Evaluating Web Mapping Applications for Visualising Demographic Diversity

A GEOG593 thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Social Sciences

at

The University of Waikato

by

Tristan McHardie

2016

Abstract

The use of maps on the World Wide Web is revolutionising how people access information. These web maps harness text, audio, video, animation, graphics and, most importantly, user interaction to improve the communication and interpretation of geographic phenomena. In this interactive and dynamic environment, web application developers have the opportunity to allow map readers to also be map producers; users can be enabled to query and analyse data and then assemble maps and other visualisations themselves, enhancing their understanding of the data and its geospatial relationships. This thesis focuses on what factors are important for the design of web mapping applications that provide visualisations of demographic diversity. Demographic information is inherently geospatial, making maps the ideal tool for visualising population statistics and their spatial relationships. An outcome of this research is a set of guidelines for reviewing web mapping applications that portray demographic diversity information. This thesis also provides an insight into the current standard for web mapping of demographic information and identifies the potential for improvement. A review of the literature, as well as a reflexive implementation, are the principal methods used to develop these guidelines. The essential components include the needs of demographers, and principles of cartography and human-computer interaction, as well as a consideration of open source and proprietary software. Five prominent, national demographic web mapping applications (two Australian, and one from New Zealand, the United States of America, and Canada) are used to develop these guidelines, as well as providing an insight into the current standard. The main conclusion of this research is that the guidelines developed provide a worthwhile structure but need to be kept broad. Also, there needs to be more emphasis on user data upload facilities, which is important for demographers, and cartographic communication needs to be given higher priority.

Acknowledgements

First and foremost, I would like to extend my immense thanks to my supervisors, Dr Lars Brabyn and Dr Natalie Jackson, for their enlightenment, expertise, wisdom, advice and support. While Lars and Natalie actively allowed me to make this thesis my own work, it has benefited enormously from the numerous conversations we have had.

Secondly, thank you to Capturing the Diversity Dividend of Aotearoa/New Zealand (CaDDANZ) for the generosity that made Masters study a financial possibility for me.

Thirdly, thank you to Mum and Dad for showing their unlimited and unquestioning confidence in me, and especially Mum for her many hours of proofreading and editing.

Finally, thank you to my dear Emma for her constant love and support, for encouraging me to go back to university, for picking up the pieces when things fell apart, and for keeping me focussed on the end goal.

Table of contents

Table of figures

Table of tables

1 Introduction

1.1 Background and objective

The use of maps on the World Wide Web is revolutionising how people access information. These web maps harness text, audio, video, animation, graphics and, most importantly, user interaction to improve the communication and interpretation of geographic phenomena. In this interactive and dynamic environment, web application developers have the opportunity to allow map readers to also be map producers; users can be enabled to query and analyse data and then assemble maps and other visualisations themselves, enhancing their understanding of the data and its geospatial relationships (Buchroithner & Gartner, 2013; Cartwright & Peterson, 2007).

This thesis aims to develop a set of guidelines for evaluating web mapping applications and to answer the research question: *what factors are important for the design of web mapping applications that provide visualisations of demographic diversity?* It approaches this with the end user being a demographer wishing to perform custom analyses and visualise the results using maps. However, the creation of web maps for population data spans multiple disciplines, and this thesis attempts to answer the research question by reviewing demography, cartography, human-computer interaction and the software technology behind web mapping.

The relationships between demography, cartography, human-computer interaction and software technology are demonstrated in Figure 1. Demographers are the subject user group, who also provide the information to be mapped and any interpretations of this information. Cartography defines the correct way to produce visualisations of the information using maps. This forms the link between demography and cartography, in the form of a workflow from data selection to data visualisation. Human-computer interaction defines the interface by which computers, in particular, web technology, enable the users to explore, select and analyse demographic data and create dynamic and sophisticated cartographic visualisations of that data. Human-computer interaction is underpinned by the software technology powering web applications.

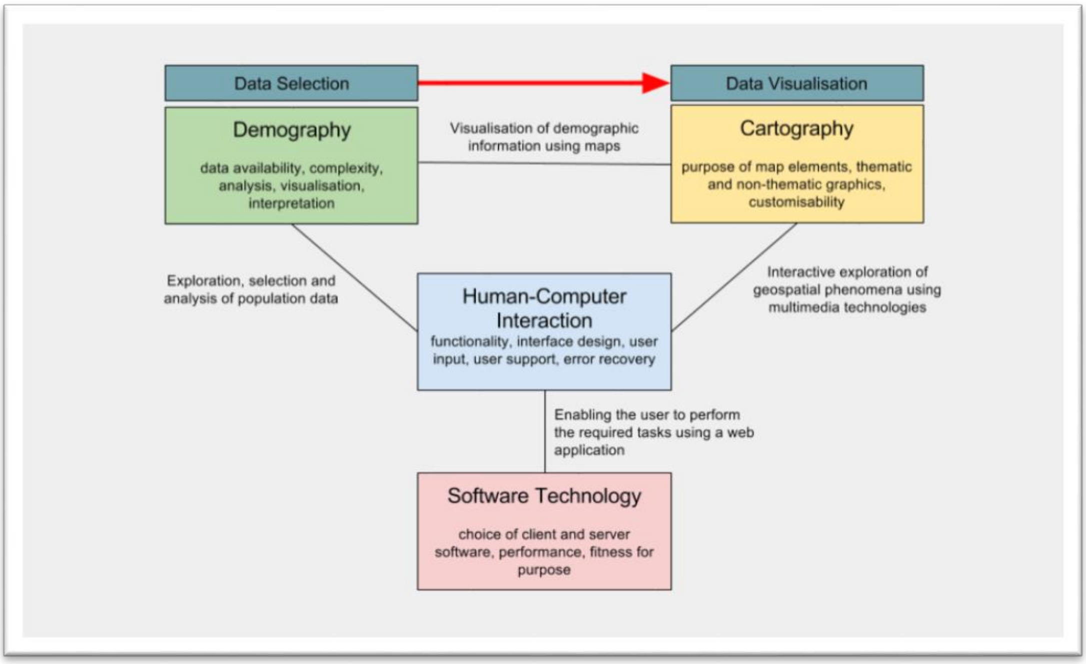

Figure 1: Relationships between disciplines of demography, cartography and human-computer interaction, underpinned by software technology.

The approach taken to all four disciplines is from a high-level point of view while acknowledging that there are many nuances and facets that it has not been possible to discuss in detail. The reviews attempt to extract and distill the most fundamental concepts of these separate subjects, which are broad and complex academic disciplines in their own rights.

Investigating demography serves to explain why we would want to look at population diversity in the first place and what we should be looking for. The literature review discusses what is important to demographers, based on research by Van Dalen and Henkens (2012), in particular, population ageing, large-scale migration and urbanisation which contribute to changing diversities in our populations. As demographers are the subject user group, a typical workflow for these users is also discussed, and how this might theoretically translate into a web application.

Cartography covers how maps communicate information and how the traditional principles for static print maps are extended for interactive and dynamic web maps. The implications of user interaction and animation are briefly explored, along with essential components for web maps, the differences between cartographic convention and cartographic grammar, thematic map types and the use of colour and symbology, with the aim of establishing a critical awareness of how these factors affect the way maps communicate spatial information.

Human-computer interaction is concerned with what makes a web application’s interface usable and useful. Usability studies are already a well-studied topic and, once the foundations have been established, this thesis looks to this research in order to create a general set of guidelines for evaluating the usability of a web application. Additionally, the human-computer interaction literature highlights methods for the evaluation of software usability, and this is adapted into the method which this thesis later uses for evaluating web mapping applications.

Software technology is discussed in terms of the differences between open source and proprietary software as this is a key issue for organisations wishing to adopt web mapping.

The guidelines developed from the literature are then applied in the evaluation of five publicly accessible web applications which provide data and maps of national statistics. The results of these analyses are used to conclude what issues are currently facing these web applications in terms of communicating the spatial relationships of demographic information, and how these issues might be resolved.

1.2 Relevance

Existing research has focussed more on the usability and functionality of web applications and web maps (see, for example, Komarkova, Jedlicka, & Hub, 2010; Komarkova, Visek, & Novak, 2007) and have developed very specific evaluation criteria which overlook the importance of the map graphics themselves in terms of cartographic communication. Compared to these other studies, this thesis develops sets of evaluation criteria which take demography and cartography into account in an attempt to highlight a need for a more balanced approach to web map evaluation.

