	3	3	3	4	4	3	2	2	5	4	2	3	2	2	2	3	2	2	4	3	3
5	4	1	5	1	5	1	5	1	5	1	5	5	5	5	4	5	3	1	2	5	5	5	5
6	5	1	3	1	4	1	5	2	3	1	5	5	5	3	4	5	4	1	1	4	2	2	4
7	5	2	4	2	4	2	4	1	4	1	4	5	4	4	4	4	3	3	4	5	5	4	4
8	3	4	3	3	4	2	5	3	4	5	5	4	5	4	4	3	5	2	3	5	4	4	4
9	4	3	4	1	4	2	4	2	4	1	5	4	5	4	4	4	4	3	3	4	4	5	4
10	4	1	4	2	4	1	5	2	4	1	5	3	5	5	4	4	5	3	4	4	4	4	4
11	5	2	4	2	4	1	4	1	4	1	5	5	5	3	4	4	4	2	3	4	5	4	4
12	3	4	3	2	4	1	4	4	5	1	5	1	4	2	4	2	2	2	2	2	3	4	3
13	4	2	4	1	3	1	5	1	3	2	5	4	5	4	5	4	4	3	4	4	3	4	5
14	2	3	3	1	3	2	4	3	3	1	5	4	5	4	4	5	3	4	4	4	3	4	3
15	1	5	3	4	2	3	2	5	2	5	5	1	4	2	2	1	3	5	2	1	3	3	4
16	3	1	4	1	5	1	5	1	5	1	5	5	3	5	5	5	4	1	3	4	4	5	4

Table 5.21: Experiment 3: User performance raw data

Participant number	Completion time (seconds)		Correct answers		Help		Errors	
	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
Mobile								
1	305	290	0.5	0.5	1	0	1	1
2	157	250	1	1	0	0	1	2
3	366	227	1	1	2	0	1	2
4	165	220	0	1	0	0	2	0
5	188	114	0	1	0	0	1	0
6	198	146	0	1	0	0	1	0
7	307	279	1	1	0	0	1	1
8	248	351	1	1	0	2	1	1
9	156	168	1	1	0	0	0	0
10	364	377	1	0	0	2	0	1
11	317	305	0	0	0	0	1	1
12	333	374	0	0	0	3	3	4
13	225	320	1	1	0	0	2	2
14	272	304	1	1	0	0	0	0
15	240	202	0.5	1	3	0	3	1
16	402	231	0.5	1	0	0	3	1
Desktop								
1	160	161	1	1	0	0	1	0
2	372	128	1	1	1	0	4	2
3	220	396	1	0	0	2	2	2
4	203	250	1	1	0	0	2	0
5	90	141	0	1	0	0	0	0
6	132	247	1	1	1	0	1	0
7	191	245	0	1	0	0	2	0
8	280	305	0	0.5	0	2	0	0
9	164	203	1	1	0	1	0	0
10	279	526	1	1	1	1	0	1
11	144	287	0.5	1	0	0	1	2
12	300	181	0	1	3	0	2	1
13	193	317	0	0	0	0	0	0
14	201	232	1	1	1	0	1	1
15	207	274	0.5	0.5	0	1	3	0
16	162	231	0.5	0.5	0	1	3	1

Table 5.22: Experiment 3: Participant properties

Participant number	Started with	Tech familiar	Data expert	Preference	Previous user
1	Desktop	yes	no	Desktop	yes
2	Tablet	no	no	Tablet	yes
3	Desktop	no	no	Desktop	no
4	Tablet	yes	yes	Tablet	yes
5	Desktop	yes	yes	Tablet	yes
6	Tablet	yes	yes	Tablet	yes
7	Desktop	yes	yes	Desktop	yes
8	Desktop	yes	no	Tablet	no
9	Tablet	yes	no	Tablet	yes
10	Tablet	no	no	Desktop	no
11	Desktop	no	no	Desktop	no
12	Tablet	no	no	Tablet	no
13	Desktop	yes	no	Desktop	yes
14	Desktop	no	no	Tablet	yes
15	Tablet	yes	no	Tablet	no
16	Desktop	yes	no	Desktop	yes

5.5.9 Analysis of the Data

This section discusses the data collected throughout the experiment with respect to the three hypotheses and the research objectives.

5.5.9.1 Preferences and Comments

So far in the analysis of the experiments, the main focus was the difference between novice and expert users. It may be a possibility that usability scores originating from a person are strongly affected by a user's perception and prejudice of the application and the device being interacted with [H2]. In order to investigate this hypothesis, the participants' results was split up, for analysis, based on their preference of the applications. This is presented in Table 5.19.

F1: Overall, the tablet application was preferred over the desktop application.

F2: Focus on data presentation benefited application usability.

Data experts preferred the tablet application over the desktop one, even through other similar applications that they have had experience with were desktop based. When interviewed regarding their choice, they reasoned that the information in the tablet application was presented in more familiar terms and that it was easier to navigate with the information being presented much more nicely. It seems that a big focus on the data instead of the interface and the controls benefit usability more than having the query interface and the data on one screen.

Table 5.23: Percentage of users who have stated one of the comments

Comment	Percentage of users	Application
Range select view switches back to individual value selection	31.25	Desktop
Very informative	37.5	Common
Cool	68.75	Common
No clear all button	25	Desktop
Easier to set range	75	Desktop
Easier to undo	25	Mobile
Easier to navigate	43.75	Mobile
Scale on map would be useful	12.5	Common
Tutorial a bit intrusive	31.25	Mobile
Biased towards desktop	37.5	Desktop

On the other hand, the preference of casual users was not clearly separable. Exactly half of the participants, who were not experts of the data set, preferred the tablet application. One was expecting that novice users would prefer the tablet application because of the simpler interface, given that the desktop application focuses heavily on the data set and the interface presents little help. When asked regarding their choice, some users stated that they are either biased towards computers or simply don't like tablets. Other users found the context switch between the query interface and the map confusing and preferred to have the information and the controls on one view.

F3: The preference of users did not depend on their skill levels or their previous experience with other similar applications

One would have expected to have a clear separation of preference based on the skill levels of the users. Based on the literature review in Section 3, previous research showed that expert users would prefer the more complex desktop application as it pro vices faster one-screen controls and immediate full results views. While on the other hand, the mobile application would be more preferred by novice users as it has clearly laid out and explained controls along with a tutorial and has a bigger focus on the exploration aspect.

FX: Overall, the mobile application achieved a 74.06 SUS score and the desktop application achieved a 75.63. This shows that the two applications have achieved an equivalent usability. These scores are backed up by the after scenario ones.

5.5.9.2 Usability of the Applications

In the previous experiment it was determined, that there was a difference between the perceived usability between novice and expert users. As presented by Table 5.24, the perceived usability was present in P3 as well with a wider score gap of 17.5, up from the previous 10. However it was interesting to see that the practical usability and the after scenario scores have not changed significantly for casual users. In fact it has increased for experts, indicating an improvement. These trends may support Finding 5 of Experiment 2, which states that there was an observable relationship between the decrease in usability and the introduction of new features.

F4: The perceived usability of the mobile application has decreased, while user satisfaction remained the same.

Table 5.24: Average usability scores for the mobile and desktop applications

	SUS	Practical Usability	After Scenario
Mobile			
Novice users	71.25	33.83	14.75
Expert users	88.75	40.5	17.4
Tablet preference	81.94	36.89	17.22
Desktop preference	67.5	33.71	13.86
Average	75.63	35.5	15.75
Desktop			
Novice users	72.5	34.084	15.5
Expert users	78.75	33	15.5
Tablet preference	66.39	32.11	14.67
Desktop biased	83.93	36	16.57
Average	74.06	33.81	15.5

F5: The method of the introduced tutorial was not the optimal approach to increase overall usability.

The practical usability rating reflects the users' perception of the usability of the different system components. When inspecting the individual entries on the scale, one was able to see that the entries which have scored lower from the previous experiment are the users' ability to understand the information that is being presented and the functions available. It is surprising to see these entries to score lower than previously, as the added tutorial was created to address the issue of confusion and uncertainty with respect to the information and interface presented. Given that these attributes are low, the introduced tutorial does not give the right suggestion to users and either an alternative method of introduction or the addition of less intrusive hints should be created to combat users skipping the tutorial and omitting the help given to them.

F6: The mobile application was well suitable for expert users

The usability scores indicate that the changes have created a tablet application which is very effective and easy to use by data experts. In fact, the usability scores for experts correspond to an A grade, therefore it can be assumed that it possesses all the required functionalities that an expert user would need for the purposes that the application was created for. However given that the usability score obtained by novices is a C+, the application requires a more easier to understand presentation that is more optional for casual users while still maintaining the current level of usability by experts.

F7: The application was also suitable for novice users who do not have a bias towards the desktop application

When considering simply the users who prefer the tablet application, the results obtained show a SUS usability rating of 81.94, which is above the minimum of an A grade (exceptional) usability, hinting that these participants have found their experience with the application beneficial and would recommend its use to others[3]. They were equally as much satisfied with their performance in completing their tasks as the data experts, and have found only minor issues with the application.

F8: The desktop application was suitable for participants who prefer to use it over the mobile app

The usability of the desktop application was similar to the tablet application when considering non data experts, however when considering only the expert users, the desktop usability was significantly lower than the tablet's. When looking at the usability of the desktop application based on users' preference, one was able to see that the SUS scores are significantly higher for people who preferred it over the tablet app, in fact the scores broken down per preference groups are almost the inverse of the tablet usability. Considering this breakdown of users, one was able to see that there was a relationship between the usability of the applications and the user preferences. The reason for the high difference between the SUS scores by experts when considering the two platforms may be explained by the fact that 75% of the expert users preferred the tablet application.

F9: Preference for an application influenced it's perceived usability and how satisfied users were with it.

5.5.9.3 User Performance

In the previous experiment, it was shown that expert users complete tasks faster and with less errors than novice users. It was expected that this behaviour translated onto the desktop application as well. Given that expert users were more used to desktop applications, one was expecting that they would be able to complete the set of tasks faster on the desktop than the tablet.

In addition, the singular interface should have allowed for a better overview of the information, benefiting completion time, however it was unclear how experts were affected by the more complex user interface. It was possible that they would be making more mistakes, but

Table 5.25: Average completion time (in seconds) per task on the tablet (T) and desktop (D)

	T1	**T2**	**T Average**	**D1**	**D2**	**D Average**
Novice users	282.083	283.25	282.67	223.5	270.09	246.79
Expert users	214.5	189.75	202.13	154	220.75	187.36
Tablet preference	217.44	236.56	227	215.67	217.89	217.22
Desktop preference	326.57	289.86	308.21	192.71	309	250.86
Overall	265.19	259.88	262.53	206.13	257.75	231.94

it was expected that they would be able to complete the tasks with a similar correct answer rate.

Table 5.25 shows the average completion time of each task per group split as novice/expert and desktop/tablet preference, and tables 5.26 and 5.26 show the number of errors and helps given per question as well as the number of correct answers for the tasks for the tablet and the desktop application respectively.

Table 5.26: Average number of correct answers and errors per question on the mobile and desktop applications

	Correct answers	**Errors**	**Help per question**
Mobile			
Novice users	1.42	1.33	1.08
Expert users	1.25	0.75	0
Tablet preference	1.39	1.11	0.88
Desktop preference	1.36	1.29	0.71
Overall	1.335	1.04	0.54
Desktop			
Novice users	1.33	1.13	1.167
Expert users	1.5	0.63	0.25
Tablet preference	1.5	0.94	1.11
Desktop preference	1.21	1.07	0.71
Overall	1.415	0.89	0.7085

F10: Data set experts completed tasks faster on the desktop and with fewer errors than others.

Following the trend confirmed in the previous experiment, expert users have completed the tasks about 25-30% faster than others and have made less errors and required less help.

F11: Participants got familiar using the desktop application faster, however task completion times seemed to be equivalent on the two applications.

As presented by Table 5.24, on average the completion times on the desktop application were about half a minute shorter than on the tablet. However when looking at the per task

[3]http://www.measuringusability.com/sus.php

break down, one can see that the time it took to complete task 2 on both devices is identical and that the extra time is accountable only towards the first question. This difference is visible on both the expert/novice and tablet/desktop preference breakdowns. A possible source for this increased time was the presence of the tutorial, which in its current form was slightly intrusive as users were not able to disable it, taking time away from focusing on the task at hand. In order to investigate the effect of the tutorial, an a/b testing with the tutorial on and off should be run in future work.

F12: Participants' confidence in using an application depended on how much they liked to use it, which in turn affected the time it took users to complete the tasks and the number of errors they have made during the experiment.

In the previous section, the effect of users' preference on the received SUS grade was discussed, and based on the data presented by Table 5.24 shows that users' confidence with the application has an effect on their performance and the ability to complete tasks. Participants who preferred the tablet application have completed the tablet tasks faster than the other group. Participants who preferred the desktop application have completed the desktop tasks much faster than the ones who did not. No similar correlation was noticeable for the number of errors, the answer rate, and the number of helps given.

5.5.10 Conclusions

Data shows that users were capable to complete tasks much faster using the desktop application [Table 5.25], indicating that Hypothesis 3 (H3) is true. However the number of errors that users have committed per question is higher on the desktop-based application than on mobile, so is the number of times help has been requested or needed. These figures indicate (as presented in Table 3.3) that the desktop application may be harder to use or that the way information is presented on the desktop is more difficult to grasp. Interestingly, the usability of the desktop and tablet applications were equivalent [Table 5.24]. This is reflected in the achieved SUS scores, which is backed up by the After-Scenario scores which reflect the same relationships.

It was shown that H2 holds true, as the usability rating for the mobile application was high when the participant liked the application. On the other hand, the users who preferred the desktop application found *PVGeoVisualisation* to be more usable. Following on the findings [F9], it was shown that there was no noticeable relationship between participants' experience with the data set and which application they preferred, hence H1 was shown to be not true.

Most users preferred the tablet application to explore the data set, including data experts [Table 5.19]. However comments by users highlight that the unified view of the desktop-based application was preferred over two separate views, but the interface of the mobile application, in general, was more welcoming and simpler. Once users got familiar with the mobile application, they achieved the same level of performance than on the desktop equivalent [F11]. Users' confidence in completing the tasks was affected by their preference for an application, which in turn had an effect on how fast they were able to complete the given exercises [F3, F9, F12, Table 5.26].

The introduced tutorial had the inverse effect of what was desired; instead of improving the overall usability, its inability to be dismissed caused frustration with users and decreased the overall perceived usability of the application [Tables 5.12 and 5.24].

The application possessed some minor issues that made it more difficult for novice users to use, however given the data, it was shown that it possessed the minimal set of features that is required by expert users to use such an application. Overall, the application received very positive feedback from the participants and seemed to be suitable for the exploratory study of the underlying data set.

Chapter 6

Conclusions

In this study, the usability of the *PVGeoVisualisation mobile* was investigated through several usability studies. The application aimed to address the need for a mobile Linked Data application capable to visualise geographical information and enable usage by novice users, reducing the amount of prerequisites for using the application to almost nothing (users were required to be able to read in English and be able to interact with the application through hand gestures).