These previous studies also ask very specific questions, often regarding the inclusion or exclusion of specific functionality. However, web applications are hugely variant and there may be perfectly valid reasons why specific items might be limited or missing. Therefore, this thesis develops deliberately broad criteria that are more aptly named *guidelines* (sometimes called heuristics). These guidelines allow evaluators more flexibility and subjectivity. More specific criteria are often irrelevant if the web application does not make use of the subject of a particular criterion, and this can unfairly affect the final scores. This thesis will look at what appears to be common amongst web mapping applications, particularly around available map types, and capture these commonalities into the guidelines for evaluation. To achieve this, this thesis takes a fresh look at the fundamentals of the four topics of demography, cartography, human-computer interaction and software technology for web mapping and uses this to make an informed decision on producing its own set of guidelines.

1.3 Structure of thesis

This thesis commences with a literature review chapter containing sections for each of the four subjects of demography, cartography, human-computer interaction and open source and proprietary software. As discussed previously, the literature review takes a high-level look at these subjects and attempts to extract the fundamental principles in order for these to be developed later into evaluation guidelines.

Following the literature review is a chapter that is written as a standalone paper to be submitted for peer review. The paper is the core of the research, taking the fundamental principles established by the literature review and developing them into the evaluation guidelines. It is also here that the method is developed and the five web applications are evaluated. The paper includes a summary of the evaluation results and concludes with a discussion of these results and how different map types should be used to better communicate geospatial relationships. It must be noted that the paper contains some duplicated content from the surrounding chapters but in summary form for publication.

After the paper, the thesis itself concludes with a full discussion of the results. Again, these are structured under headings for each of the four subjects and discuss what was discovered during the evaluations. The conclusion chapter also includes a discussion of the limitations of this research and how this research has contributed to filling the knowledge gap, followed by thoughts on future research and suggestions for a prototype web application.

2 Literature review

2.1 Demography

Demography is a science of populations (Lundquist, Anderton, & Yaukey, 2015; Weinstein & Pillai, 2016) and is explicitly linked to very basic aspects of humanity – when and where we are born, where we are from, where we live, when and where we die. The patterns which emerge when this data is analysed across a whole population provides answers to questions about what has happened, what is happening, and, perhaps most importantly, what might happen to the shape of our groups, communities, regions and nations. Demography is interested in the size and composition of populations, the change in the composition of populations, and the relationships between this change and composition. Demography is a discipline which takes the data and demands to know "so what?" (Lundquist et al., 2015). The answer to this question can inform decisions from government policy through to advertising and marketing.

2.1.1 Demographers

The breadth of application of the discipline is reflected in the disciplines it overlaps, with sociology, economics, anthropology, geography, history and biology (Lundquist et al., 2015) being the subjects with the most commonalities. This variation is reflected within demographers themselves, being "a 'melting pot' of different sciences and corresponding approaches" (Van Dalen & Henkens, 2012, p. 365). In a survey done in 2009 across 970 demographers worldwide, this variation within the profession was made apparent. For example, 53 percent of respondents graduated in demography, with others being predominantly from sociology, geography and mathematics and statistics (Van Dalen & Henkens, 2012). Of the respondents, Van Dalen and Henkens found that 46 percent described their practice as applied, 19 percent were fundamental, and 36 percent identified in the middle of the two. This uneven split between academic and applied is important to note. As this research could be used to create a tool for demographers, it is essential to discover what type of practice demographers engage in, and what the profession considers important. Van Dalen and Henkens' survey did show that demographers enjoy analysing data (59 percent) describing it as "highly rewarding", which indicates there is a market for a tool to aid them in this.

Commonalities in the profession exist at a fundamental level – "it is the norm among demographers is to stay close to the data" (Van Dalen & Henkens, 2012, p. 398). The survey also found that successful demographers must be highly empirical. When asked specific questions relating to their profession, the most important population issue for Asia/Oceania in the next 20 years was identified as population ageing followed by urbanisation and large-scale migration flows.

2.1.2 Population ageing, migration and diversity

With population ageing being at the top of the list of what is on a demographer's mind, it deserves some explanation here. Holdsworth (2013), Jackson (2007), Lundquist et al. (2015) and Weinstein and Pillai (2016) describe population ageing as the phenomenon where a population's median age increases over time, leading to a point where the population as a whole begins to decline. The process by which this occurs is known as the *demographic transition*; a four-stage process where the population moves from balance to a decline in mortality (improved life expectancy in infancy, childhood and late adulthood), to a decline in fertility (fewer individuals of child-bearing age), to a state of overall decline. During the shift from the decline in mortality to the decline in fertility, the population grows overall (see Figure 2). Jackson (2007) describes the four "dimensions" of population ageing as:

1. **Numerical:** absolute increases in the number of elderly
2. **Structural:** proportional increases in the number of elderly
3. **Natural decline:** fewer births than deaths
4. **Absolute decline:** inward migration cannot keep up with natural decline

Knowledge of the nature of population ageing and the demographic transition are critical in planning, for example, labour markets.

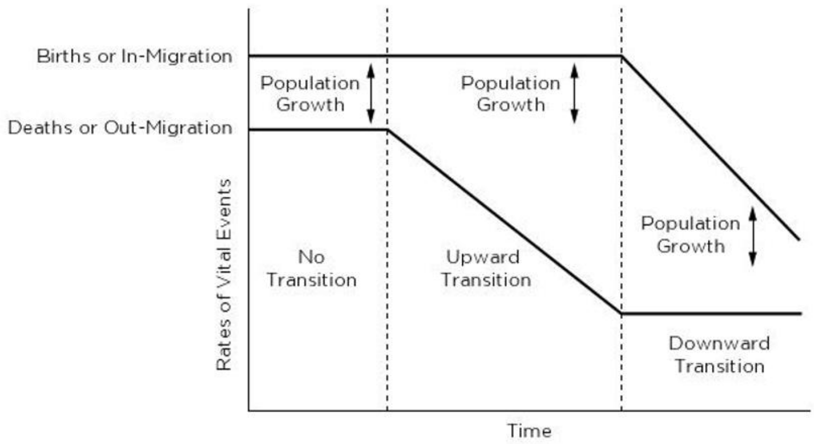

Figure 2: The Demographic Transition *(Weinstein & Pillai, 2016, p. 190)*.

Once geographically selective, with migrants predominantly moving to colonial settlements such Australia, South Africa, New Zealand, Canada, and the United States of America, migration has now shifted to a worldwide process. The reasons for migration are changing, and the patterns are no longer simply one-way permanent migration, with circular and temporary patterns emerging (Bakens, Nijkamp, & Poot, 2015). The way in which migration occurs impacts both the sending and receiving nations, changing communities, not just culturally, but socio-economically and spatially (Nijkamp & Poot, 2015). For example, migrants seem to settle in urban environments (Bakens et al., 2015), contributing to the phenomenon of urbanisation, one of the priorities identified by Van Dalen and Henkens' survey. A major outcome of these changes is in the way diversity shifts within a population.

Whilst migration is a major contributing factor to changing the diversity of a population, it is not the only way by which a population's diversity changes. Population diversity relates to differences occurring between various characteristics of a population, such as gender, age, ethnicity, culture, religion, language and socio-economics. It also includes how these occur both spatially and temporally, such as the changing structure of a population as it ages or the arrival of different ethnicities due to migration over time. Because of the nature of populations, and especially with increasing urbanisation, population diversity should be considered at multiple scales such as national, regional, local and neighbourhood, to best examine spatio-temporal heterogeneity (Nijkamp &

Poot, 2015). For example, when examining cultural diversity, measurement across multiple scales is important to show where the intermingling of cultures occurs rather than more homogeneous neighbourhoods in an otherwise culturally diverse region. Diversity is usually enumerated as a fractionalisation index; essentially, the probability of an individual from one cohort encountering another individual from a different cohort, and these indices are often mathematically similar to the indices used by biologists and ecologists (Nijkamp & Poot, 2015). While it is not within the scope of this thesis to discuss the components or mathematics of the many indices available (see, for example, Nijkamp & Poot, 2015), it is sufficient to say that diversity indices are commonly calculated as a value between zero and one, indicating a probability as mentioned above.

2.1.3 Typical analysis workflow

The typical analysis workflow of a demographer falls into two paths to obtain data for visualisation. This depends largely on the type or depth of analysis they are performing and whether they require the use of *raw* or *pre-processed* data. The following sections discuss a typical workflow for both of these paths from a demographer's point of view and develop this into how a web application might work, based on discussion with N. Jackson (personal communication, June 13, 2016).

For the first path, demographers are often required to answer very specific and unique research questions which require highly bespoke analysis and a demographer undertaking some deep analysis will require raw data, often from multiple sources and in different formats, and might spend considerable time cleaning, standardising and preparing this data before carrying out any analysis. Typically, such complex numerical analyses would be undertaken using software such as Microsoft Excel. However, in the second path, some demographers may have a use for pre-processed data, such as fractionalisation indices if these are available to them.