Throughout the investigation, the design of the mobile application was improved based on the analysis of these usability studies. Along with the improvements, different factors affecting the perceived usability of the application were investigated, such as the addition of features or the types of users interacting with the application. Later, this application was compared to a desktop-based *PVGeoVisualisation* application.

This section summarises the findings and results of the usability studies. Initially the evidence supporting the answer to the research question will be presented, then the findings relevant to each research objective will be stated along with how the results have been achieved. Finally, the contributions of the research will be summarised with their limitations highlighted. In addition the final remarks of the author will be presented along with suggestions for possible future work.

6.1 Research Question

To what extent can mobile applications achieve equivalent usability to existing desktop applications when visualising Linked Data geographical information?

Overall, the *PVGeoVisualisation mobile* application has achieved an equivalent usability to its desktop counterpart [Experiment 3: FX], however, there are some restrictions that need to be considered. It was found in the third experiment, detailed in Section 5.5, that the usability of the applications depended on users' platform preferences as shown by findings F9, F8, and F12. Participants who liked and preferred to use the mobile application found it to be much more usable over the desktop application, on the other hand, users who preferred the desktop application found that to be more usable [Experiment 3:F12]. Therefore we can say that one application cannot fully replace its counterpart.

It was found that the participants who were familiar with the data set, including the structure and contents of the visualised information, have performed better than novice users both in terms of task completion time and the number of errors committed [Experiment 2: F1, F6, F7]. On average, users' performance on the desktop-based application, *PVGeoVisualisation,* was better than on mobile [Experiment 3: F10, F11]. But it is important to note that once users got familiar with the mobile application, the difference in completion time has become less significant [Experiment 3: F10, F11]. While interacting with the desktop-based application, novice users requested more help on average than while using the mobile app [Experiment 3: F10].

As the applications were specifically designed for the visualisation of geographical data sets by placing markers on a map based on the location some geographical feature of data points. Hence, the *PVGeoVisualisation* applications are not suitable for investigating the properties of and the relationships in the data set, and the author is unable to conclude whether these results would be reflected in a different use-case than the applications were designed for.

6.2 Research Objectives

In this section, the work that has been carried out along with a summary of the relevant findings is presented for each research objective.

6.2.1 RO1: Review of the State of the Art

To survey the state of the art in research into Linked Data and best practices in visualisation techniques for mobile devices, including the visualisation of Linked Data and geographical information.

Overall, this research objective was fulfilled. 8 mobile Linked Data applications were revised, 5 geographical visualisation applications were reviewed, a set of 16 requirements were identified for the PVGeoVisualisation mobile application developed as part of this work.

Initially, the background of the State of the Art was presented which covered an overview of areas such as Linked Data, information visualisation, mobile UI design, and geographical information and its visualisation. After this overview, Linked Data visualisation was investigated and a set of requirements for applications was established based on the survey by Dadzie et al. [15]. A set of example applications applying visualisation techniques on Linked Data have been evaluated according to the established set of requirements. It was found that only DBpedia mobile and map4RDF were designed for use by novice users by fulfilling the requirements for targeting a wider user base [Section 3.1]. The positive aspects of the presented applications have been used in the design of the *PVGeoVisualisation* applications, allowing the researcher to identify areas of focus.

Different methods of storing geographical information using Linked Data were presented [Section 3.2]. In addition to presenting the Basic Geo Vocabulary and GeoLinkedData, a list of applications which apply these methods to manage geographical information was presented. Finally, it was discussed how the USPV data set references geographic information and what the applications would need to do to be able to display USPV events on a map. Based on this information, the USPV data set was enriched with geographical coordinates and the application was able to visualise the information effortlessly.

Mobile Linked Data applications were described in detail with a focus on their management of the Linked Data information [Section 3.3]. Based on the review of the applications, some challenges that apply to Linked Data on mobile were highlighted. The technical review of the applications allowed the author to identify a set of features that would be essential for the *PVGeoVisualisation mobile* application in order to provide a good user experience and achieve a high usability score. The presented applications have been compared to each other with respect to these required features in Table 3.2.

Finally, the practices that have been used for the usability study techniques have been introduced, including the System Usability Scale and the Think-Aloud protocol.

The literature review surveyed the relevant research domains and presented similar applications to *PVGeoVisualisation*. However it's possible that the survey is not complete and that some previous findings are missing from it.

6.2.2 RO2: Approach for Geographical Linked Data Visualisation

To design an approach to visualising geographical information obtained from Linked Data end points, focusing on the usability and the user experience on mobile devices.

This research objective was fulfilled by the design paper-based of the user interface of *PVGeoVisualisation mobile*. Based on the literature review and the detailed use-case of the *PVGeoVisualisation* applications [Section 4.1], a set of 9 to 12 requirements have been established that the created applications would have needed to satisfy. These requirements have been created based on the visualisation and linked data requirements that have been used to assess the example applications. Table 6.1 shows how the two *PVGeoVisualisation* applications satisfy these requirements.

Table 6.1: Table showing how well *PVGeoVisualisation* and *PVGeoVisualisation mobile* satisfy the requirements outlines in Section 4.2

	Requirement	**PVGeoVisualisation**	**PVGeoVisualisation mobile**
1	Visualise the USPV event instances on a map	yes	yes
2	Access the USPV data through the SCSS end point	yes	yes
3	Use SPARQL and RDF	yes	yes
4	Be able to filter information	yes	yes
5	Provide a visual SPARQL query building UI	yes	yes
6	Be able to select an event and display its properties	yes	yes
7	Be usable by novice users, not just data set experts	yes	yes
8	Hide the underlying technology and data structure	yes	yes
9	No designated server	no	yes
10	Be able to work offline	no	yes*
11	Provide the most up-to-date version of the data	yes	yes*
12	Follow the mobile UI design guidelines	-	yes

There are two remarks regarding the mobile application, it is capable to work offline and all functionality regarding the Linked Data aspects of the application will function when not connected to the Internet, however location search will not function as it is a service that required the geocoding of the inputted address. The mobile application does not necessarily provide the most up to date information. As presented in Section 4.5, it was found that a constant update process slowed the mobile application down significantly, hence the decision to only update the local data set once per session was made. When the application is launched, it will refresh the locally stored Linked Data with the newest version from the SCSS end point.

Table 6.2: A comparison of the different mobile linked data applications and *PVGeoVisualisation mobile* with respect to the Linked Data challenges on mobile

	DBpedia mobile	Ontowiki mobile	mSpace mobile	Stevie	Qpedia	More!	Who's Who	wayOU	PVGeoVisualisation mobile
Offline usage	no	yes	no	no	no	no	no	no	yes*
Server independent	no	no	no	no	no	yes	no	yes	yes
General purpose	no	yes	yes	no	yes	no	no	no	no
Local filter	no	yes	no	no	no	no	yes	no	yes
Any source	yes	yes	yes	no	no	no	no	no	yes*
SPARQL	yes	yes	yes	yes	yes	yes	yes	yes	yes
Dereferencing	yes	no	yes	no	no	no	no	no	no

Table 6.2 presents which beneficial features of mobile Linked Data applications *PVGeoVisualisation mobile* implements and compares it to the example applications. As states, the application is capable for offline use and it does not require a custom server to manage the data and create visualisations of it. It also possesses a local filtering technique through the visual query building interface, through which users are able to limit the amount of information that is presented on the map. The application, even though use case and data set specific, could be adapted and be used by any other geographical Linked Data set. Even though application does not feature URI dereferencing, it is a feature that could be to enhance the exploratory aspects of the application, including information access from other data sets.

Table 6.3 presents how the mobile application fulfils some of the Linked Data visualisation requirements. Users are unable to inspect the raw RDF data or create custom SPARQL queries on the data set. This is due to the requirements of *PVGeoVisualisation mobile*, where the underlying data structure was required to be hidden from users. The application does not visualise the structure of the underlying RDF graph, instead it focuses on presenting the meaning of the information that is used by the application. The application possessed the ability to undo actions on the query interface, which can be considered as a form of navigation history. The use-case of the application do not consider the possibility for co-operation between users, hence it does not provide functionalities for screen-sharing or exploring of the current results set. These are possible areas for future work in case the application's use-case would be extended.

There have been a set of Linked Data visualisation requirements presented in Section 3.1.1. Based on the findings of the experiments, Table 6.4 examines how each of them have applied to *PVGeoVisualisation mobile*.

Table 6.3: A comparison of different mobile linked data visualisation applications with respect to the set of outlined requirements

	DBpedia Mobile	Ontowiki Mobile	mSpace mobile	Stevie	Qpedia	More!	Who's Who	wayOU	map4RDF	GeoNames browser	PVGeoVisualisation mobile
Overview data	yes	no	yes	yes	no	no	no	no	no	yes	yes
Filter out data	yes	yes	yes	no	yes	no	no	no	yes	yes	yes
Detail/Drill down view	yes	yes	yes	no	no	no	no	no	yes	yes	yes
Information preview	yes	no	yes	no	no	no	no	no	yes	yes	yes
Visualise relationships	partial	yes	no	no	no	no	yes	no	no	no	no
Multidimensional data	yes	yes	yes	no	yes	no	no	no	yes	yes	yes
Export data-set	no	no	yes	no	no	no	no	no	no	yes	no
Intuitive navigation	yes	yes	yes	yes	yes	-	yes	?	yes	yes	yes
Navigation history	yes	yes	no	no	yes	no	no	no	no	no	yes*
Ability to explore	yes	yes	yes	partial	partial	no	partial	partial	yes	yes	partial
Raw RDF view	no	no	no	no	no	no	no	no	no	yes	no
Custom SPARQL	yes	no	no	no	yes	no	no	no	no	no	no
Large data set support	yes	no	yes	no	no	no	no	no	yes	yes	yes
Unrestricted LD navigation	yes	no	yes	no	no	no	no	no	no	yes	no
Visual query building	yes	no	yes	no	no	no	no	no	yes	yes	yes
Share current view	no	no	yes	no	no	no	no	no	no	yes	no

Given the requirements of the application, *PVGeoVisualisation mobile* fulfils most of the requirements of Linked Data mobile applications and Linked Data visualisation. While this study focuses on the design for a tablet, the findings and approaches may be translated to a smaller screened device. In addition, the visualisation issues may be less apparent on a tablet than it would be on a smartphone. The design of *PVGeoVisualisation mobile* ensured that the application achieves a high usability as most of the challenges which would decrease the perceived usability and raise issues have been addressed. The usability studies have shown that the mobile application has received a high usability rating from participants who preferred to use it [Experiment 3: F6], and has achieved an equivalent usability to its desktop counterpart [Experiment 3: F6, F7, F9].

6.2.3 RO3&4: Iterative Development and Usability Evaluation of the Mobile Application

This research objective was fulfilled by implementing the design of *PVGeoVisualisation mobile* and improving it through several prototype iterations. These prototypes were evaluated through usability studies where different issues issues with the prototype application were identified along with areas of focus for future iterations, such as additional features or feature rework. In addition, to this the experiments were used to investigate different aspects of the mobile application. Initially two alternative designs for the query interface have been created and the winning design was selected based on the usability scores and user feedback [Section 5.3]. It was shown that a list based presentation of options was more usable over a tag cloud based approach. The experiment also highlighted the need for strong visual feedback on actions, as users forgot their previous actions or were unsure if the state of the application has changed after they have done something. Therefore the mobile application received improvements to address these issues including a results counter was added indicating the number of events that satisfy the current selection. In addition, some features have been added to facilitate geographical reasoning for users unfamiliar with US geography.

The second usability experiment focused on evaluating the difference between novice and data set experts, as well as analyse the difference between participants who previously used the application and the others who were new to it. It was found that users who were familiar with the data set have performed better overall [Sections 5.4 and 5.5]. They have completed tasks much faster than novice users and in addition experts have committed less errors than the others. This experiment also showed the need for hints and visual cues on the user interface. Tool-tips, helpful messages, and a tutorial have been added in order to improve the usability of the application for novice users.

The multiple usability studies allowed to gain a much needed insight into the issues with the mobile application and enabled the identification of issues. This resulted in a mobile application that is on par, in terms of usability, with the desktop-based PVGeoVisualisation application [Section 5.5.10]. However these experiments have not been carries out with a large number of participants, therefore the results are not necessarily accurate. It would have been beneficial to run these experiments with a much bigger sample of participants in order to gain further insight into issues, have the chance to investigate more user-application interactions, and collect more data. Overall, the dissertation fulfilled RO3 and RO4.

Table 6.4: Table indicating how the Linked Data visualisation requirements apply to *PVGeo-Visualisation mobile*

	Requirement	**Status**	**Comment**
A1	Data overview	yes	The overview and the off-screen visualisations were combined into one method [Section 5.3.9]
A2	Information filtering	yes	Demonstrated by the continuous use of the query UI
A3	Drill-down into areas of interest	yes	Corresponding to zooming on the map to spread out markers.
B1	Visualise relationships	no	The application did not visualise relationships between events. It has achieved a high usability nontheless.
B2	Display multidimensional data	no	The application did not visualise multidimensional data.
B3	Export visualised data	yes	This requirement was not in scope of the application, however the author assumes that users (especially experts) would want to reuse the results in other applications
C1	Clear and intuitive navigation through the web of data	partial	This requirement depends on the use case of the application. Clear and intuitive navigation is essential, however if the use-case of the application does not require it, the ability to navigate the Linked Data cloud is not essential.
C2	Explore the data without restrictions	yes	Users were able to explore the USPV data without restrictions. Upon clearing the query UI, all events on the map were shown and users could explore the data set in its entirety.
C3	Inspect underlying raw data	no	This requirement is use-case specific, however as seen with the mobile application, if the use-case of the application is exploration, there was no need to access the raw data
C4	Option to run custom SPARQL	partial	Users want to be able to define their filtering options on the data. The users were able to create custom queries visually. Experts did not require the ability to write SPARQL through a text editor.
C5	Extract raw RDF data	yes	Given that the author assumes data exportation would be useful, exporting the raw data in its original format would be equivalently useful
D1	Nagivate easily through a large data set	yes	Users were able to navigate the USPV data set through an intuitive map based interface.
D2	Allow exploration of the data to gain understanding of it	yes	Initially users explored the information on the map prior to accessing the query UI in order to understand what sort of information was being presented.
D3	offer the creation of queries through helper methods	yes	Without the visual query builder and the hints, novice users would not have been able to filter the presented information.
D4	Allow analysis of regions of focus	?	This area has not been investigated. While the highlight of the selected location helped in the visual analysis of the data, the application did not feature helper methods (e.g. count the number of fatalities) to analyse the presented information. This sinvestigation is presented as possible future work [Chapter F].
D5	Present the results of queries and usage to others	?	The mobile application did not feature collaboration techniques, hence this point is unresolved. The inclusion of collaboration methods is one of the proposed future features [Chapter F].

6.2.4 RO5: Comparison of the Mobile and Desktop Applications

The mobile application was compared to its desktop-based counterpart in order to determine whether *PVGeoVisualisation mobile* is capable to achieve an equivalent usability to the desktop-based version of the application. This comparative study was the main aim of this dissertation and it was concluded in the previous section [Section 6.1] that overall an equivalent usability is possible. The study investigated how the different types of users interacted with both *PVGeoVisualisation* applications and it was found that the user's platform preferences affected the usability rating of the applications [Section 5.5]. As with the previous experiments, this study was done only with a small number of participants [Section 5.5.10], hence the conclusions and the analysis may not be representative.

6.3 Future Work

As the experiments have been carries out with only a small number of participants, possible future work following on this study would possibly need to focus on the evaluation and the design of the *PVGeoVisualisation* application through longer, more exhaustive usability tests involving significantly more participants that can represent the targeted user base of the application. In addition, to evaluate whether the findings of this dissertation apply to any geographical Linked Data set, the evaluations may investigate the use of different data sets in order to verify the findings.

There have been some issues that have been identified in Prototype 3 that would need to be addressed in order to further improve the mobile application's usability, such as improving the on-screen tutorial to make it easily dismissible, this is to correct the drop in usability caused by the current version of the tutorial. Other possible future features, such as additional visual cues and advances visualisation options are detailed in Chapter F found in the appendices, these features would enhance users' performance and make it easier to understand the data. Additionally, the mobile application may benefit from other visualisation techniques, however this may affect the usability negatively.

In addition to changes to the mobile application, the desktop visualisation tool may also benefit from some of the features introduced in this dissertation such as the indication of the selected states on the map and the map based area selection.

6.4 Final Remarks

Overall, the presented *PVGeoVisualisation mobile* application possessed the set of minimum features that experts users required in order to use the application well and perform their tasks with a minimum number of errors [Section 5.5.10]. The general feedback from users was positive. They have thought that the mobile application was fun to use, and it was informative. It allowed them to learn about the presented information and the political violence history of the US. They liked that they had the option to explore the data set while completing their tasks. They have complemented the design of the application and its simplicity. Other than a few minor issues that may be addressed with minimal effort, the application suited its use-case well [Section 5.5.10].

Appendix A

Generic Desktop Linked Data Visualisation Applications

A.0.0.1 Fenfire

Fenfire [23] is a desktop application for the interactive visualisation of linked data in a graph format. It is a generic tool for visualise any RDF graph independent of the source. Users are able to begin exploring the web of data by specifying an origin URI.

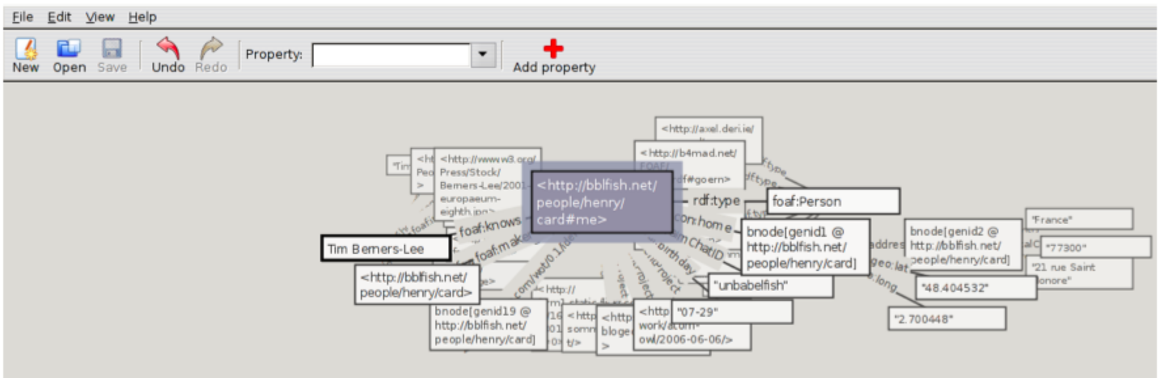


Figure A.1: Image of the user interface of *Fenfire*. [23]

A.0.0.2 IsaViz

IsaViz[1] is a desktop semantic web application that allows the exploration and navigation of RDF resources. Similar to Fenfire, IsaViz displays the RDF information in a graph-based visualisation, and like Fenfire the users are required to supply a URI as a starting point from which one may begin exploring the data set.

The application allows the usage of the underlying data by both both expert and non-expert users. The application allows the users to interact with the RDF data directly, giving more flexibility for expert users, while the visualisations allow non-experts to better understand the data and the relationships between entries. Users are able to search and filter the data as required, reducing the amount of information present, which is useful in case of an abundance of entities displayed.

[1] http://www.w3.org/2001/11/IsaViz/

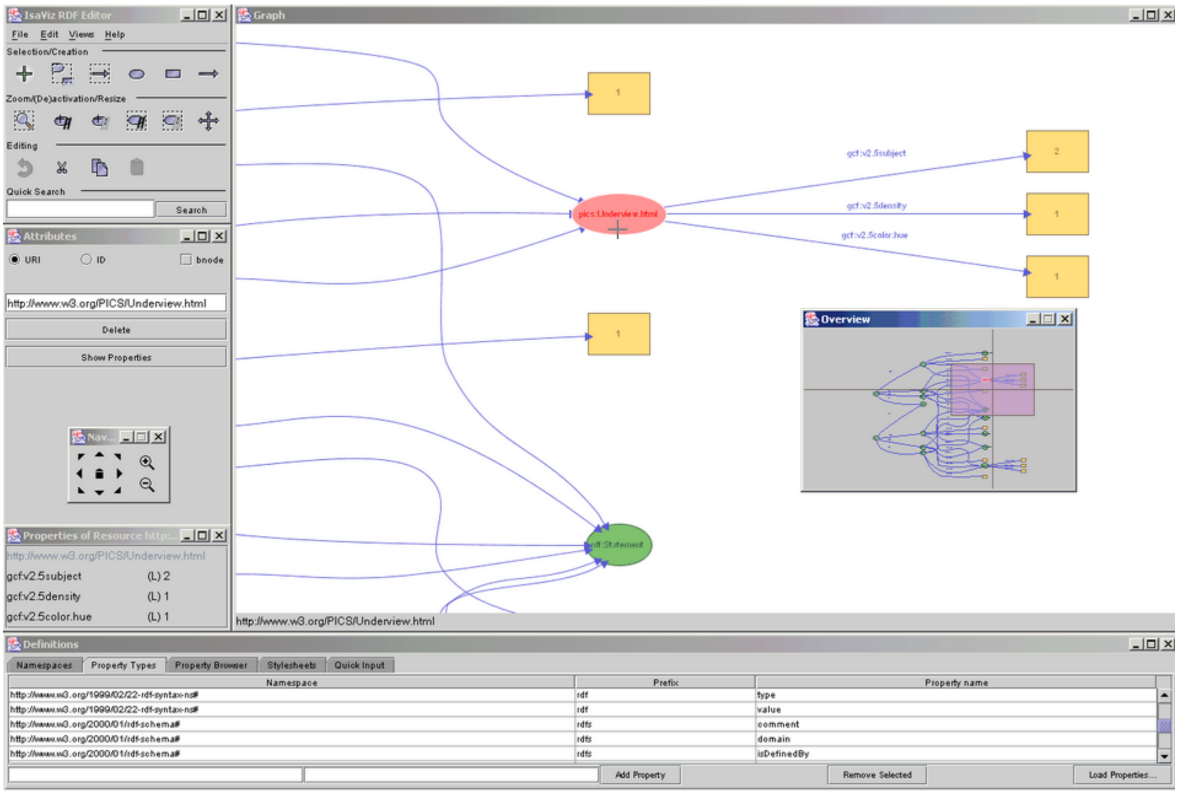


Figure A.2: Images of the user interface of *IsaViz*

In addition to the visualisation and exploration options, the application allows the data to be exported as required.

A.0.0.3 LESS

LESS [2] is an approach which aims to visualise RDF resources using templates. The overall aim of the approach is to create visualisations of linked data that are easily usable and understood by non-expert users. The approach has some basic templates for different types of data, which then can be combines together to make more complex visualisations. Due to the flexible nature of this approach, LESS is capable for any type of visualisation given that a template has been created for it.

A.0.0.4 OpenLink

OpenLink is a linked data browser which allows users to explore the linked data web starting from a URI that has been specified by the user. The application has different views based on which data entries are visualised. It allows users to filter data and save SPARQL queries for later use.

The application offers different types of visualisations based on the type of information. Some examples of available visualisations are:

- Subject based grouping is displayed in a table
- Location based data is displayed on a map
- Time based information is shown on a linear timeline

A.0.0.5 RDF Gravity

RDF Gravity[2] is a visual RDF browser, which presents the underlying RDF data in a graph based visualisation, reflecting the structure of the data. Information regarding the nodes is presented using a label next to the node of the graph. The user is capable to zoom in and out of the graph, allowing to see an overview or a detailed view of a region of interest.

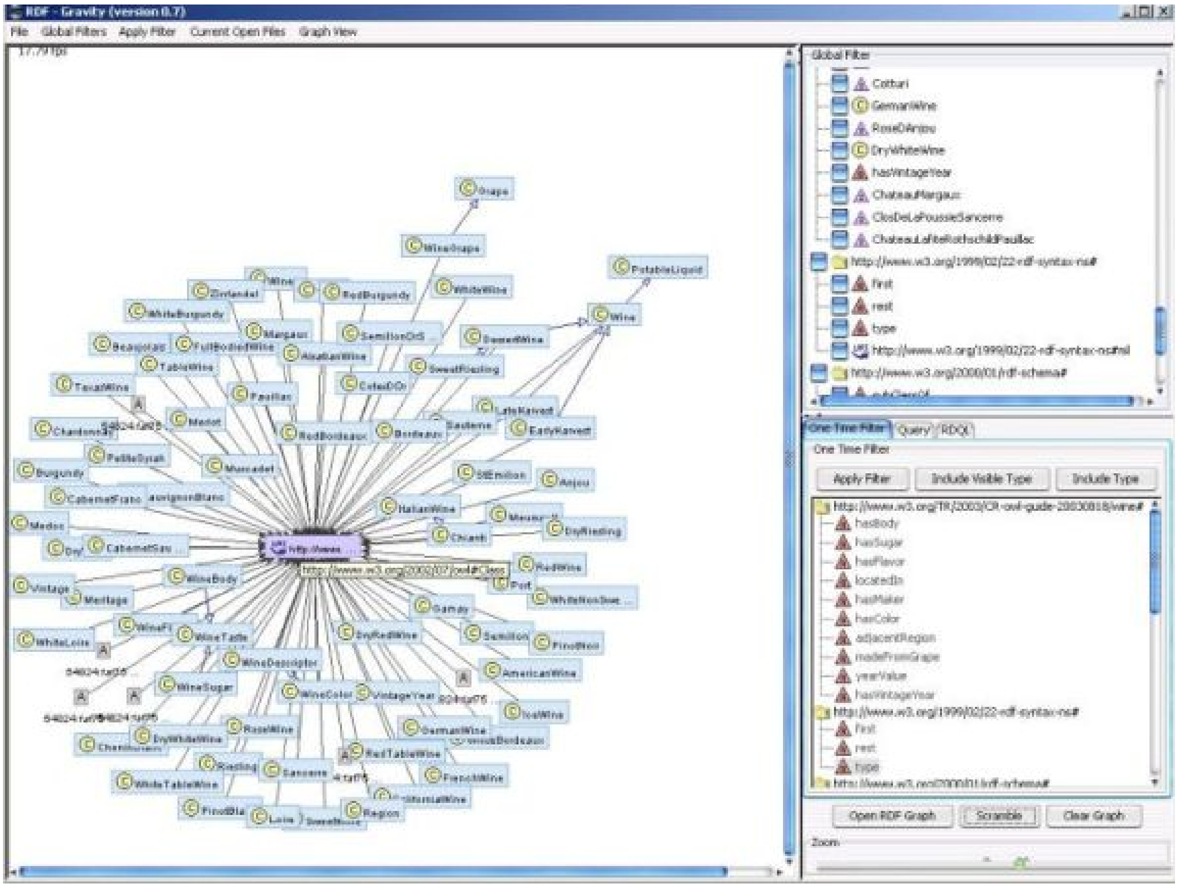


Figure A.3: Image of the user interface of *RDF Gravity*

In addition to the visualisation of the data, the application allows for users to search the presented graph using keyword based search or SPARQL queries.

A.0.0.6 RelFinder

RelFinder[3] provides graphical visualisation of the relationship between two semantic web nodes. The application aims to find a path between two specified URIs in the linked data web. The user is required to provide information such as the two URIs and a query end point in order for the application to work, due to this requirement, the application is less non-expert user friendly, as most non-expert users are not aware of the existence of these concepts.

The application focuses on the discovery of the link between two resources and does not support well the exploration of the data by the user.