Raw population data typically refers to absolute counts for a particular variable, for example, identified ethnicity. This is opposed to pre-processed data, which might include pre-calculated ratios, rates, proportions and indices. Population data is most often aggregated by some areal unit (e.g. census unit, city, region or country) to protect the privacy of individuals within the counting area. Smaller counts might also be rounded or randomised (see, for example, Statistics New Zealand, 2015) to add further privacy protection. In addition, the units of the variable itself may be aggregated; for example, ages grouped in five-year ranges, or broadly categorised ethnicities. While the exact locations of individuals in a dataset naturally provides the best spatial resolution, it has to be accepted that population data, especially from censuses, will be aggregated and analysts must work within this constraint.

2.1.4 Web application workflow

So, what does a demographer need from a web application in order to visualise population diversity using maps? This thesis breaks this question into a workflow of three phases:

Phase 1. Data procurement
Phase 2. Data analysis
Phase 3. Visualisation and interpretation

Ideally, a single web application could be used to deliver this entire workflow. However, the data analysis phase would usually be performed using software like Microsoft Excel as clarified by N. Jackson, (personal communication, June 13, 2016). A key issue here is that a web application attempting this phase would need to emulate the complex functionality of this software, and this is by no means a trivial undertaking. Therefore, it is suggested that the notion of performing the data analysis phase within the web application is abandoned, and left to more appropriate software like Microsoft Excel. The single-application workflow described above now becomes fragmented into three standalone tasks; phase two is outside the scope of the web application, and phases one and three are standalone functions that can be invoked optionally and independently. This is especially true when considering that the data may be acquired partly from sources outside of this hypothetical web application. For now, we will consider just the requirements for the data procurement and visualisation and interpretation phases, and treat these as independent of the overall workflow. The paper component of this thesis develops these phases in the context of the hypothetical web application.

2.2 Cartography

Since the Internet became a medium for map distribution in the mid-1990s, the audience for interactive maps has been increasing. In fact, in 2001, the rate at which web map use was increasing was stronger than the growth of the Internet use overall. With continuous improvements in mobile technology, map use is also prospering in this area (M. P. Peterson, 2007b). M. P. Peterson also noted that it is interesting how readily people have adapted to web maps. This new, interactive, *multimedia cartography* has grown from a user requirement for intuitively presented geographic information by combining extra elements such as text, audio, video and animation to "enhance [the] user experience and make it easier and faster to grasp information" (Cartwright & Peterson, 2007, p. 1). Ormeling (2007) succinctly describes the use of interactive cartography as allowing the user to not just ask 'what is there' but also 'what more is there?'

Geographic information "has become increasingly 'scientific' and 'objective' in nature" (Taylor & Lauriault, 2007, p. 519) but the focus still needs to remain on qualitative properties such as *artistic* and *emotional* to "more effectively portray and represent a world rich in colours, texture, sounds and smells" (Taylor & Lauriault, 2007, p. 519). Taylor and Lauriault note that the way digital cartography is currently displayed "is emotionally stark, largely visual and often in black and white." Whilst information displayed via a computer does not replace the real world, they say that multimedia cartography can attempt to create a more "interesting, engaging and effective" experience.

According to MacEachren, Crawford, Akella, and Lengerich (2008), more emphasis has been placed on developing the tools necessary for web mapping rather than the design principles. Peterson states that "cartography today is as much about [computer programming] as it is about design and data analysis skills" (2015, p. 240). However, Peterson also states that cartographers will still have much to tweak and tune until software improves to the point where most facets will be automatable.

Kraak and Ormeling (2011) note that a geographic information system (GIS) is generally not a drawing tool. Printed maps would normally be created in a GIS and then exported to specialised graphics software for further cartographic refinement. Web maps, on the other hand, are generated directly from GIS software and cannot take advantage of specialised graphics software. This limitation has led to two different paths for the creation of print and web maps.

Cartography aims to reduce reader error and misinterpretation by using suitable graphic presentation. Kraak and Ormeling state that *cartographic grammar* is well-defined and that the basics do not change for the web, saying "well-designed web maps can be recognised as relatively 'empty'" (2011, p. 79). Beyond these basics, Peterson asserts that digital map design is still in its youth, and that "common rules or standards for cartographic presentations on screen displays are not defined yet" (2015, p. 55) and there is minimal published literature on the subject, leaving much to trial and error. M. P. Peterson (as cited in Cartwright & Peterson, 2007) asserts that "paper thinking" still over-influences multimedia cartography.

According to Miller (2007), a web map interface can be broken down into:

- **Informational objects**: the map object itself, direct spatial elements and their symbologies
- **Functional objects:** marginalia such as scale and direction indicators, titles, legends and controls for querying and manipulating the thematic data

- **Aesthetic objects:** for example, borders and logos.

2.2.1 Essential map components

2.2.1.1 Basemaps

Maps can be broadly categorised as either *general reference* or *thematic* (Robinson, 1995). General reference maps describe the location and character of physical objects, and thematic maps describe the distribution of measurable phenomena, such as human populations. Thematic maps still require a degree of general reference detail, in the form of a *basemap*, so the reader is able to spatially orient and appreciate the location and distribution of the phenomena in relation to the physical world. The level of basemap detail required for a thematic map might be as minimal as political boundaries but, depending on the needs of the theme and user, it might also include roads, place names, coastlines buildings and bodies of water or other physical objects (Cammack, 2007; Kraak & Ormeling, 2011). An example cited by Kraak and Ormeling (2011) is a map of relative fish catches by country; a landlocked country displayed a relatively high catch when viewed in relation to other landlocked countries. If the topographic detail did not show the presence of a large lake, the reader might interpret the high catch as an anomaly in the data.

2.2.1.2 Map elements

According to cartographic texts such as Kraak and Ormeling (2011), G. N. Peterson (2015) and Robinson (1995), the first essential consideration of map design is that the map must communicate a clear purpose. In the case where mapping, for example, demographic diversity, the demographers are providing meaning to numeric data and this needs to be clearly portrayed by the map's theme and, in particular, by the map's title. The map's title should include the area, the common denominator of the theme and the time. The purpose of the map needs to be clear to the target audience and they must be kept in mind during the design process.

Other cartographic elements which are considered essential to paper maps by Kraak and Ormeling (2011), G. N. Peterson (2015) and Robinson (1995) are north arrows, insets, scale bars, graticules and legends. However, this thesis suggests that, because web maps are a different environment to print maps, the cartographer should feel free to make a judgement call as to whether these elements are truly required or if they could negatively impact the overarching aim of effective communication. For example, a legend might not be required if the web map uses interactive pop-ups to describe a symbol. Miller (2007) states that a coverage map, or map inset, should be used due to the ability for web maps to be panned and zoomed at the user's whim. Again, this thesis argues that this is not necessarily true, depending on the context of the map and the availability of other orienting properties, such as place labels and other marginalia. Symbols, lines and text should be visually distinct and readable. Good maps will establish a clear visual hierarchy by contrasting the theme from the basemap. The map should only convey one or two phenomena as trying to include too many variables can be detrimental to communication. To overcome the issue of poor quality data, it should be mapped in less detail and at smaller scales to prevent an inaccurate interpretation of quality. Additionally, data sources should be appropriately attributed.

2.2.1.3 Scaling and generalisation

Miller (2007), Kraak and Ormeling (2011), Peterson (2015) and Gaffuri (2011) all discuss a very important and unique factor of web map design: scaling. A user can zoom in and out of a map at will and expect immediate results and this is not a trivial matter like enlarging or reducing a photograph. At smaller scales, irrelevant objects should be removed, e.g. minor roads. At larger scales, some objects may require more detail, e.g. rivers might need to be represented as polygons rather than simple lines. To achieve this range of requirements, data is commonly served for 20 scale sets (sometimes grouped) where the cartographic features are generalised as the scales get smaller (G. N. Peterson, 2015). Generalisation is the result of elimination, simplification, displacement and enlargement of cartographic features to improve readability. After generalisation, the symbols themselves may also be scaled to aid in legibility and prevent overcrowding such as line thickness, font size and point size. The aim is to provide a "consistent look and feel across all zoom levels" (G. N. Peterson, 2015, p. 247).

Research by Cecconi and Galanda (2002), while dated, highlights that generalisation and scaling are about compromise. When scale sets are generated beforehand, performance is fast and cartographers can easily tune the results, but at the expense of updateability. Conversely, on-the-fly generalisation will seamlessly handle data updates at the expense of speed and intervention from cartographers. This research is backed by Gaffuri (2011) who assert that pre-processed generalisation is not interactive enough. Cecconi and Galanda proposed a combination of on-the-fly adaption and pre-built scale sets.

2.2.1.4 Projections

The standard projection used by web mapping software is Web Mercator and this is often the only option available. While Web Mercator preserves direction and shape, and results in a squareness which is useful for tiling purposes, it can produce noticeable distortion at a continental level and nearer the poles (G. N. Peterson, 2015). This effect may be noticed when comparing the shape of areas on small-scale overviews to the shape of those same areas when viewed at larger scales.