[2]http://semweb.salzburgresearch.at/apps/rdf-gravity/
[3]http://www.visualdataweb.org/relfinder.php

Appendix B

Application Screenshots

B.1 Dekstop application

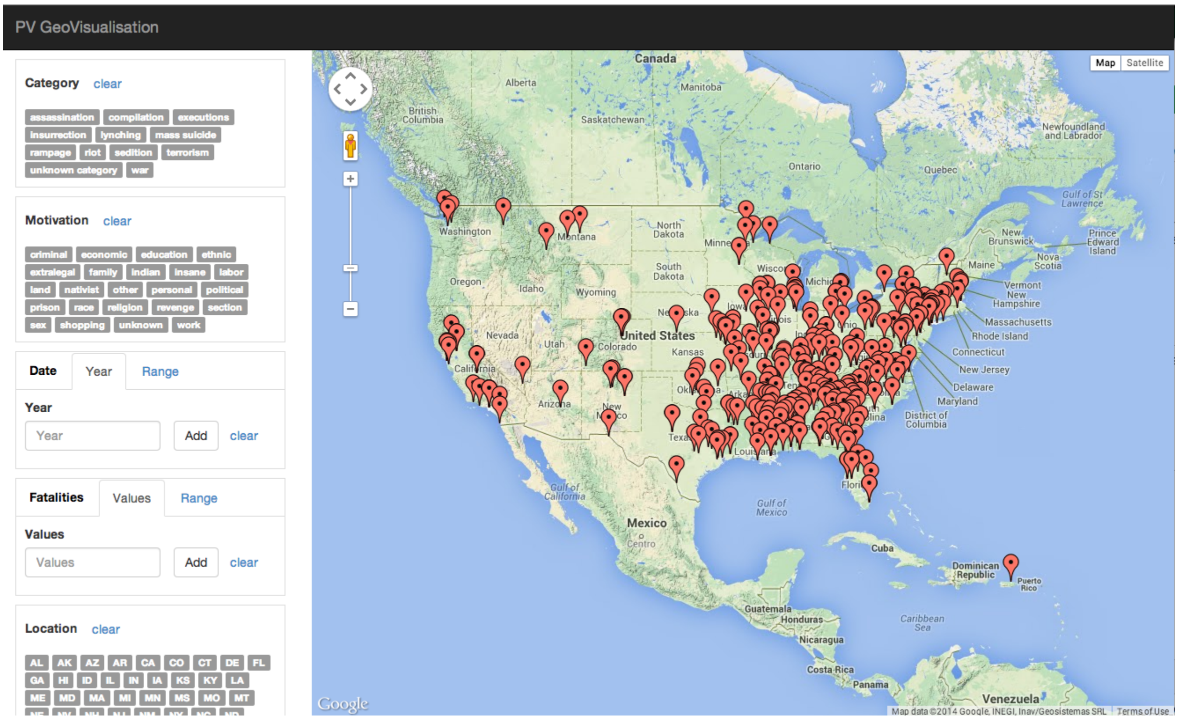


Figure B.1: An overview creenshots of the desktop USPV visualisation application

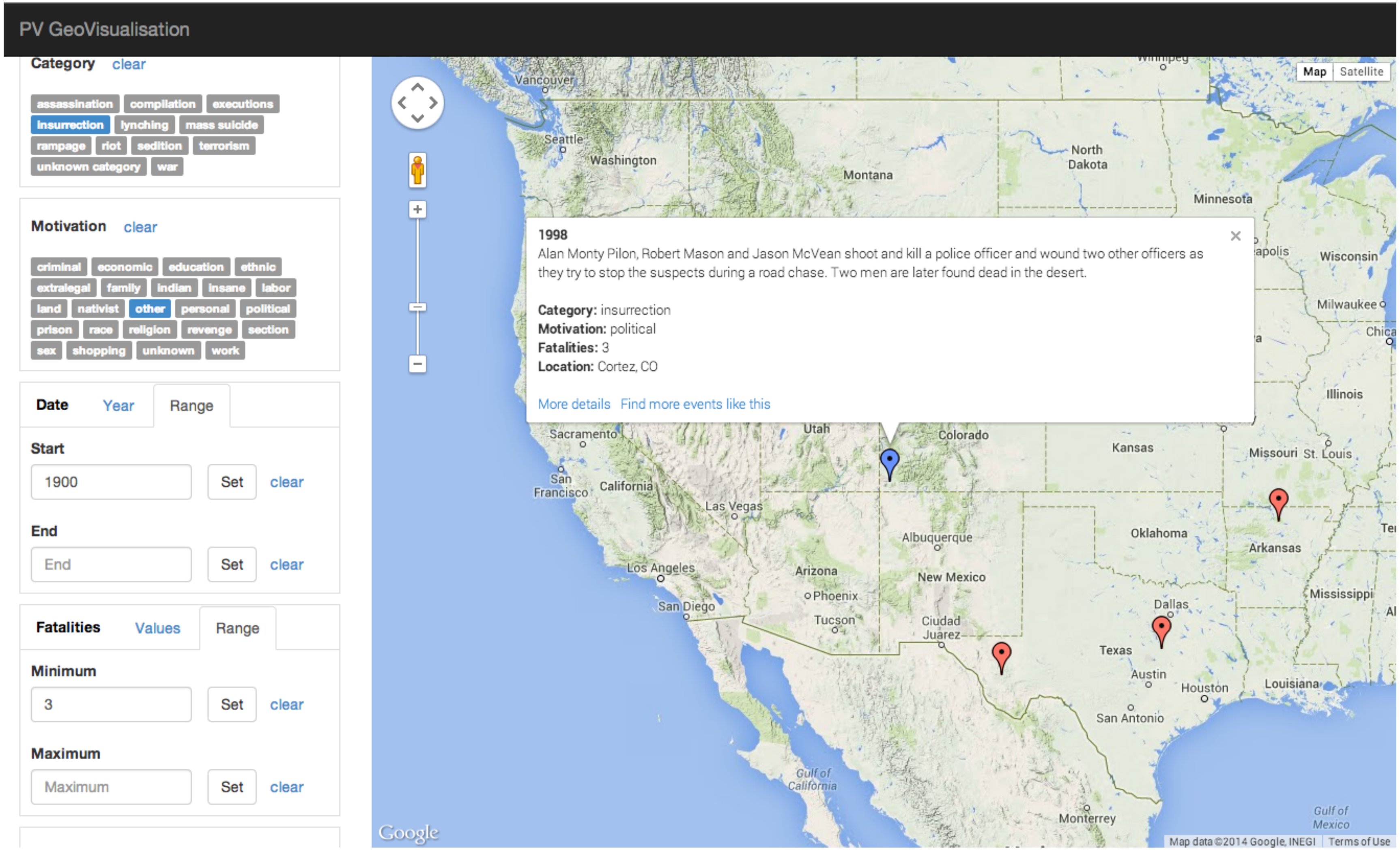


Figure B.2: Screenshot of the desktop USPV visualisation application with an event callout

B.2 Mobile Application: Prorotype 1

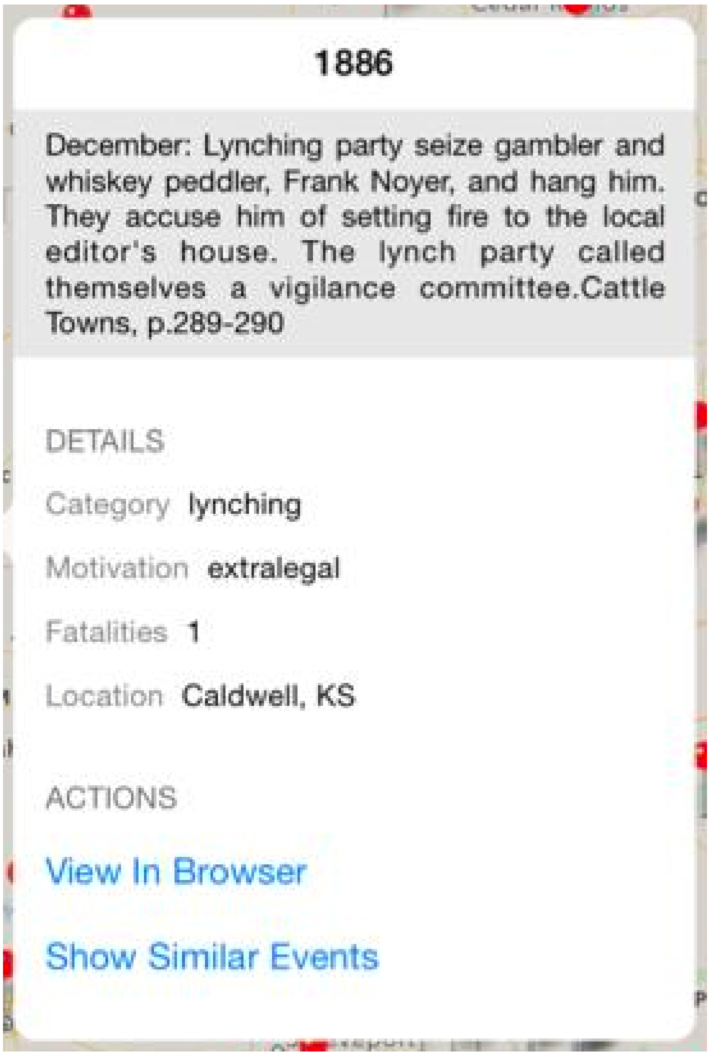


Figure B.3: Screenshot of the event marker callout of PVGeoVisualisation mobile

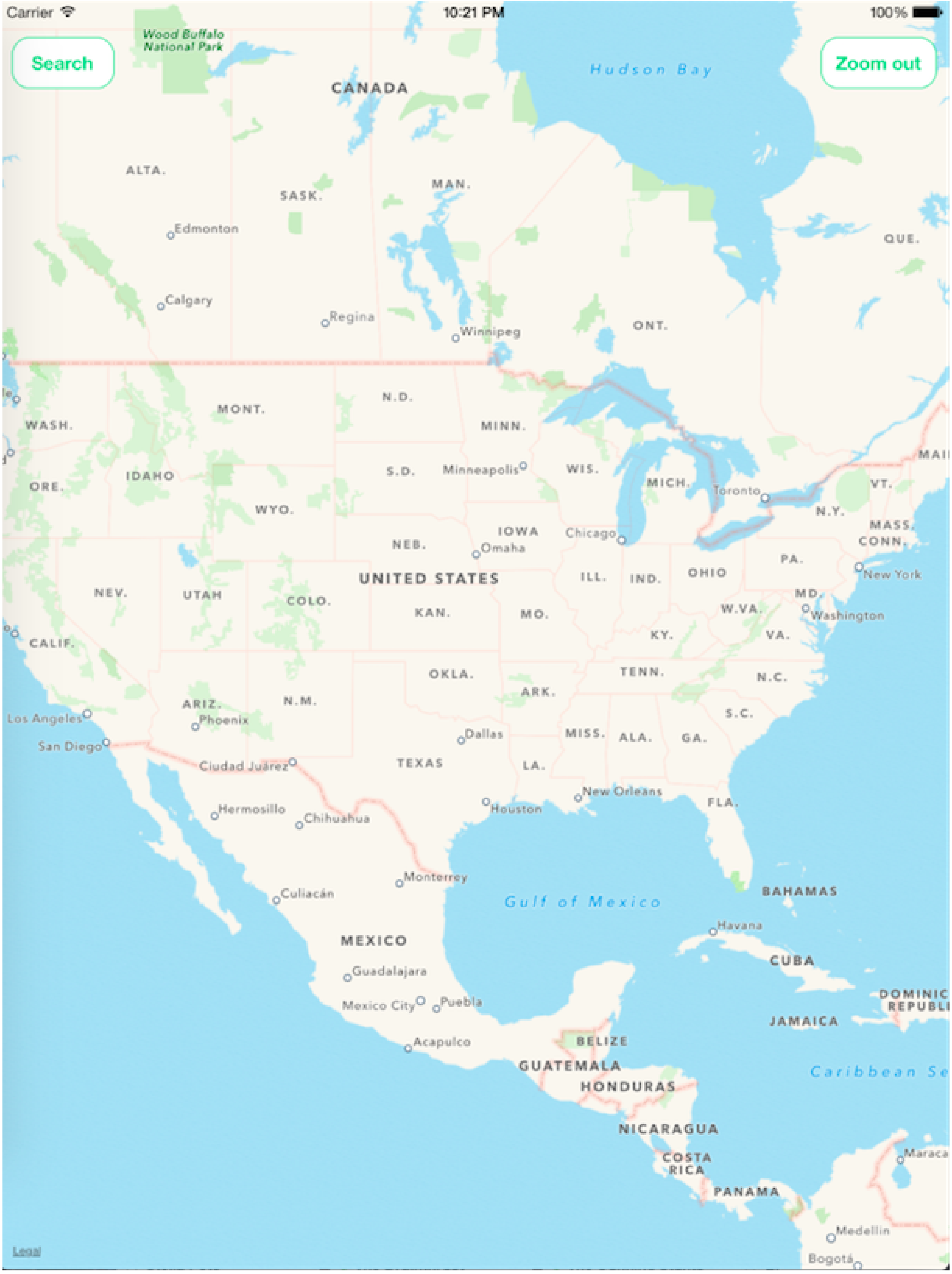


Figure B.4: Screenshot of the map UI (P1) of PVGeoVisualisation mobile

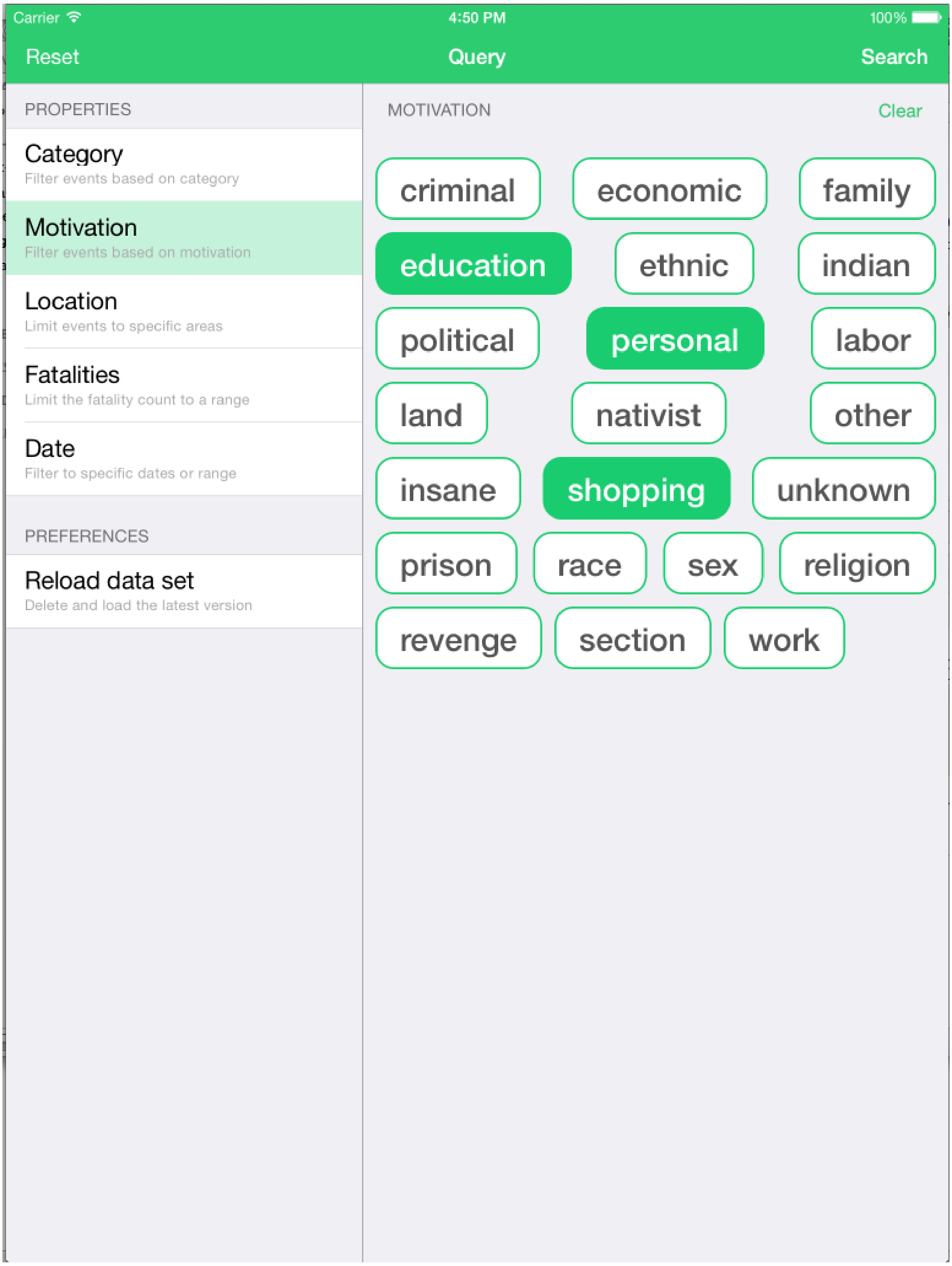


Figure B.5: Screenshot of UI1 (P1) of PVGeoVisualisation mobile

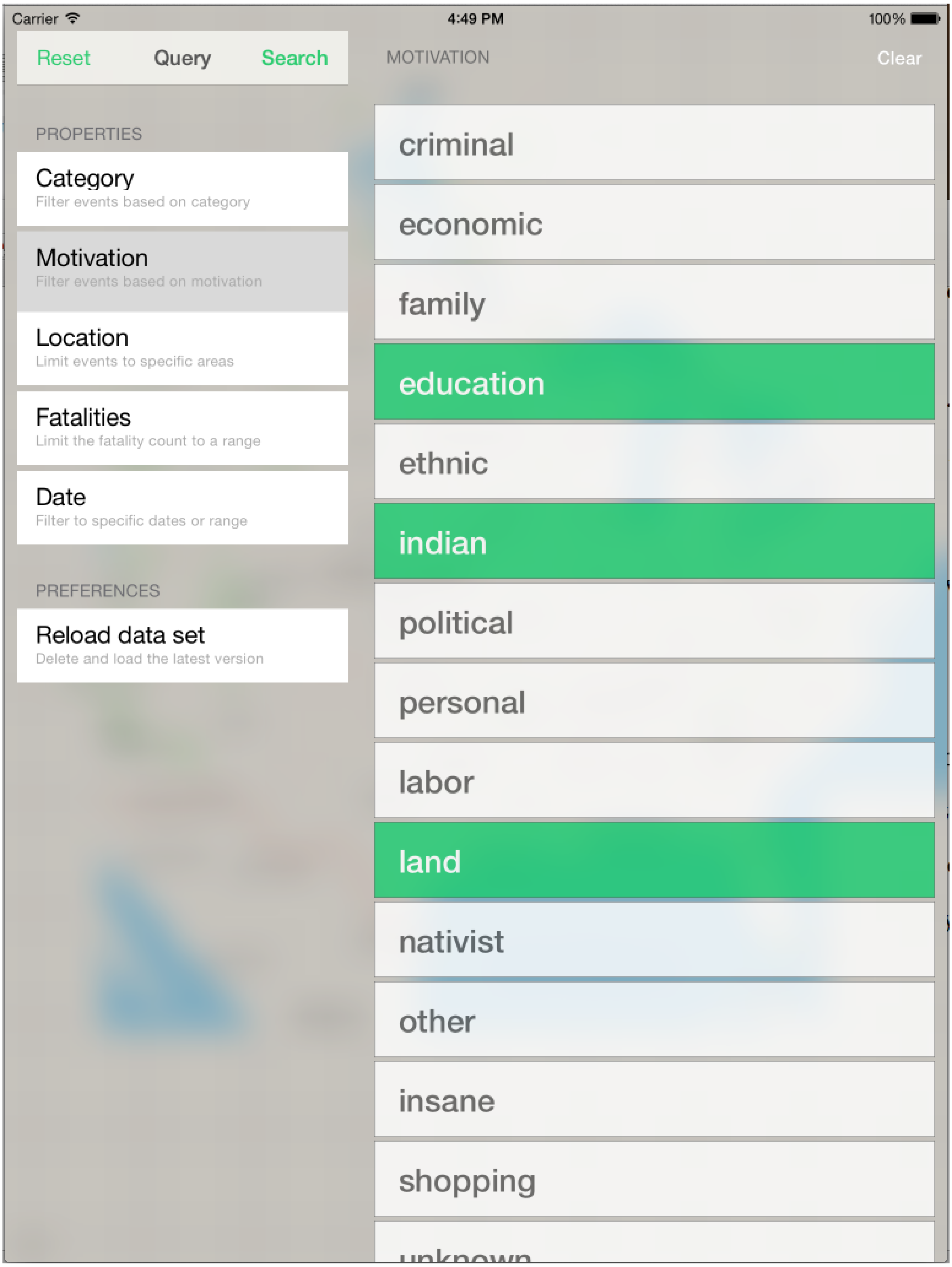


Figure B.6: Screenshot of UI2 (P1) of PVGeoVisualisation mobile

B.3 Mobile Application: Prorotype 2

Figure B.7: Screenshot of the map view (P2) of PVGeoVisualisation mobile

Figure B.8: Screenshot of the query UI (P2) of PVGeoVisualisation mobile

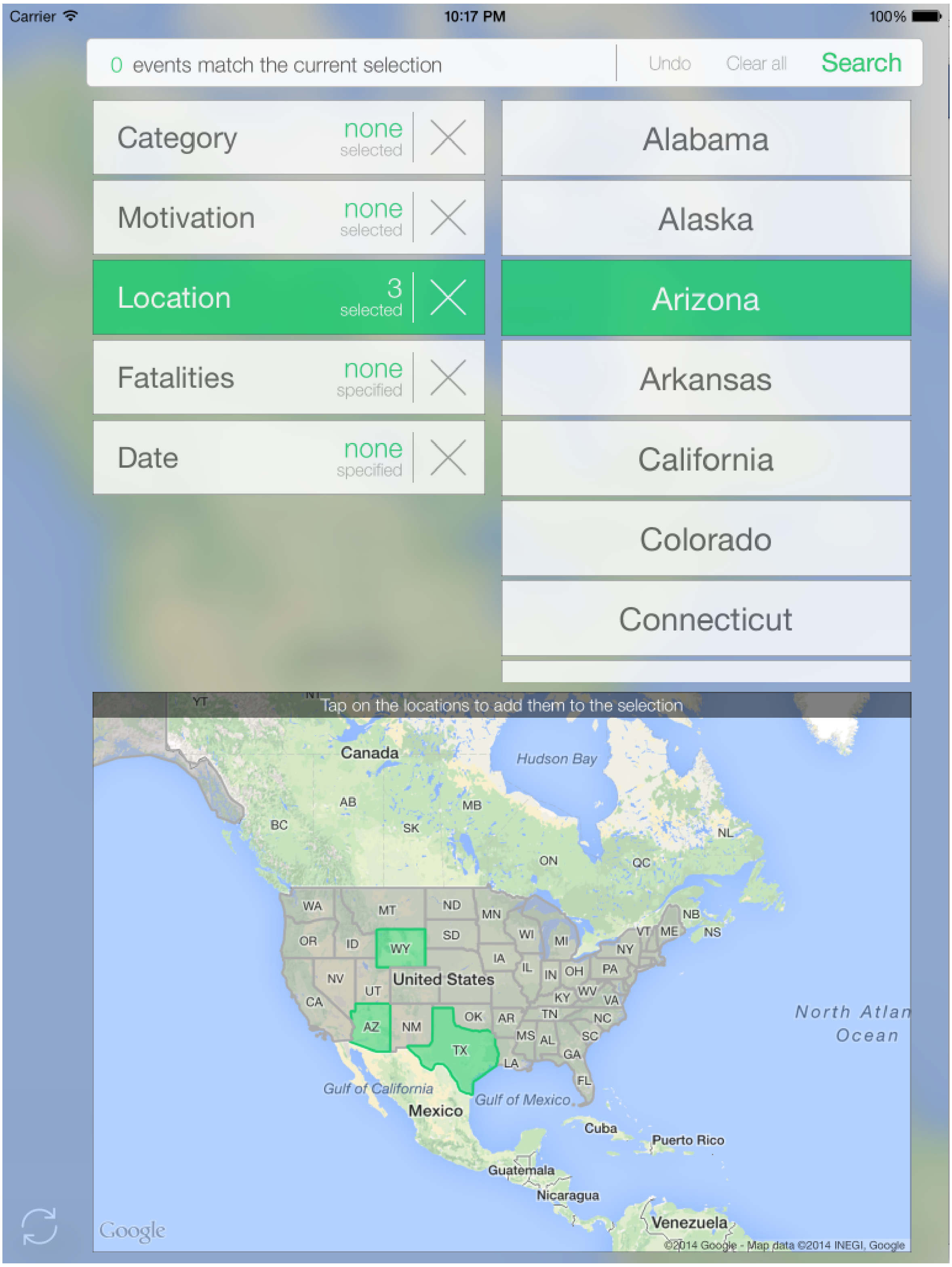


Figure B.9: Screenshot of the location picker in the Query UI (P2) of PVGeoVisualisation mobile

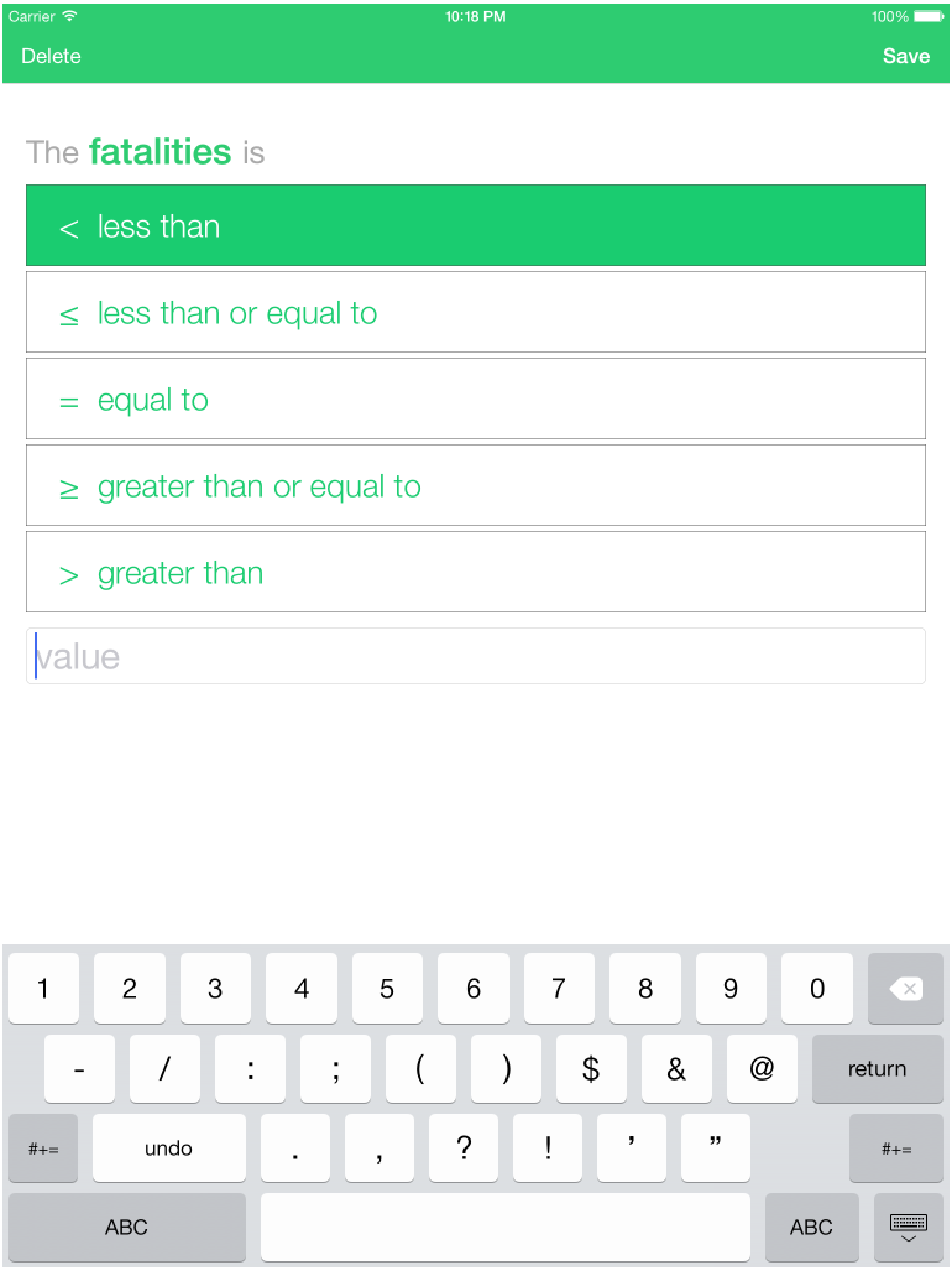


Figure B.10: Screenshot of the condition picker (P2) of PVGeoVisualisation mobile