2.2.1.5 Multiple maps

For the purposes of comparison, it can be useful to display two or more maps simultaneously, each map showing a different location or moment in time. MacEachren et al. (2008) created a system for analysing cancer statistics aggregated by area where they used series of multiple linked *micromaps*. Not only can micromaps such as these show a single phenomenon across different areas or times but they can also be swapped to show a single area with multiple phenomena or times. In cases such as these and, indeed, any situation where two or more maps are being compared, Ormeling (2007) asserts that the theme or the area (and scale) must remain constant when the opposing element changes.

2.2.2 Cartographic communication

Like print maps, the primary function of (most) web maps will be the *communication* of spatially related information. In the case of communicating demographic information – among others – this means using some manner of *thematic map* to portray the spatial occurrence of a phenomenon.

Maps communicate using the principles of *cartographic grammar* and *cartographic convention*. General reference maps, such as basemaps and topographical maps use cartographic convention, for example, the forest is green, the water is blue and north is up. By comparison, when designing thematic maps cartographers look to the broader rules of cartographic grammar, as convention cannot capture the many different scenarios which thematic maps will potentially encounter (Kraak & Ormeling, 2011). Substantial research on the neurophysiological, psychological and cognitive backgrounds of cartographic grammar underpins cartographic literature and has informed the general understanding of how humans see and interpret maps. While this thesis will not delve into the specific ways in which human visual hardware turns seeing into perceiving (although human-computer interaction, a related topic, is discussed later in the thesis), it is important to note that this research exists and informs cartography at its most fundamental level. The interested reader should look to works such as MacEachren (1995) for an explanation of this topic, which is multi-disciplined in itself.

2.2.2.1 Map types

Map themes can be qualitative or quantitative. Before choosing a map type, spatial data should be assessed to determine its nature, length and range. Kraak and Ormeling (2011) categorise data into five types:

- **Geographical:** physical features, usually reserved for basemaps

- **Nominal:** qualitative, e.g. primary spoken language
- **Ordinal:** quantitative, ordered classes, e.g. cold, cool, warm, hot
- **Interval:** absolute quantitative intervals, e.g. the total population of an area
- **Ratio:** relative quantitative intervals, e.g. the proportion of an ethnicity to the total population of an area.

It is also useful to categorise quantitative data as either *spatially extensive* or *intensive*. Spatially intensive data is calculated using a spatial variable, for example, population density, where the denominator is the actual area of the aggregation unit. Conversely, spatially extensive data is not derived from a spatial variable and an example of this is absolute population counts (Buckley, 2013; Haining, 2003). The data's scale, and whether it is spatially intensive or extensive, should inform the choice of map type. Generally, the mapping of population statistics leads to the use of choropleth or proportional symbol maps (Kraak & Ormeling, 2011).

Choropleth maps use discrete, stepped colours to represent quantitative differences between areas. Typically, choropleths are used to visualise statistics that have been aggregated by counting areas, such as census units. This aggregation yields maps which tend to convey homogeneity in the distribution of the statistic, and it is important that choropleths be used with caution when mapping spatially extensive data. For example, if an area with an absolute population count of 10,000 were divided into two equal parts, it would not necessarily be true to say that both parts contain 5,000 people. Further, choropleth boundaries may be arbitrary, and may not necessarily represent the distribution of the phenomenon, or may be subject to the modifiable area unit problem[1] and ecological fallacies[2]. Another issue with choropleths is a tendency for readers to perceive the size of geographic areas as carrying ordinal value,

[1] Modifiable area unit problem: summarising data by different areal units produces different results for the same location.

[2] Ecological fallacy: drawing conclusions about individuals from aggregated data.

even if this is not relevant. Despite their drawbacks, choropleths are the most commonly used map for visualising socio-economic phenomena due to their ease of construction by computers, and ease of interpretation by readers. A distinction should be made between choropleths and **chorochromatic** maps; the latter uses colour to represent qualitative, rather than quantitative data. Colour is discussed in more detail below, but choropleths should use differences in lightness or saturation to represent different values, while chorochromatic maps should only use changes in hue (Buckley, 2013; Haining, 2003; Kraak & Ormeling, 2011; Meirelles, 2013).

Cartogram maps are a clever alternative to choropleths where the sizes of areas are changed relative to the occurrence of the phenomenon being mapped, in order to overcome the issue of false ordinal importance mentioned above (Kraak & Ormeling, 2011; Meirelles, 2013). However, this thesis argues that cartograms can distort the size and shape of areas so much that they should be treated more as a supplementary tool.

Another alternative approach to choropleths are **dasymetric maps** where additional topographic information is used to create areas which better represent the true distribution of the phenomenon being mapped, in order to reduce perceptions of homogeneity. However, because of the limited GIS software options for producing them, dasymetric maps are not likely to be used for dynamic web maps (Kraak & Ormeling, 2011; Meirelles, 2013).

Dot density maps use dots to convey the density of a phenomenon in a particular area. However, the use of dots can lead to a false perception of absolute location when used with aggregated data, and Meirelles claims users have a "tendency to underestimate the number of dots and the differences of densities between areas" (2013, p. 132). When the distribution of values varies too greatly, some areas will appear too dense (overlapping dots) and some too sparse. Dot size and colour can also be used to introduce further dimension and flexibility. Testing is required to determine the best proportion of dots to use, which limits use in automated symbolisation.

Proportional symbol maps use geometric shapes of varying size to convey difference in value. The use of simple geometric shapes as symbols, such as circles, squares and triangles, is important; the more complex the shape, the more difficult it is to perceive a difference in size. The range of values is important and the minimum and maximum values should be shown in the legend. Kraak and Ormeling (2011) warn that three-dimensional symbols can lead to serious misinterpretation because the viewer perceives the area covered by the symbol before the actual size of the symbol (Kraak & Ormeling, 2011; Meirelles, 2013).

Isopleth maps use *isolines* to partition intervals of a continuous phenomenon where there is a continuous distribution in all directions. While isopleths are not as intuitive as choropleths, they are appropriate for comparing different phenomena and are therefore more effective than choropleths because they can be used to demonstrate trends and the direction of change. Because of the continuous nature of their surfaces, isopleths can display more than the seven classes recommended for choropleths (see below) (Kraak & Ormeling, 2011; Meirelles, 2013).

Statistical surface maps are useful as dramatic three-dimensional visualisations of the peaks and troughs of a phenomenon. However, the rendered image may become overly complex and background areas can be obliterated by peaks in the foreground (Kraak & Ormeling, 2011). Statistical surfaces are more difficult to construct, especially dynamically, but there may be merit to further research in their use for population mapping, especially in a web mapping environment where the user can pan the three-dimensional model. Colour ramps may help the viewer decipher the complexity more readily, as is the case with other three-dimensional graphs. A downside to statistical surfaces is that it becomes more difficult to include a basemap, thus making it harder for users to spatially orient themselves.

Another option for showing movement of a phenomenon are **flow line maps** and these are useful for portraying migration trends. Lines of varying thickness represent value while arrowheads represent direction. The length of the lines can also be used to approximate the distance of movement in the phenomenon. Thick bands may obscure the underlying topography and indeed other bands (Kraak & Ormeling, 2011; Meirelles, 2013).

2.2.2.2 Colour and symbology

Once the correct map type is determined, the way the map is coloured and symbolised needs to be considered. Cartographic convention plays an important role in how maps are interpreted, particularly in topographic maps, as discussed above. However, convention is only part of the story and the "general character of graphics cues used" (Kraak & Ormeling, 2011, p. 62) also help improve interpretation. For example, points representing businesses use the business' logos, railways by a dual track with sleepers or hospitals by red crosses. Miller (2007) asserts that these "self-describing symbols" may also negate the need for legends. While conventions are helpful in topographic mapping, Kraak and Ormeling state that thematic maps do not necessarily have conventions, meaning cartographers must look to the broader guidelines of cartographic grammar when choosing map types and symbolisations for these.

In the context of multimedia cartography, Miller (2007) remind us that exact figures can always be made available by interacting with the map, e.g. clicking or hovering over an area or point of interest.

Bertin (as cited in Kraak & Ormeling, 2011; Robinson, 1995) listed six graphic variables which can be exploited by cartographers:

Qualitative	**Quantitative**
• Hue • Shape • Textural orientation	• Symbol size • Lightness • Texture

Table 1: Cartographic variables for representing qualitative and quantitative differences.

Kraak and Ormeling list three other proposed graphic variables in addition to Bertin's six: *saturation*, *arrangement* and *focus*. In the context of a web browser, Kraak and Ormeling also suggest using *opacity*, *shadow* and *shading*, although these new variables still require more research to determine their cartographic effectiveness. Before choosing which graphic variables to use, the cartographer must determine the hierarchy of information, which will translate into the visual hierarchy. This visual hierarchy will determine the cartographic methods used to demonstrate the importance of each element, e.g. the colour of the sea should not dominate a map explaining strictly land-based phenomena (Kraak & Ormeling, 2011; Miller, 2007; G. N. Peterson, 2015).