B.4 Mobile Application: Prorotype 3

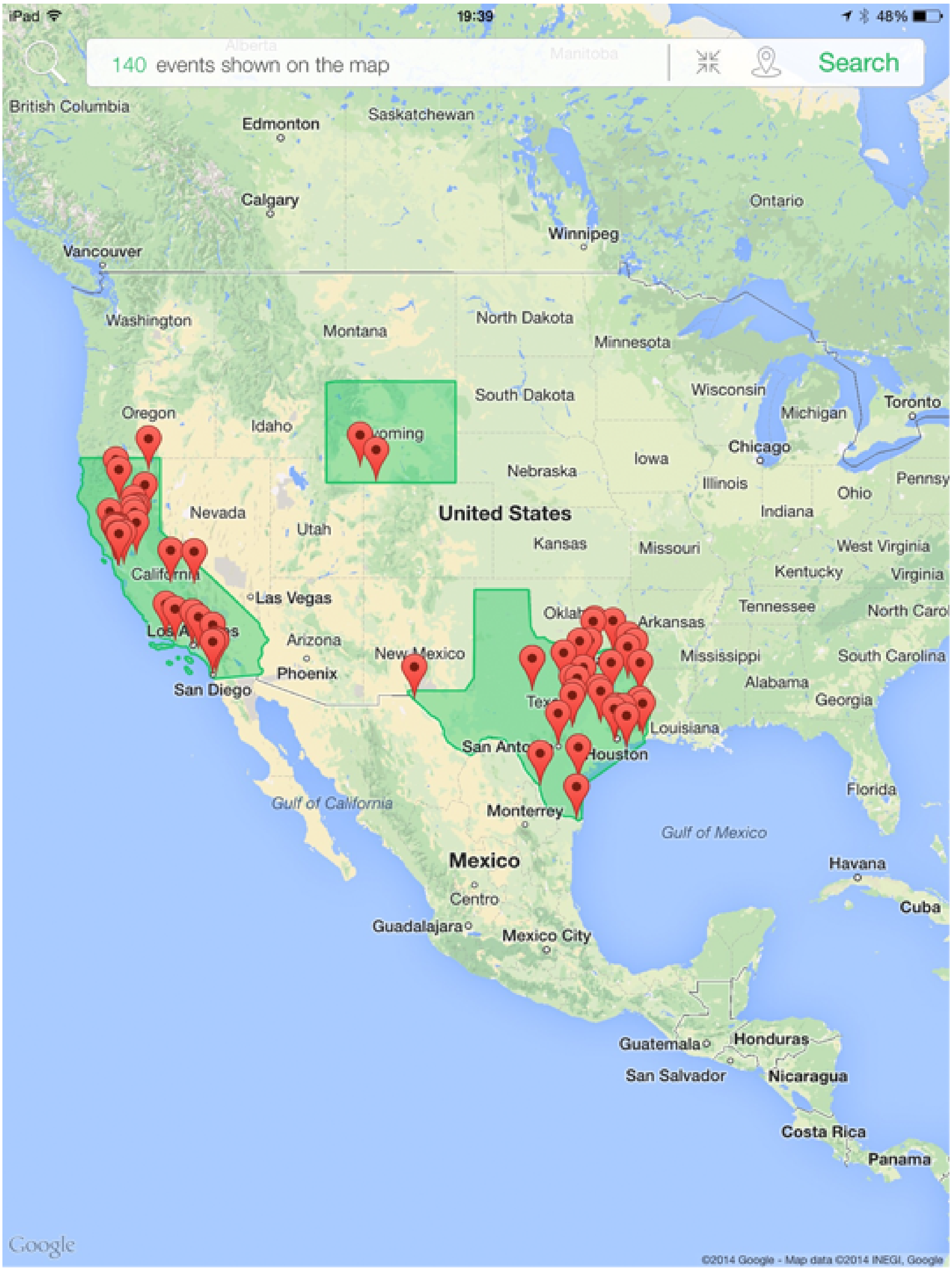


Figure B.11: Screenshot of the map view (P3) of PVGeoVisualisation mobile

Figure B.12: Screenshot of the query UI (P3) of PVGeoVisualisation mobile

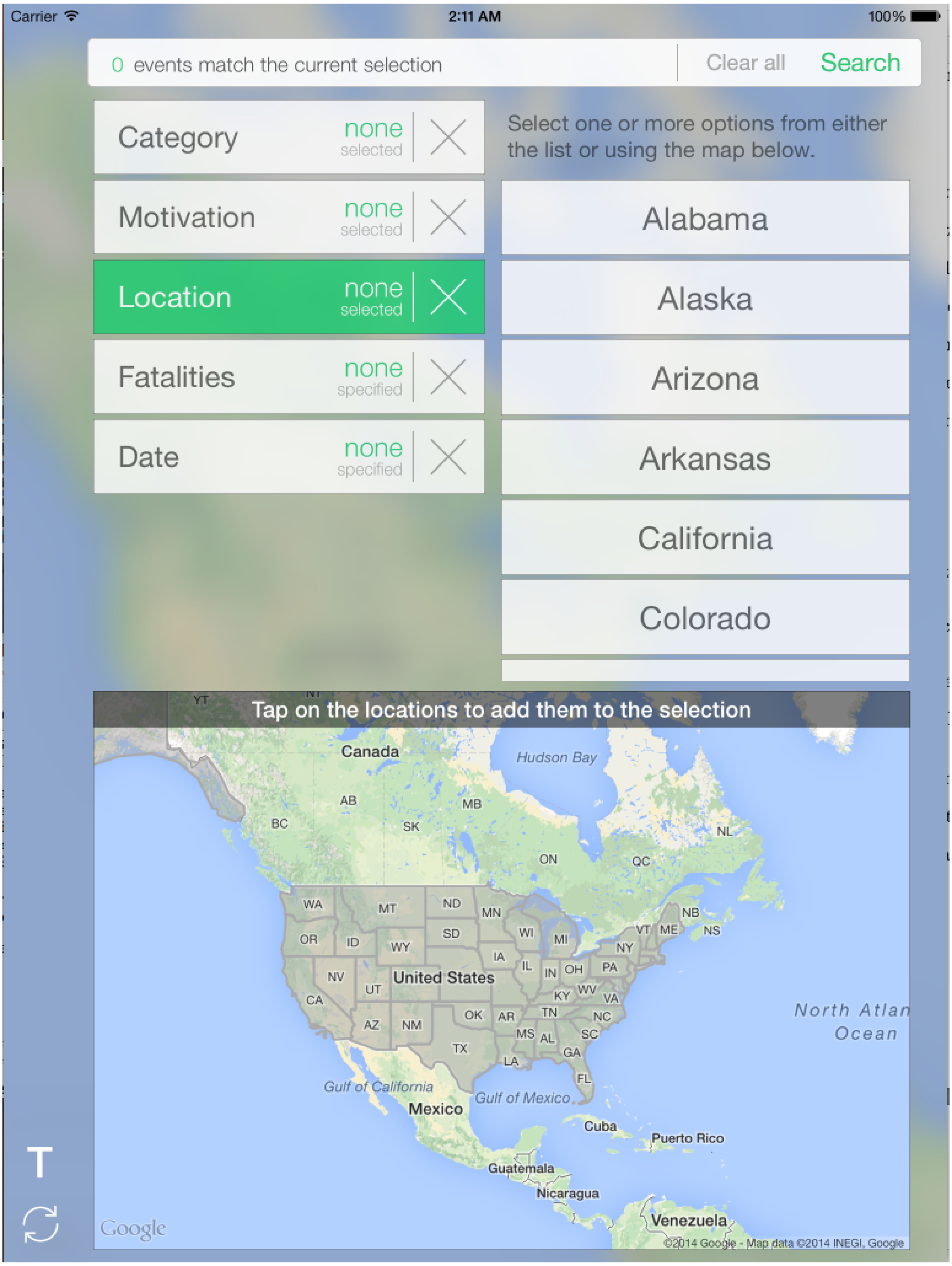


Figure B.13: Screenshot of the location picker in the Query UI (P3) of PVGeoVisualisation mobile

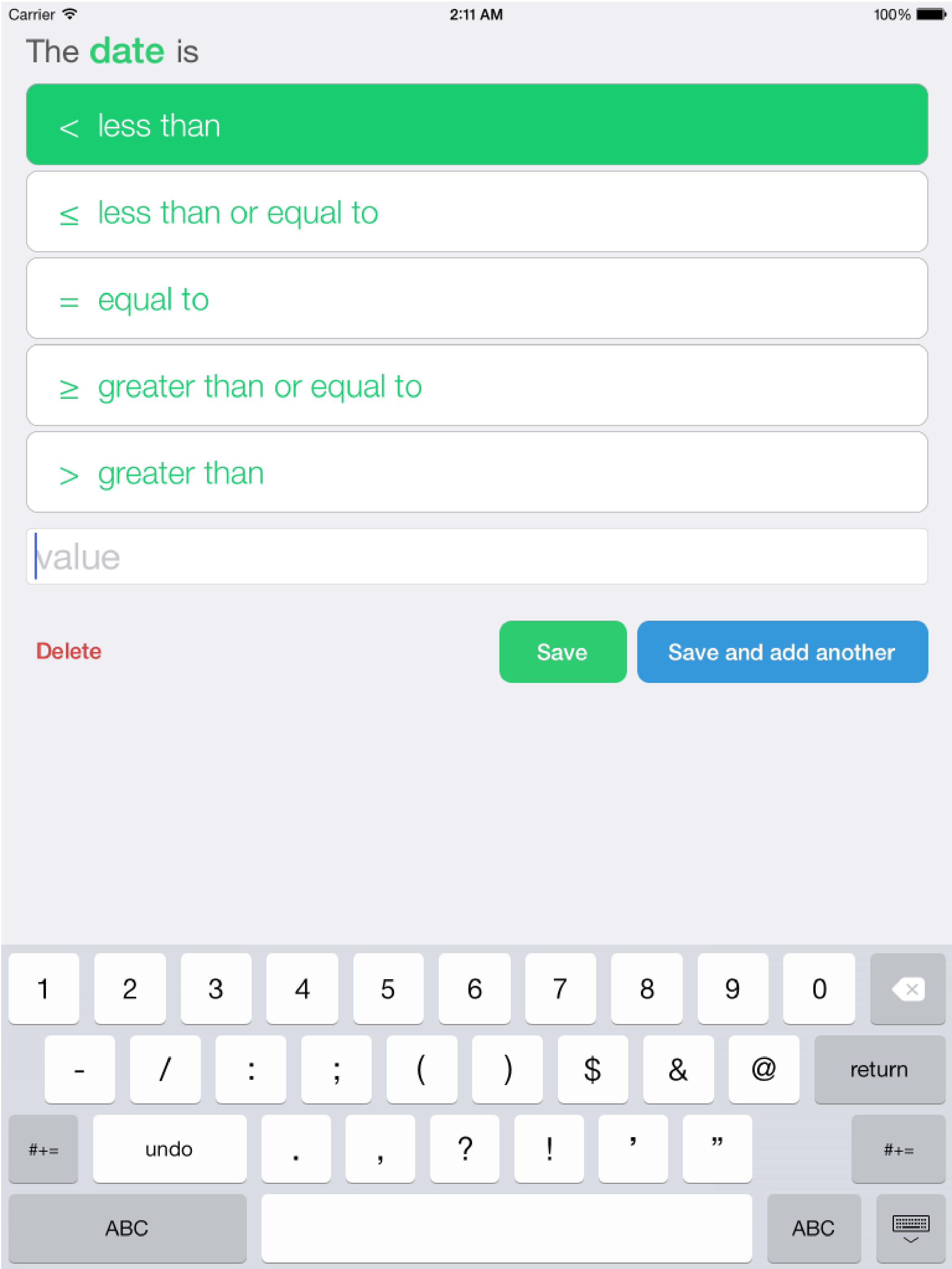


Figure B.14: Screenshot of the condition picker (P3) of PVGeoVisualisation mobile

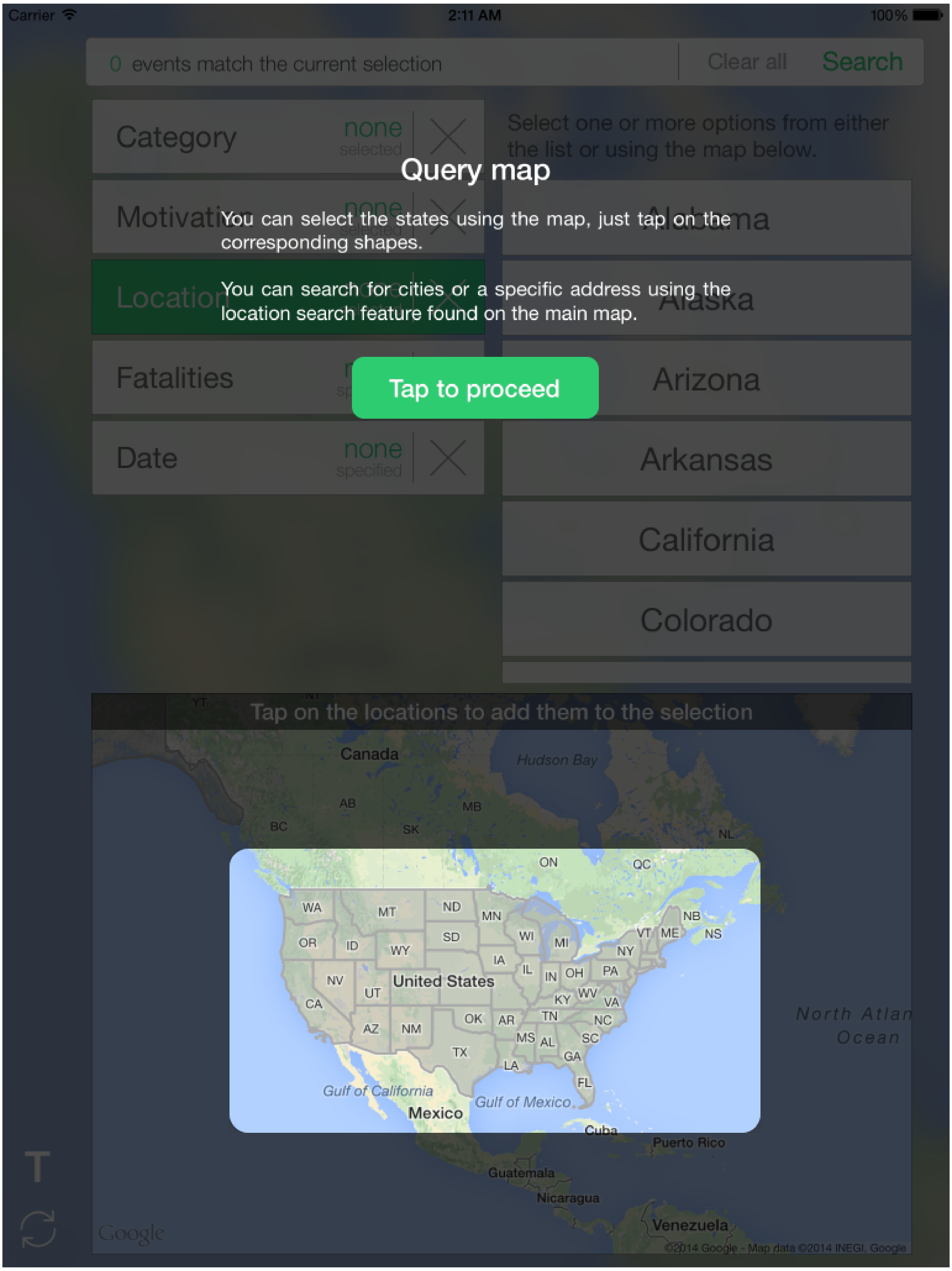


Figure B.15: Screenshot of a tutorial dialog (P3) of PVGeoVisualisation mobile

Appendix C

List of Tasks per Each Experiment

This section presents the list of tasks per each experiment.

C.1 Experiment 1

Task 1

In a prison riot in 1959, 2 inmates, 3 guards, and the deputy warden have been killed. Where did this event occur?

Task 2

How many people have been killed due to events with religious motivation in California after 1990?

Task 3

Who was assassinated in the second part of the 20th century around the Washington Metropolitan Area?

C.2 Experiment 2

Task 1

Find all 20th century events that have occurred in a costal city of California. How many such events were there? And how many of them had a political motivation?

Task 2

The Kansas City Metropolitan Area includes the two Kansas Cities from Kansas and the neighbouring state, as well as the nearby cities. Which motivation for these events was the most common?

Task 3

Route 65 is a north-to-south US highway from Gary, Indiana to Mobile, Alabama. Count the number of fatalities along the route which have occurred either in the 19th or the 20th century.

C.3 Experiment 3

Practice Task

Find all assassinations with extralegal motivations. When did they occur?

Tablet Task 1

Consider the states on the west coast of the US in the 21st century. How many events were there where more than one person has died? What was the most common motivation of these events?

Tablet Task 2

Route 5 is a North to South US highway from the Canadian border in Washington state to San Diego in California. How many riots were there along the route in the frst part of the 20th century? How many people have died due to these events?

Desktop Task 1

Consider the states that share a border with Mexico in the 20st century. How many terrorism related events were there where more than 2 people have died? What was the total fatality count?

Desktop Task 2

Route 25 is a North to South US highway from Buffalo, Wyoming to Las Cruces, New Mexico. How many riots were there along the route in the frst part of the 20th century? How many people have died due to these events?