The use of colour has psychological, physiological, cognitive and associative (e.g. convention) issues (Kraak & Ormeling, 2011; Meirelles, 2013). Kraak and Ormeling state that the colours used to symbolise qualitative and quantitative differences need to be handled differently; normally, changes in lightness are perceived as differences in quantitative value, e.g. pink implies a lower density than red, while changes in hue are perceived as qualitative differences, provided they have a similar lightness. Kraak and Ormeling go on to say that colour is more difficult to perceive in smaller areas and that darker colours tend to dominate an image; therefore, highly saturated colours should be reserved for smaller areas. A system which automates symbolisation would need to take areal size into account so no false ordinal relationship is perceived.

There seems to be no hard-and-fast rule for the maximum number of colours which can be used on a single map. According to Kraak and Ormeling (2011), the maximum number of discernible colours at a glance is eight, although this can be increased using textures. However, G. N. Peterson (2015) claims this number is 10 – 12 before human vision becomes fatigued and that too much colour, including colour in the marginalia, is tiring. G. N. Peterson also suggests the maximum number of discernible shades (lightness) of the same hue is five.

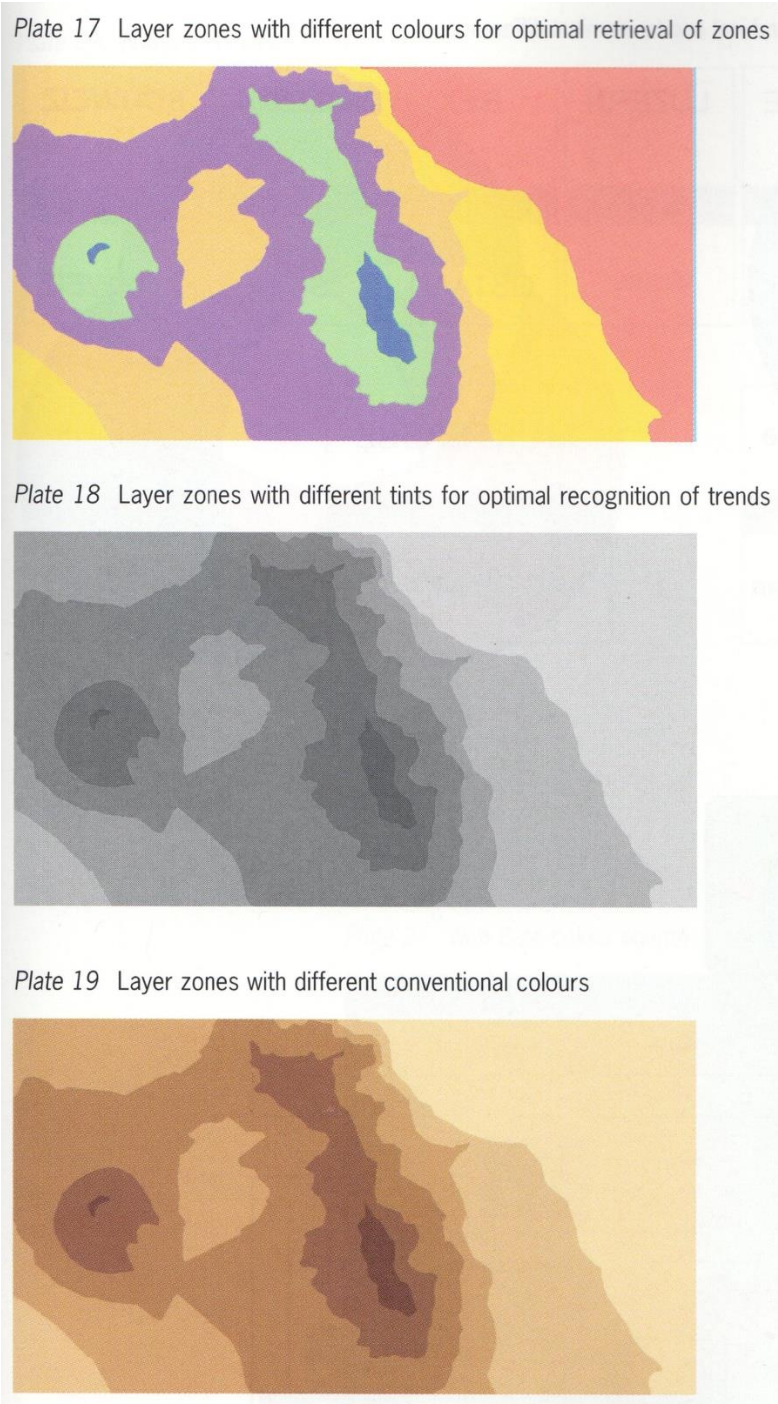

Figure 3: Examples of representative differences in colour (Kraak & Ormeling, 2011).

2.2.2.3 Spatio-temporal relationships

Static maps can only display a single snapshot in time, making it difficult to represent change. Change is more difficult to describe cartographically, for both the cartographer and the user. While change can be represented using a series of static maps, as would be required if the maps were printed, a major benefit of on-screen cartography is that map themes themselves can be *animated*. In fact, an animation can be thought of as a series of individual maps displayed frame-by-frame and offers a very expressive way to represent changes in location and value (Kraak, 2007; Kraak & Ormeling, 2011). Interpolation can be used to smooth the transitions between frames (M. P. Peterson, 1995), and, with modern cascading style sheets (CSS) technology, vector data can be animated directly in the browser, allowing the possibility of animating live streaming data. It is also worth noting that the variable of change does not necessarily have to be time; non-temporal variables, such as a change of projection, can also be animated.

Where animations are provided, the user should be able to control the display with controls for play, pause, fast-forward, rewind and even the possibility of moving through the timeline manually using a slider. In an exploratory environment, the user will likely have queried the data, making them the animator, so care should be taken to ensure there is adequate flexibility in control. (Kraak & Ormeling, 2011; MacEachren et al., 2008).

While the rules for static maps are well defined, animation opens up a new realm of cartographic possibility, the effects of which are not yet fully understood and have been researched (Kraak, 2007; Kraak & Ormeling, 2011). Kraak and Ormeling note that there are perceptual and cognitive limitations with animations where cognitive overload may mean the user is not able to fully process meaning from the data, due to the sheer number of frames. Kraak and Ormeling also suggest that the symbols used in animations should themselves represent change, e.g. arrows for movement or shapes which change in size.

2.2.3 Map interactivity

In an interactive environment, users become the map *producers*; they want to assemble the pieces themselves and they may want to spend considerable time customising and exploring (Buchroithner & Gartner, 2013; Cartwright & Peterson, 2007). Cartwright and Peterson claim this is due to how humans learn and structure knowledge because “interaction is the key to knowledge formation” (2007, p. 2). With an estimate of more than half of the educated population not being competent with maps, interactive multimedia cartography may help improve knowledge acquisition amongst these users (M. P. Peterson, 2007a). Web maps allow the user to query the data and perform analysis, enabling them to explore and understand data, determine geospatial relationships, synthesise the results of analysis and communicate new geospatial knowledge (Kraak & Ormeling, 2011).

In comparing the difference between static maps and today’s dynamic and interactive web maps, Miller (2007) talks of the “hypermedia paradigm” by which users of such maps are now conditioned to expect extra information to be available by interacting with the map. Interaction with web maps ostensibly falls under human-computer interaction but, as asserted by Cartwright and Peterson (2007), interaction plays a fundamental role in knowledge transfer. To be effective, any interactive map elements must balance their cartographic symbology and iconology with the need to depict that the element is, in fact, interactive; web maps should actively promote interaction by indicating which features are interactive, such as symbols or the cursor changing on hover. The map object should be manually navigable by the user with pan and zoom tools, possibly alongside links that zoom to predefined areas. In fact, Ormeling (2007) asserts that it is “unethical” to only provide fixed frames.

Web map interaction is discussed here under cartography, but it is recognised that this is also rooted in human-computer interaction (Kraak, 2007; Kraak & Ormeling, 2011; MacEachren et al., 2008). In this respect, cartography and human-computer interaction have a lot in common and *multimedia cartography* has much to learn from human-computer interaction research (Taylor & Lauriault, 2007).

2.3 Human-computer interaction

Human-computer interaction (HCI) is a medley of several separate disciplines such as psychology, cognitive science, ergonomics, sociology, computer science, engineering, business and graphic design, to name a few (Dix, Finlay, Abowd, & Beale, 2005). HCI "is both a craft and a science" and does not have a "general and unified theory" (Dix et al., 2005, p. 5). A fundamental principle of HCI is that computers and software should be designed in a user-centric way (Taylor & Lauriault, 2007) and with the understanding that users have specific tasks in mind and want to achieve those tasks in ways which are seamless with respect to their lives outside the computer (Dix et al., 2005). HCI is not just about pretty interfaces; it is integral to the design of systems as a whole and successful products have three *use* words: *useful*, *usable* and *used* (Dix et al., 2005).

The cartography of web maps and HCI have a lot in common. The crossover is mentioned by Taylor and Lauriault: "Seeing the user as a consumer is important but the future of multimedia cartography also depends on a much deeper understanding of the user in terms of both use and usability issues. Here, multimedia cartography has much to learn from human factors psychology, cognitive science, and from [HCI] research in general" (2007, p. 510).

2.3.1 Input and output

The science of HCI is broken down into *input-output channels* between humans and computers. A typical consumer computer has a monitor and speakers (output) and a keyboard and mouse (input). Conversely, humans have eyes and ears (input) and limbs (output). Although many other inputs and outputs are possible, this example demonstrates the input-output channels of monitor-to-eyes, speakers-to-ears and limbs-to-keyboard-and-mouse (Dix et al., 2005).

To web map creators, vision is the human input channel of primary importance. Human vision is limited to what can be physically seen but our brains are also very good at interpretation – we can use incomplete information to derive a complete image (Dix et al., 2005). Distance and size are dominating factors in how humans interpret images; how much of the field of vision an object occupies, along with other cues such as foreground objects overlapping background objects, determines how we perceive (Dix et al., 2005). In reading text, human eyes move in periods of saccades (rapid movements) and fixations. Recognition and interpretation occur during fixations, which occupies 94 percent of viewing time. Common words are recognised by shape, allowing a reader to scan. Deviating from the common shape or nature of words or symbols can affect the speed of interpretation (Dix et al., 2005; MacKenzie, 2013). When searching lists (text or otherwise), there is a linear relationship between the number of items to scan and the time taken to find the item (MacKenzie, 2013). Dix et al. cite how research has proven that backlit screen reading is slower so text interfaces require careful design. A negative contrast (black on white) improves luminance and therefore legibility.

2.3.2 Usability

The usability of a system is governed by how well the user can interact with it. In web mapping, pointing devices (mouses, trackpads, etc.) play the most important role, allowing the user to move the cursor, operate controls to query the data, pan and zoom map objects and interact with points or areas of interest.

Reaction to a stimulus and the time taken to move in response is largely affected by factors such as age, fitness, fatigue and skill (Dix et al., 2005). MacKenzie (2013) backs this up but adds that interface design is a speed-accuracy tradeoff and that performance improves noticeably with practice. Fitts' Law (cited in Dix et al., 2005) is a formula derived from an empirical study of human accuracy when moving or reaching for targets at increasing distances, such as moving a mouse. While Fitts' research was in ergonomics, the formula also holds true in HCI. The formula is:

$$D = a + b * \log_2(\frac{A}{W} + 1)$$

Where D is an index of difficulty (usually time), a and b are constants related to the particular task, A is amplitude (usually the distance to the target) and W is the reachable width of the target. Fitts' Law is an important concept for analysing user interface design as it shows that relatively smaller targets at relatively larger distances are increasingly difficult to reach, meaning commonly used controls should be grouped near each other or, if they are distant, targets should be of a sufficient size to hit them easily. An example in a web map context might be zoom buttons that are placed in the bottom right corner of the display when most other controls are located near the top left corner. The user will spend the majority of their time near a group of primary controls, so when they do move to the remote zoom buttons, these buttons ought to be large enough so that the index of difficulty in reaching them is appropriate. Alternatively, the zoom buttons can be made smaller but therefore will need to be closer to the primary centre of the user's attention.

2.3.3 Design principles

Dix et al. (2005) discuss design principles and guidelines for applying HCI in practice without needing to understand the various disciplines which underpin them. The principles often overlap and may enhance or interfere with each other, meaning they need to be applied with careful thought. A distilled list of HCI principles from Dix et al. is below.

Learnability:

- **Predictability:** No surprises. Past interaction guides future interaction as recognition is more powerful than recall. All operations a user can (or cannot) perform should be visible.
- **Synthesisability:** Users should be able to determine that change has occurred as a result of an operation. This helps build a mental model of how the system operates.
- **Familiarity:** The user should be able to apply knowledge of other systems (computer or not) in the operation of the current system.

- **Consistency:** Similar to convention in cartography. Red-coloured feedback means ‘bad’, green-coloured feedback means ‘good’. Arrow keys perform directional operations.

Flexibility:

- **Dialogue initiative:** Interaction can be initiated by the system or by the user. User pre-emptive dialogue, where the user initiates action, is preferred over system pre-emptive, where system initiates action. This gives the user a feeling that they are in control and have the ability to perform actions as they need.
- **Multi-threading:** Allow multiple tasks to occur at the same time. Windowing systems are the most common example of this in practice. Modern web browsers allow this to occur seamlessly in a web context without having to build this into the product.
- **Task migratability:** Task execution can be transferred from the user to the system and vice versa. Mundane actions, like spell checking, can be performed by the system, but it would be unwise to allow the system to always autocorrect.
- **Substitutivity:** Allowing flexibility in the way a user interacts, e.g. when entering values either “2.5 cm” or “1 inch” are acceptable.
- **Customisability:** User-initiated customisability is *adaptability,* and system-initiated customisability is called *adaptivity*. Users should be able to customise the way they use the system, such as moving toolbars, while the system might automatically adjust itself if it detects an advanced user over a novice. Adaptivity is likely to interfere with predictability if it is not done well.

Robustness:

- **Observability:** The user can determine the internal state of the system by its output.

    - **Browsability:** The state of the system can be explored via the interface.
    - **Defaults:** Suggested input can quickly be perceived as correct or not and reduces physical action. May evolve adaptively.
    - **Reachability:** Ability to navigate through observable states.
- **Persistence:** System outputs should represent system state persistently. For example, a user may hear a 'beep' indicating a new email has arrived but will soon forget. A persistent visual notification will serve as an additional reminder.
- **Recoverability:** Humans make mistakes, and a system should allow recovery from these. Undo, redo, back, forward commands are examples of this.
- **Responsiveness:** Users, particularly on the web, are impatient, and the performance of the system is key to usability. Response time should be fast or at least indicate that the system has not stalled.
- **Task conformance:** Does the system perform all the required tasks? Does it fulfil the user's requirements and expectations based on their understanding of the tasks?

2.3.4 Evaluating usability

Beyond the HCI fundamentals above, there are already established approaches to how the usability of software is evaluated. The terminology in this area varies, depending on the author and their approach. For example, Komarkova et al. (2007) use the term *heuristics* for specific criteria which enable unskilled evaluators to make quick decisions. On the other hand, Nielsen (1995) refers to heuristics as broad rules of thumb which can be used by evaluators who have more background knowledge of the area of evaluation, and this allows more subjectivity in decision making. Travis (2016) uses the term *guidelines* in the same way Komarkova et al. talk of heuristics. In the interests of consistency and to use a plain English term, this thesis uses the term **guideline** to refer to the criteria for evaluating usability.

Komarkova et al. (2007) have developed 138 guidelines for evaluating the usability of web-based GIS. However, this list focusses in detail on the usability of the GIS itself rather than the cartographic communication of the output. Travis (2016) develops a more general approach with a comprehensive 247-point list of guidelines for evaluating usability, task orientation, navigation, data entry, content quality, layout and design, user support and fault tolerance, specifically for web applications. While many of Travis' guidelines are related to the marketing aspects of website design, the usability aspects can be adapted for the purposes of performing web application evaluations and these aspects have been included in the HCI factors in Table 4 on page 60 of this thesis.

Also adapted into the HCI factors in Table 4 are the more fundamental and broad usability guidelines from Nielsen and Mack (1994). These guidelines look at the usability of software in very broad terms and reinforce Dix's et al. (2005) design principles from above. These guidelines are distilled on Nielsen's (1995) website:

- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognise, diagnose, and recover from errors
- Help and documentation

2.4 Open source and proprietary software

The free and open source movement has become a cultural revolution, and this is not limited to software; we now have open source cars ('OSVehicle', n.d.), electronics ('Arduino', n.d.), and even firearms (Defense Distributed, n.d.). The decision between significant capital expenditure in more traditional commercial software, versus using a competing and ostensibly free alternative is not obvious, and organisations initially approached this choice in a “reluctant and uncomfortable manner” (Fitzgerald, Kesan, & Russo, 2011, p. vii).

The terminology used to describe software distribution models can be a point of confusion, and definitions vary between sources, which may be a symptom of how rapidly software technology is changing. For example, the opposite of open source is not usually called closed source, and free software might not mean free of charge. This is not an exhaustive list, but it does serve to generalise the most common paradigms of software licencing models.