Appendix D

Documents Used in the Experiments

D.1 Conscent Form and Informatio Sheet

TRINITY COLLEGE DUBLIN
INFORMATION SHEET FOR PARTICIPANTS

Usable Mobile Geographical Linked Data Visualisation

Background of the Research

Information visualisation allows users to consume complex data much more easily, as information is presented in such a way that underlying relationships and concepts can be identified intuitively.

Linked data is a way to access and use information over the World Wide Web that is similar to that of webpages. Each element, concept has a link, like a webpage, and each element has links to other elements.

Currently there is minimal research into the visualisation of linked data that is geographical in nature, and there is even less focus on the usability of such visualisations.

As mobile devices become more ubiquitous, the availability of tools that are present on traditional environments will be come more necessary.

The purpose of the project is

- To investigate the difference in the usability of traditional and mobile geographical linked data visualisation tools.
- To design an approach to visualising geographical information obtained from linked data end points, focusing on the usability and the user experience on mobile devices.

The project aims to contribute to the linked data research area by providing

- An approach to visualising geographical linked data sets on mobile devices
- A tool to visualise geographical linked data sets on mobile based on the developed approaches
- An evaluation methodology for the visualisation of geographical linked data on mobile devices
- Experimental results in optimal visualisations and interface design for linked data visualisation and exploration on mobile

Procedures of this Study

In this study you will be interacting with two applications, one on a laptop and one on a tablet. You will be requested to execute a few tasks on each application.

During the experiment, you will be asked to think aloud and say what your intentions are, what you are currently thinking and explain any difficulties that you may encounter. The investigator may ask you questions or request some explanation of your actions. The information that is gathered during your interaction with the system allows the researchers to identify issues and areas where further focus is required. **You are not required to answer** these questions and **you may at any point request the experiment to stop**. The investigator will aim to interfere with your experience as little as possible, this means that you will not receive a lot of help from the investigator as the aim is to assess whether the application can be understood by users without outside help. **This is not an assessment of your capabilities** to use the software, but an assessment of the ease of use of the application.

The investigator will be taking notes throughout the experiment, which may include your actions, issues you've encountered or any comments that you have made. This information helps the researchers to evaluate the application and the user experience. **You may comment on your experience at any point**, and it is welcomed from you to do do.

Your actions with the applications during the experiment may be recorded through software. **Your actions will only be recorded if you explicitly agree to it**.

You will be asked to execute a series of tasks on two separate applications, after concluding your tasks with one application, you will be asked to fill out a short survey regarding your experience with the application. The next stage of the experiment involves the execution of similar tasks on the other application, then once again you will be asked to fill out the survey. Your answers to the survey will allow the researchers to compare the two applications.

At the end of the experiment, the investigator will ask you to comment on your experience with the applications and the experiment. You may voice any additional issues, suggestions, or comments that you think will be beneficial to the study.

You may request a copy of your data from the researchers and it will be given to you.

Conflicts of Interest

You are requested to spare your own time for the purposes of this investigation. The investigators will ask you to interact with an application that you may be unfamiliar with and may cause you to be uncomfortable in from of another person.

You will be asked questions regarding your interaction with the applications and in addition it may be recorded if you agree to it.

In case you are acquainted with one of the researchers, s(he) is using your relationship in order to facilitate the recruitment procedure and to obtain the desired amount of participants for the study.

The information collected from you will be aggregated and used for the purposes of this research.

Right to Withdraw

Your participation in this study is voluntary and you have the right to withdraw from the study at any time you wish.

You may request the researchers for the deletion of all information that has been collected from you throughout the research at any time.

Duration of Participant's Involvement

This experiment and your involvement will last for about 20 to 30 minutes. You may be asked to attend another session of the experiment at a later stage of the study, which will involve the same exercises that you are asked to do now.

Your participation in this study is voluntary and you may decide to stop collaborating at any time.

Anticipated Risks

You will be interacting with computers (a laptop and a tablet) that present and visualise information on the screen. This may cause episodes of epilepsy.

Provisions for Debriefing

Prior to beginning the experiment you will be debriefed regarding the purpose of the experiment and the procedures that you are about to undertake. The investigator will clearly state that your participation is voluntary and that you may decide to not answer questions or stop the experiment at any time. The investigator will ask you whether you agree to the recording of your actions.

You will be briefed regarding the tasks you will be asked to do as well as the questionnaires that you will be asked to complete.

Just before beginning the experiment the investigator will walk you through this document and will explain and ask you to sign the consent form. You may ask questions regarding the experiment at any time.

Preservation of Participant's Anonymity

Your information will be kept and used anonymously. Individual results will be aggregated anonymously and the research will report on the aggregate results. At no point will the research refer to you or data that is specific to you.

Your information will be used in line with the Data Protection Acts of 1988 and 2003 and the college's Data Protection Policy. All information collected will be destroyed at the end of the study and no part of this information will be shared with any third party.

Illicit Activities

In the extremely unlikely event that illicit activity is reported the researchers will be obliged to report it to appropriate authorities.

Recordings

During your interaction with the prototypes and the applications, your actions will be recorded using on-screen recording software, if you agree to it. You will be briefed regarding these recordings and you will be asked for your consent before beginning the experiments. **You may opt out of the collection of this information**.

If you agree, the application will capture and record your interaction with the application allowing the researchers to replay your interaction with the system. This will allow the researchers to create more realistic simulations for testing purposes in later stages of the study and the development of the tablet application.

These recordings or any part of them will not be shared with any third party and all copies will be destroyed once the study has concluded. No such recordings will be replayed in any public forum or presentation of the research.

Selection Procedure of Participants

The research aims to have a diverse set of participants within the 18-65 age group. Participants consists of both domain experts, people who are knowledgeable of linked data, and non-experts, people who are not aware of linked data or its structure.

The research aims to have at least 10 participants for the experiments at each stage of the study. This is a minimum requirement in order to eliminate any bias, and to be able to have a wide set of different use-cases and interactions with the application which would allow for issues with the application to surface. This number of participants allows the researchers to identify and verify the existence of any recurrent and common issues.

You have been selected as you fit these criteria. You are either a friend of college of one of the researchers, and you have been contacted and asked to participate in the study through an internal mailing list, such as the KDEG mailing list, email, social media, or in person.

In addition, you do not have a medical condition that may be induced or aggregated by the experiments, such as epilepsy.

TRINITY COLLEGE DUBLIN
INFORMED CONSENT FORM

Usable Mobile Geographical Linked Data Visualisation

Lead Researchers

Balazs Pete
Supervisor: Rob Brennan

Background of the Research

Information visualisation allows users to consume complex data much more easily, as information is presented in such a way that underlying relationships and concepts can be identified intuitively.

Linked data is a way to access and use information over the World Wide Web that is similar to that of webpages. Each element, concept has a link, like a webpage, and each element has links to other elements.

Currently there is minimal research into the visualisation of linked data that is geographical in nature, and there is even less focus on the usability of such visualisations.

As mobile devices become more ubiquitous, the availability of tools that are present on traditional environments will be come more necessary.

The purpose of the project is

- To investigate the difference in the usability of traditional and mobile geographical linked data visualisation tools.
- To design an approach to visualising geographical information obtained from linked data end points, focusing on the usability and the user experience on mobile devices.

The project aims to contribute to the linked data research area by providing

- An approach to visualising geographical linked data sets on mobile devices
- A tool to visualise geographical linked data sets on mobile based on the developed approaches
- An evaluation methodology for the visualisation of geographical linked data on mobile devices
- Experimental results in optimal visualisations and interface design for linked data visualisation and exploration on mobile

Procedures of this Study

In this study you will be interacting with two applications, one on a laptop and one on a tablet. You will be requested to execute a few tasks on each application.

During the experiment, you will be asked to think aloud and say what your intentions are, what you are currently thinking and explain any difficulties that you may encounter. The investigator may ask you questions or request some explanation of your actions. The information that is gathered during your interaction with the system allows the researchers to identify issues and areas where further focus is required. You are not required to answer these questions and may at any point request the experiment to stop. The investigator will aim to interfere with your experience as little as possible, this means that you will not receive a lot of help from the investigator as the aim is to assess whether the application can be understood by users without outside help. This is not an assessment of your capabilities to use the software, but an assessment of the ease of use of the application.

The investigator will be taking notes throughout the experiment, which may include your actions, issues you've encountered or any comments that you have made. This information helps the researchers to evaluate the application and the user experience. You may comment on your experience at any point, and it is welcomed from you to do do.

Your actions with the applications during the experiment may be recorded through software. Your actions will only be recorded if you explicitly agree to it.

You will be asked to execute a series of tasks on two separate applications, after concluding your tasks with one application, you will be asked to fill out a short survey regarding your experience with the application. The next stage of the experiment involves the execution of similar tasks on the other application, then once again you will be asked to fill out the survey. Your answers to the survey will allow the researchers to compare the two applications.

At the end of the experiment, the investigator will ask you to comment on your experience with the applications and the experiment. You may voice any additional issues, suggestions, or comments that you think will be beneficial to the study.

You may request a copy of your data from the researchers and a copy will be given to you.

Publication

This research is part of the dissertation for the MSc in Computer Science (Mobile and Ubiquitous Computing) course. The research report will be available through the Trinity College Library.

The initial progress and results of the research may be submitted to the 1st International Workshop on Geospatial Linked Data (GeoLD 2014) in conjunction with the annual SEMANTiCS conference.

The dissertation may also be submitted for publication at the International Semantic Web Conference or the European Semantic Web Conference.

Individual results will be aggregated anonymously and the research will be reported on the aggregate results.

Declaration

1. I am 18 years or older and am competent to provide consent.
2. I have read, or had read to me, a document providing information about this research and this consent form.
3. I have had the opportunity to ask questions and all my questions have been answered to my satisfaction and understand the description of the research that is being provided to me.
4. I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications in a way that does not reveal my identity.
5. I understand that if I make illicit activities known, these will be reported to appropriate authorities.
6. I understand that I may stop electronic recordings at any time, and that I may at any time, even subsequent to my participation have such recordings destroyed (except in situations such as above).
7. I understand that, subject to the constraints above, no recordings will be replayed in any public forum or made available to any audience other than the current researchers/research team.
8. I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
9. I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
10. I understand that my participation is fully anonymous and that no personal details about me will be recorded.
11. I understand that I may opt out of the recording of my interactions with the applications.
12. I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk.
13. I have received a copy of this agreement.

Details of the Participant

- Participant's name:
- Participant's signature:

- Date:
- Do you agree to the recording of your interactions?

Statement of investigators responsibility

I have explained the nature and purpose of this research study, the procedures to be undertaken and any risks that may be involved. I have offered to answer any questions and fully answered such questions. I believe that the participant understands my explanation and has freely given informed consent.

Researchers' Contact Details

Balazs Pete
peteb@tcd.ie
085 706 6460

Rob Brennan
rob.brennan@scss.tcd.ie
01 896 8426

Details of Investigator

- Name:
- Signature:
- Date:

D.2 Usability Evaluation Questionnaire

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

System Usability Evaluation Questionnaire

Section 1- Overall system usability

Please answer the following questions on the scale between *strongly disagree(1)* and *strongly agree(5)* by marking your answer with an **X** in the appropriate box next to the question.

		Strongly Disagree				Strongly Agree
		1	**2**	**3**	**4**	**5**
1	I think I would like to use this application frequently					
2	I found the application unnecessarily complex					
3	I though the application was easy to use					
4	I think that I would need the support of a technical person to be able to use this application					
5	I found the various functions in the application were well integrated					
6	I thought there was too much inconsistency in this application					
7	I imagine that most people would learn to use this application very quickly					
8	I found the application very awkward to use					
9	I felt very confident using the application					
10	I needed to learn a lot of things before I could get going with this application					

Section 2- Components specific usability

Please answer the following questions on the scale between *bad(1)* and *good*(5) by marking your answer with an **X** in the appropriate box next to the question.

		Bad				Good
		1	**2**	**3**	**4**	**5**
Learning						
1	The application used familiar language and terms.					
2	I understood the information the application presented. I did not have to take unnecessary steps or actions.					
Adapting to the User						

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

		1	2	3	4	5
3	The application allowed me to undo actions or to go back and correct mistakes.					
4	I was aware which actions were allowed by the application.					
5	The application responded to my actions as I have expected it to.					
6	I did not have to remember a lot of my previous actions or decisions.					
Feedback and Errors						
7	The application gave appropriate feedback on my actions.					
8	I have made a lot of errors while using the application					
9	The application showed helpful error messages, helping me to find the issues.					

Section 3- *Task specific questionnaire*

Please answer the following questions on the scale between *strongly disagree(1)* and *strongly agree(5)* by marking your answer with an **X** in the appropriate box next to the question.

		Strongly Disagree				Strongly Agree
		1	**2**	**3**	**4**	**5**
1	Overall, I am satisfied with the ease of completing the tasks					
2	Overall, I am satisfied with the amount of time it took to complete the tasks					
3	Overall, I am satisfied with how well I have completed the tasks					
4	Overall, I am satisfied with how well the application helped me in completing the tasks					

D.3 Worksheets

In this section the worksheets used int he experiments are presented.

Experiment 1 worksheet

Question 1

In a ***prison riot*** in ***1959***, 2 inmates, 3 guards, and the deputy warden have been killed. Where did this event occur?

Hint: The terms in bold may be useful in your search. There have been 6 deaths in total.

	Answer:

Question 2

How many people have been killed due to events with religious motivation in California after 1990?

Hint: Count the fatalities.

	Answer:

Question 3

Who was assassinated in the second part of the 20^{th} century around the Washington Metropolitan Area?

Hint: The Washington Metropolitan Area comprises of the District of Columbia and the nearby cities from the neighbouring states.

	Answer:

Experiment 2 worksheet

Question 1

Find all 20th century events that have occurred in a costal city of California. How many such events were there? And how many of them had a political motivation?

Hint: You must determine whether a city is coastal by looking at the map.

	Answer:

Question 2

The Kansas City Metropolitan Area includes the two Kansas Cities from Kansas and the neighbouring state, as well as the nearby cities. Which motivation for these events was the most common?

Hint: Use the map. Consider all cities in the nearby counties.

	Answer:

Question 3

Route 65 is a north-to-south US highway from Gary, Indiana to Mobile, Alabama. Count the number of fatalities along the route which have occurred either in the 19th or the 20th century.

Hint: Use the pinch-to-zoom gesture to get more or less details on the map.

	Answer:

Experiment 3 worksheet – Practice

Question

- Find all assassinations with extralegal motivations. When did they occur?

	Answer:

Experiment 3 worksheet – Computer

Question 1

Consider the states that share a border with Mexico in the 20st century. How many terrorism related events were there where more than 2 people have died? What was the total fatality count?

	Answer:

Question 2

Route 25 is a North to South US highway from Buffalo, Wyoming to Las Cruces, New Mexico. How many riots were there along the route in the first part of the 20th century? How many people have died due to these events?

	Answer:

Experiment 3 worksheet – Tablet

Question 1

Consider the states on the west coast of the US in the 21st century. How many events were there where more than one person has died? What was the most common motivation of these events?