Open source software means the product’s source code has been made publicly available. Open source software is usually non-commercial — although there are exceptions to this — and it is not necessarily free software (see definition below) (Fitzgerald et al., 2011; ‘gnu.org’, n.d.; Open Source Initiative, n.d.).

Proprietary software is the opposite of open source and is sometimes referred to as closed source (not to be confused with commercial software). Proprietary software may be freeware, commercial, or a mix of both (Esri, 2011; Fitzgerald et al., 2011).

Free software does not necessarily mean unpaid or non-commercial but rather free as in liberty. The Free Software Foundation (‘gnu.org’, n.d.) lists the four “essential freedoms” of free software as:

- **Freedom 0:** The freedom to run the program as you wish, for any purpose.
- **Freedom 1:** The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this.
- **Freedom 2:** The freedom to redistribute copies so you can help your neighbour.

- **Freedom 3:** The freedom to distribute copies of *your* modified versions, too. By doing this, the whole community has the opportunity to benefit from your changes. Access to the source code is a precondition for this.

Because being open source is a precondition, free software can be considered a subset of open source software. However, not all open source software is free software but, because most is, the two terms almost describe the same thing and are often used interchangeably (Fitzgerald et al., 2011).

Not to be confused with free software, **freeware** is proprietary software distributed without direct cost. Although it may still be commercialised indirectly, freeware cannot be classified as commercial software. Freeware can generally be considered a subset of proprietary software (Esri, 2011; Fitzgerald et al., 2011).

Software is classed as **commercial** if its purpose is to generate revenue directly and it can be free software (as defined above), open source, proprietary or a mixture (Esri, 2011; Fitzgerald et al., 2011).

2.4.1 Commercialisation

Commercial software, as defined above, is sold directly under some form of licence. Licence terms may include the length of time the software may be used, the maximum number of installations or concurrent users, or may even be pay-per-use (Fitzgerald et al., 2011). Whatever the terms, commercial software incurs direct financial costs. Non-commercial software, such as free and open source and even freeware, may still be commercialised indirectly, even if the software itself is non-commercial. Some examples of indirect revenue streams are:

- professional services
- support
- in-app advertising

- promotion of commercial versions of the same product, such as in the case of feature-limited freeware

It is a common misconception that free and open source software is not commercially viable, but the rewards for developing or contributing to open source software can be as intangible as ‘paying it forward’. The commercialisation of non-commercial software does not have to come from the owners themselves. For example, PostgreSQL is a free and open source database server, but many businesses and individuals earn revenue through professional services as product experts.

2.4.2 The pros and cons of open source software

With all the options available, it can be difficult to choose which software to use, and this problem is not unique to GIS and web mapping. It would be unwise to discount any particular software product purely on its distribution model, but rather a more holistic approach is required.

Pros:

- **Cost-free:** No initial capital expense.
- **Full customisability:** Those with the necessary skills can make changes at a code level.
- **Agility:** Tend to adapt to changing trends faster than proprietary systems.
- **Openness:** No motivation for proprietary formats and protocols.
- **Faster bug and security fixes:** Shorter release cycles than proprietary systems.
- **Code quality:** Continuous community review and a more diverse range of developers mean open source software can be of a higher quality than proprietary software.
- **Trust:** Because the code is open source, it is less likely it will contain spy features.
- **Teamwork:** The burden of development is shared amongst many developers.

- **Understanding:** Users can read the source code to see algorithms and functionality, bringing a better understanding of how the product works.

Cons:

- **Support:** Proprietary software more often has extensive product support available, including comprehensive manuals and training. Conversely, free and open source software commonly has little more than community support.
- **Development continuity:** Open source products often rely on a few developers who champion their continued development. If these developers lose interest, there are no guarantees the project will remain active.

(Fitzgerald et al., 2011; IDEA, 2011; Param, 2013; Synopsys, 2013).

Each product in a software stack or workflow needs to be chosen on the merits of functionality, compatibility, convenience and support, and these merits should then be weighed with a cost-benefit analysis. The shape of a software stack or workflow may incorporate a mix of distribution models, depending on requirements. Esri (2011) encourages hybrid models and, from a cartographic perspective, Jobst (2007) leans on the side of open source software, citing sustainability, spread to a wider audience, open documentation and expandability. For example, a GIS user's workflow may consist of a paid version of ArcGIS for routine work, the free view-only version of LP360 for quality checking of point cloud data, and the free and open source QGIS for custom or unusual functionality, which ArcGIS does not provide. Similarly, a server stack may use the commercial-proprietary ArcGIS for Server, running on the free and open source Ubuntu Linux operating system and PostgreSQL database server, simply because the systems administration team has vast experience with Linux and PostgreSQL but not with any free and open source GIS services. The overarching theme from the literature is to use whatever combination of products necessary to achieve the desired outcomes and that open source software can be just as a viable.

3 Web mapping of demographic diversity (paper)

This chapter is intended to serve as a standalone paper. It contains repetition from other chapters of this thesis with changes to tone and formatting to suit publication in a journal. As expected in journal articles, the results are concisely summarised. A more detailed description of the results, including comments, is provided as an appendix to this thesis.

3.1 Abstract

The use of maps on the World Wide Web is revolutionising how people access information. These web maps harness text, audio, video, animation, graphics and, most importantly, user interaction to improve the communication and interpretation of geographic phenomena. In this interactive and dynamic environment, web application developers have the opportunity to allow map readers to also be map producers; users can be enabled to query and analyse data and then assemble maps and other visualisations themselves, enhancing their understanding of the data and its geospatial relationships. This paper focuses on what factors are important for the design of web mapping applications that provide visualisations of demographic diversity. Demographic information is inherently geospatial, making maps the ideal tool for visualising population statistics and their spatial relationships. An outcome of this research is a set of guidelines for reviewing web mapping applications that portray demographic diversity information. This paper also provides an insight into the current standard for web mapping of demographic information and identifies the potential for improvement. A review of the literature, as well as a reflexive implementation, are the principle methods used to develop these guidelines. The essential components include the needs of demographers, and principles of cartography and human-computer interaction, as well as a consideration of open source and proprietary software. Five prominent, national demographic web mapping applications (two Australian, and one from New Zealand, the United States of America, and Canada) are used to develop these guidelines, as well as providing an insight into the current standard. The main conclusion of this

research is that the guidelines developed provide a worthwhile structure but need to be kept broad. Also, there needs to be more emphasis on user data upload facilities, which is important for demographers, and cartographic communication needs to be given higher priority.

Keywords: Demography diversity, cartography, human-computer interaction, web mapping.

3.2 Introduction

This paper develops a set of guidelines for evaluating web mapping applications and to answer the research question: *what factors are important for the design of web mapping applications that provide visualisations of demographic diversity?* It approaches this with the end user being a demographer wishing to perform custom analyses and visualise the results using maps. The creation of web maps for population data spans multiple disciplines, and this paper answers the research question by performing high-level literature reviews of demography, cartography, human-computer interaction and the software technology behind web mapping.

These reviews inform the development of evaluation guidelines which are then applied in the evaluation of five publicly accessible web applications which provide data and maps of national statistics. The results of these analyses are used to conclude what issues are currently facing these web applications in terms of communicating the spatial relationships of demographic information, and how this might be improved.

The relationships between demography, cartography, human-computer interaction and software technology are demonstrated in Figure 1. Demography is the subject user group, who also provide the information to be mapped and any interpretations of this information. Cartography defines the correct way to produce visualisations of the information using maps. This forms the link between demography and cartography, in the form of a workflow from data selection to data visualisation. Human-computer interaction defines the interface by which computers, in particular, web technology, enable the users to explore, select and analyse demographic data and create dynamic and sophisticated cartographic visualisations of that data. Human-computer interaction is underpinned itself by the software technology powering web applications.