	Answer:

Question 2

Route 5 is a North to South US highway from the Canadian border in Washington state to San Diego in California. How many riots were there along the route in the first part of the 20th century? How many people have died due to these events?

	Answer:

Appendix E

Experiment Participant Information

E.1 Experiment 1

Table E.1: Age group based breakdown of the participants of Experiment 1

Age group	**Number of participants**
18 - 20	1
21 - 25	7
45 - 50	1
55 - 60	1
TOTAL	**10**

Table E.2: Participant breakdown based on familiarity with similar tools in Experiment 1

Familiar with similar tools	**Number of participants**
Yes	5
No	5
TOTAL	**10**

E.2 Experiment 2

Table E.3: Age group based breakdown of the participants of Experiment Two

Age group	Number of participants
18 - 20	1
21 - 25	2
26 - 30	5
55 - 60	1
TOTAL	**9**

Table E.4: Participant breakdown in each group based on whether they have prior experience with the application or not

	Prior experience with the app	No prior experience with the app
Novice users	4	0
Expert users	0	5

E.3 Experiment 3

Table E.5: Age group based breakdown of the participants of Experiment Two

Age group	**Number of participants**
21 - 25	12
26 - 30	4
TOTAL	**16**

Table E.6: Breakdown of the number of users based on skill level and whether they have been new to the tablet application or not

	Previous users	**New users**	**TOTAL**
Novice users	6	6	12
Expert Users	4	0	4
TOTAL	10	6	**16**

Appendix F

Possible Future Features for PVGeoVisualisation Mobile

This section presents the possible features that may be added to the prototype to increase its functionality and usability. However whether these features are needed and should be added would be shown by the results of the usability studies.

F.1 Map overview

It was shown by Burigat et al. [10] that Overview+Detail is beneficial for users in realising the presence of off-screen elements and that is was better than other approaches such as Wedge or Halo. However the overview pane takes up previous screen space, in addition the amount of data that may be presented by the application is significant and presenting this information twice adds significant load on the amount of processing the device would need to handle.

A possible approach which should be investigated would be the use of a subtle overview techniques, where users would be initially shown an overview of the data, a zoomed out version of the information on the map, initially indicating to them where data points are.

F.2 Visual cues on event markers

In prototype 1, the data points on the map are uniform markers and do not give any indication of the event to users except the location. A possibility would be the use of custom markers which present additional information about events, such as the year or the fatality count. Other option would be the use of different colours to represent different categories or motivations.

However it is not clear which approach would benefit users, hence a possible usability study would be the investigation of the difference between two approaches.

F.3 Visualisation of invalid events

Burigat et al. [9] argues that it was better if geographical applications did not remove data points from the map in case they did not satisfy the visualisation conditions, instead they should be represented differently. It may be worth exploring different visualisation techniques for "invalid" events, however since the data that is presented is multidimensional and heterogeneous in nature it is unclear how this should be achieved.

The events possess properties which are continuous and non continuous in nature, which cause potential issues in the representation of the "invalid" events. In case users define a range for the date of the events, it is difficult to decide whether all "invalid" events should be represented the same way or whether events that satisfy all events but the date should be shown differently. Considering that the same issue is present with the location, the visualisation of these events become increasingly difficult. In addition the number of event in the data set is significant and not eliminating "invalid" events may not benefit users in finding the desired information.

F.4 Other features

Other possible additions would be the ability to run functions on events properties of the selected events. Examples include counting, averaging, etc of the fatality count. To address the amount of markers present on the map, the application could cluster them based on the zoom level of the map, possibly improving the performance of the application.

While the main focus of the application is geographic visualisation, it may be possible that other visualisation techniques would be beneficial for certain use cases.

Appendix G

USPV Geocoding

G.1 USPV SPARQL query

The SPARQL query to execute on the *http://dacura.cs.tcd.ie* Linked Data end point. The results should be requested in the JSON format and saved as ”results.json”, this is the input to the Node.js script in Section G.2.

```
PREFIX pv:<http://dacura.cs.tcd.ie/data/politicalviolence#>
SELECT ?url ?location
FROM <http://dacura.cs.tcd.ie:3030/politicalviolence/sparql>
WHERE {
   ?url
      pv:unstructuredLocation ?location.
}
FROM 0
LIMIT 5000
```

G.2 Node.js script

Generates Turtle output for the results of the SPARQL query in Section G.1. This Turtle file can be used to enrich the USPV data set though a SPARQL UPDATE operation.

```
// The MIT License (MIT)
// Copyright (c) 2014 Balazs Pete
//
// Permission is hereby granted, free of charge, to any person obtaining
// a copy of this software and associated documentation files (the "Soft-
// ware"), to deal in the Software without restriction, including without
// limitation the rights to use, copy, modify, merge, publish, distribute,
// sublicense, and/or sell copies of the Software, and to permit persons
// to whom the Software is furnished to do so, subject to the following
// conditions:
//
// The above copyright notice and this permission notice shall be included
// in all copies or substantial portions of the Software.
//
// THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS
// OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABI-
// LITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT
// SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES
// OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE,
// ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR
// OTHER DEALINGS IN THE SOFTWARE.

var API_KEY = "__THE_API_KEY__";
// Get a Google Geocoding API key from the developers console
// and replace __THE_API_KEY__

var INPUT = "results.json";
var OUTPUT = "output.ttl";
var fs = require('fs');
var https = require('https');
var querystring = require('querystring');

// Load a JSON file from the specified path
function loadJSON(path, callback) {
    fs.readFile(path, function(err, data) {
        if (err) {
            return callback('Error while loading JSON: ' + err);
        }
        var json = JSON.parse(data);
        return callback(null, json);
```

```
    });
}
// Geocode the input address (string)
function geocode(address, callback) {
    var query = {
        address: address,
        sensor: false,
        key: API_KEY
    };
    var path =
        '/maps/api/geocode/json?' +
        querystring.stringify(query);
    var options = {
        hostname: 'maps.googleapis.com',
        port: 443,
        path: path,
        method: 'GET'
    };
    var req = https.request(options, function(res){
        var _data = "";
        res.setEncoding('utf8');
        res.on('data', function(chunk) {
            _data += chunk;
        });
        res.on('end', function() {
            var data = JSON.parse(_data);
            console.log(data);

            if (!!data && data.status == "OK") {
                callback(null, data);
            } else {
                callback(null, null);
            }
        });
    });
    req.on('error', function(err) {
        callback('Failed to send the request to Google: ' + err);
    });
    req.write('');
    req.end();
}
// Geocode the list of addresses
function createGeoData(data, callback) {
    var creator = function(i){
```

```
        if (i == data.length) {
            return callback(null, true);
        }
        var element = data[i];
        var url = element.url.value;
        if (!url) {
            return callback('Unspecified URL in geocode results');
        }
        var location = element.location.value;
        if (!location) {
            return callback('Unspecified location in geocode results');
        }
        geocode(location, function(err, result) {
            if (err) {
                return callback(err);
            }
            if (!result || !result.results.length) {
                return creator(i+1);
            }
            var geometry = result.results[0].geometry;
            var lat = geometry.location.lat;
            var lng = geometry.location.lng;
            fs.appendFileSync(OUTPUT, formatOutputLine({
                url: url,
                lat: lat,
                long: lng
            }));
            creator(i+1);
        });
    }
    creator(0);
}
// Create RDF output
function formatOutputLine(entry) {
    return '<' + entry.url + '> \n'+
    '  geo:lat  ' + entry.lat + ' ;\n' +
    '  geo:long  ' + entry.long + ' .\n'
}
// RUN
loadJSON(INPUT, function(err, data) {
    if (err) {
        return console.log(err);
    }
    fs.appendFileSync(OUTPUT,
```

```
        '@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>\n\n'
    );
    createGeoData(data.results.bindings, function(error, result) {
        if (error) {
            return console.log(error);
        } else {
            console.log('done');
        }
    });
});
```

Bibliography

[1] Gennady L Andrienko and Natalia V Andrienko. Interactive maps for visual data exploration. *International Journal of Geographical Information Science*, 13(4):355–374, 1999.

[2] Sören Auer, Raphael Doehring, and Sebastian Dietzold. Less-template-based syndication and presentation of linked data. In *The Semantic Web: Research and Applications*, pages 211–224. Springer, 2010.

[3] Aaron Bangor, Philip T Kortum, and James T Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[4] Christian Becker and Christian Bizer. Dbpedia mobile-a location-aware semantic web client. *Proceedings of the Semantic Web Challenge at ISWC*, 2008, 2008.

[5] Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked data-the story so far. *International journal on semantic web and information systems*, 5(3):1–22, 2009.

[6] Max Braun, Daniel Schmeiß, Ansgar Scherp, and Steffen Staab. Stevie-collaborative creation and exchange of events and pois on a mobile phone. In *Proceedings of the 2nd ACM international workshop on Events in multimedia*, pages 35–40. ACM, 2010.

[7] Rob Brennan and KEVIN CHEKOV FEENEY. Publishing social sciences datasets as linked data: a political violence case study. 2013.

[8] John Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189:194, 1996.

[9] Stefano Burigat and Luca Chittaro. Interactive visual analysis of geographic data on mobile devices based on dynamic queries. *Journal of Visual Languages & Computing*, 19(1):99–122, 2008.

[10] Stefano Burigat and Luca Chittaro. On the effectiveness of overview+ detail visualization on mobile devices. *Personal and ubiquitous computing*, 17(2):371–385, 2013.

[11] Stefano Burigat, Luca Chittaro, and Silvia Gabrielli. Navigation techniques for small-screen devices: An evaluation on maps and web pages. *International Journal of Human-Computer Studies*, 66(2):78–97, 2008.

[12] Mike Cammarano, Xin Dong, Bryan Chan, Jeff Klingner, Justin Talbot, Alon Halevy, and Pat Hanrahan. Visualization of heterogeneous data. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1200–1207, 2007.

[13] A Elizabeth Cano, Aba-Sah Dadzie, and Melanie Hartmann. Whos who–a linked data visualisation tool for mobile environments. In *The Semanic Web: Research and Applications*, pages 451–455. Springer, 2011.

[14] Luca Chittaro. Visualizing information on mobile devices. *Computer*, 39(3):40–45, 2006.

[15] Aba-Sah Dadzie and Matthew Rowe. Approaches to visualising linked data: A survey. *Semantic Web*, 2(2):89–124, 2011.

[16] Harsha De Silva, Dimuthu Ratnadiwakara, and Ayesha Zainudeen. Social influence in mobile phone adoption: Evidence from the bottom of the pyramid in emerging asia. *Information Technologies & International Development*, 7(3):pp–1, 2011.

[17] Andrea Dessi, Andrea Maxia, Maurizio Atzori, and Carlo Zaniolo. Supporting semantic web search and structured queries on mobile devices. In *Proceedings of the 3rd International Workshop on Semantic Search Over the Web*, page 5. ACM, 2013.

[18] Joseph S Dumas and Janice Redish. *A practical guide to usability testing*. Intellect Books, 1999.

[19] Mathieu dAquin, Fouad Zablith, and Enrico Motta. wayou–linked data-based social location tracking in a large, distributed organisation. In *The Semanic Web: Research and Applications*, pages 461–465. Springer, 2011.

[20] Timofey Ermilov, Norman Heino, Sebastian Tramp, and Sören Auer. Ontowiki mobile–knowledge management in your pocket. In *The Semantic Web: Research and Applications*, pages 185–199. Springer, 2011.

[21] Alvaro Graves. Creation of visualizations based on linked data. In *Proceedings of the 3rd International Conference on Web Intelligence, Mining and Semantics*, page 41. ACM, 2013.

[22] Martin J Halvey and Mark T Keane. An assessment of tag presentation techniques. In *Proceedings of the 16th international conference on World Wide Web*, pages 1313–1314. ACM, 2007.

[23] Tuukka Hastrup, Richard Cyganiak, and Uldis Bojars. Browsing linked data with fenfire. 2008.

[24] Apple Inc. ios human interface guidelines. `https://developer.apple.com/library/iOS/documentation/UserExperience/Conceptual/MobileHIG/index.html`. Accessed: 27 August 2014.

[25] Google Inc. Android design guideline. `https://developer.android.com/design/index.html`. Accessed: 27 August 2014.

[26] Daniel A Keim. Information visualization and visual data mining. *Visualization and Computer Graphics, IEEE Transactions on*, 8(1):1–8, 2002.

[27] Alexander de Leon, Filip Wisniewki, Boris Villazón-Terrazas, and Oscar Corcho. Map4rdf-faceted browser for geospatial datasets. 2012.

[28] Clayton Lewis. *Using the" thinking-aloud" method in cognitive interface design.* IBM TJ Watson Research Center, 1982.

[29] James R Lewis. Ibm computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1):57–78, 1995.

[30] Riccardo Mazza. Introduction to information visualisation. *Faculty of Communication Sciences, University of Lugano*, 2004.

[31] Fatih Nayebi, Jean-Marc Desharnais, and Alain Abran. The state of the art of mobile application usability evaluation. In *CCECE*, pages 1–4, 2012.

[32] Jakob Nielsen. *Usability engineering.* Elsevier, 1994.

[33] University of Mannheim. Adoption of linked data best practices in different topical domains. `http://data.dws.informatik.uni-mannheim.de/lodcloud/2014/ISWC-RDB/`. Accessed: 27 August 2014.

[34] Gonzalo Parra, Joris Klerkx, and Erik Duval. More!: Mobile interaction with linked data. *structure*, 6:7, 2010.

[35] Gary Perlman. Practical usability evaluation. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, pages 168–169. ACM, 1997.

[36] Emmanuel Pietriga, Christian Bizer, David Karger, and Ryan Lee. Fresnel: A browser-independent presentation vocabulary for rdf. In *The semantic web-ISWC 2006*, pages 158–171. Springer, 2006.

[37] Helen Sharp. *Interaction design.* John Wiley & Sons, 2003.

[38] M Skjæveland. Sgvizler: A javascript wrapper for easy visualization of sparql result sets. In *Extended Semantic Web Conference*, 2012.

[39] Aaron Smith. Smartphone ownership–2013 update. *Pew Research Center: Washington DC*, 2013.

[40] Claus Stadler, Jens Lehmann, Konrad Höffner, and Sören Auer. Linkedgeodata: A core for a web of spatial open data. *Semantic Web*, 3(4):333–354, 2012.

[41] IDC Corporate USA. Smartphone os market share, q1 2014. `http://www.idc.com/prodserv/smartphone-os-market-share.jsp`. Accessed: 27 August 2014.

[42] IDC Corporate USA. Worldwide tablet shipments miss targets as first quarter experiences single-digit growth, according to idc4. `http://www.idc.com/getdoc.jsp?containerId=prUS24833314`. Accessed: 27 August 2014.

[43] W3C. Sparql 1.1 overview. `http://www.w3.org/TR/sparql11-overview/`. Accessed: 27 August 2014.

[44] Max L Wilson, Alistair Russell, Daniel A Smith, Alisdair Owens, et al. mspace mobile: A mobile application for the semantic web. 2005.