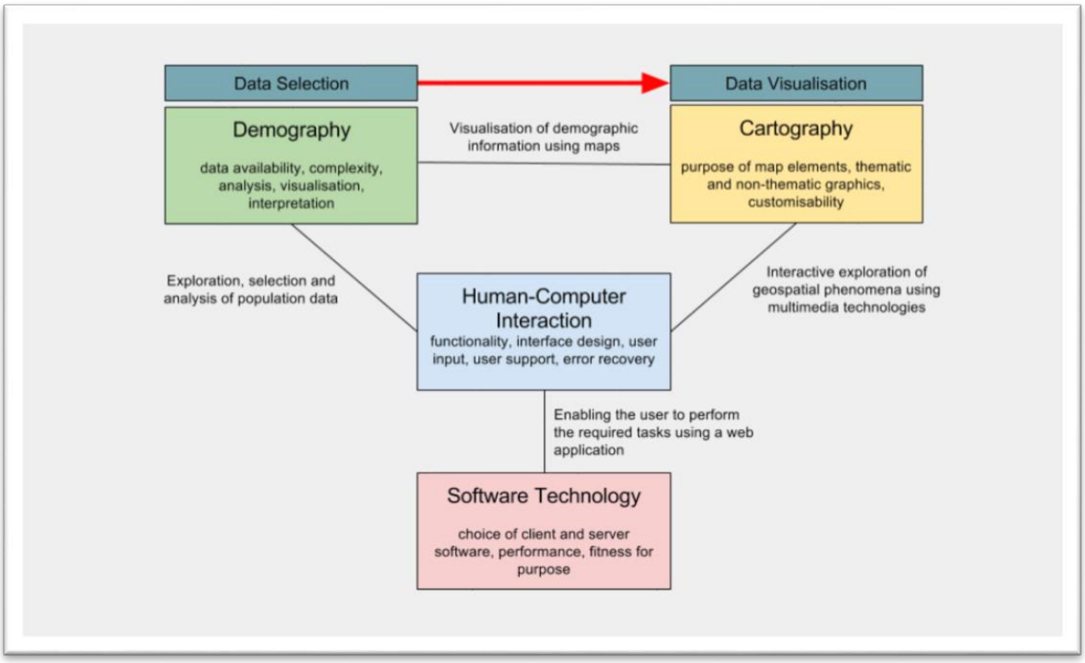

Figure 4: Relationships between disciplines of demography, cartography and human-computer interaction, underpinned by software technology.

Previous research has focussed more on the usability and functionality of web applications and web maps (see, for example, Komarkova et al., 2010, 2007) and has developed very specific evaluation criteria which overlook the importance of the map graphics themselves in terms of cartographic communication. This paper takes demography and cartography into account in an attempt to highlight a need for a more balanced approach to web map evaluation beyond just interface usability and functionality. These previous studies also ask very specific questions, often regarding the inclusion or exclusion of specific functionality. However, web applications are hugely variant and there may be perfectly valid reasons why specific items might be limited or missing. This paper develops deliberately broad guidelines that allow evaluators more flexibility and subjectivity. To achieve this, this paper takes a fresh look at the fundamentals of the four topics of demography, cartography, human-computer interaction and software technology for web mapping and uses this to make an informed decision on producing its own set of guidelines.

3.3 Demography

Demography is a science of populations (Lundquist et al., 2015; Weinstein & Pillai, 2016) and is explicitly linked to very basic aspects of humanity – when and where we are born, where we are from, where we live, when and where we die. The patterns which emerge when this data is analysed across a whole population provides answers to questions about what has happened, what is happening, and, perhaps most importantly, what might happen to the shape of our groups, communities, regions and nations.

In a 2009 survey by Van Dalen and Henkens (2012) across 970 demographers worldwide, 46 percent of respondents described their practice as applied, 19 percent fundamental, and 36 percent as a mixture of the two. When asked specific questions relating to their profession, demographers indicated the most important population issue for Asia/Oceania in the next 20 years as population ageing, followed by urbanisation and large-scale migration flows. Of particular interest from this survey was that demographers prefer to stay close to the data with 59 percent of respondents describe data analysis as highly rewarding.

Population diversity relates to differences occurring between various characteristics of a population such as gender, age, ethnicity, culture, religion, language and socio-economics. It also includes how this occurs both spatially and temporally, such as the changing structure of a population as it ages or the arrival of different ethnicities due to migration. Because of the nature of populations, and especially with increasing urbanisation, population diversity should be considered at multiple scales, such as national, regional, local and neighbourhood, to best examine spatio-temporal heterogeneity (Nijkamp & Poot, 2015). For example, when examining cultural diversity, measurement across multiple scales is important to show where the intermingling of cultures occurs rather than more homogeneous neighbourhoods in an otherwise culturally diverse region. Diversity is usually enumerated as a fractionalisation index; essentially, the probability of an individual from one cohort encountering another individual from a different cohort, and these indices are often mathematically similar to the indices used by biologists and ecologists (Nijkamp & Poot, 2015).

3.3.1 Typical analysis workflow

The typical analysis workflow of a demographer commonly falls into two paths to obtain data for visualisation and this depends largely on the type or depth of analysis they are performing, and whether they require the use of *raw* or *pre-processed* data. For the first path, demographers are often required to answer very specific and unique research questions which require highly bespoke analysis and a demographer undertaking some deep analysis will require raw data, often from multiple sources and in different formats, and might spend considerable time processing and analysing this data. Typically, such complex numerical analyses would be undertaken using software such as Microsoft Excel. However, in the second path, some demographers may have a use for pre-processed data, such as fractionalisation indices if this is available to them.

Raw population data typically refers to absolute counts for a particular variable, for example, identified ethnicity. This is opposed to pre-processed data, which might include pre-calculated ratios, rates, proportions and indices. Population data is most often aggregated by some areal unit (e.g. census unit, city, region and country) to protect the privacy of individuals within the counting area. In addition, the units of the variable itself may be aggregated, for example, ages grouped in five-year ranges. While the exact locations of individuals in a dataset naturally provides the best spatial resolution, it has to be accepted that population data, especially from censuses, will be aggregated and analysts must work within this constraint.

3.3.2 Web application workflow

So, what does a demographer need from a web application in order to visualise population diversity using maps? This paper breaks this question into a workflow of three phases:

Phase 1. Data procurement

Phase 2. Data analysis

Phase 3. Visualisation and interpretation

Ideally, a single web application could be used to deliver this entire workflow. A key issue here is that a web application attempting this phase would need to emulate the complex functionality of spreadsheet software. Therefore, it is suggested that the notion of performing the data analysis phase within the web application is abandoned, and left to more appropriate software like Microsoft Excel. Phases one and three are now standalone functions which can be invoked optionally and independently. We will consider the requirements for the data procurement and visualisation and interpretation phases, and treat these as independent of the overall workflow.

3.3.3 Data procurement

When considering a single web application that both provides data and generates visualisations, one might initially assume that the data selected will be the data used by the visualisation. However, because an analysis might use data from several sources, the original data does not necessarily have to be sourced from this web application.

Because analysts are particularly interested in sourcing raw data, the user should be able to explore and choose the variables as required for their analysis, and to preview the data in tabular form. There should be a wide variety of characteristics available as absolute counts, such as age, ethnicity, language and gender. It is helpful for these characteristics to be available in varying levels of complexity, for example, age in one-, five- and ten-year groups. Additionally, the data might include pre-calculated ratios, rates, proportions or indices. Metadata should be available, including the source of the data and any caveats or limitations, such as ethnic overcounts.

In order to assess heterogeneity across multiple levels, the spatial resolution of the data should be available in a variety of levels and complexities, such as national, regional, city and neighbourhood, and these areas might not necessarily overlap.

The web application should allow the data to be selected, grouped and summarised with a high degree of flexibility. In order to analyse change across time, the data should be available for a range of time periods and, for population data, this will probably mean the most recent census and several censuses into the past or future. Finally, in order to be analysed using third-party software, the data must be downloadable in a common data format, such as Excel Workbook (xls or xlsx) or comma-separated values (csv).

The visualisation and interpretation phase can be performed by a separate web application to the data procurement phase. The demographer should be able to upload the results of their analysis for visualisation. Additionally, the uploaded data would need to have a spatial element so it can be mapped. Aside from maps, other graphics, such as charts and infographics, can be used to enhance the interpretation of the data and the map.

3.4 Cartography

According to MacEachren, Crawford, Akella, and Lengerich (2008), more emphasis has been placed on developing the tools necessary for web mapping rather than the design principles. Peterson states that “cartography today is as much about [computer programming] as it is about design and data analysis skills” (2015, p. 240). However, Peterson also states that cartographers will still have much to tweak and tune until software improves to the point where most facets will be automatable.

Cartography aims to reduce reader error and misinterpretation by using suitable graphic presentation. Kraak and Ormeling (2011) state that *cartographic grammar* is well-defined and that the basics do not change for the web. Beyond these basics, Peterson asserts that digital map design is still in its youth, and that "common rules or standards for cartographic presentations on screen displays are not defined yet" (2015, p. 55) and there is minimal published literature on the subject, leaving much to trial and error. M. P. Peterson (as cited in Cartwright & Peterson, 2007) asserts that "paper thinking" still over-influences multimedia cartography.

3.4.1 Essential map components

3.4.1.1 Basemaps

Maps can be broadly categorised as either *general reference* or *thematic* (Robinson, 1995). General reference maps describe the location and character of physical objects, and thematic maps describe the distribution of measurable phenomena, such as human populations. Thematic maps still require a degree of general reference detail, in the form of a *basemap*, so the reader is able to spatially orient and appreciate the location and distribution of the phenomena in relation to the physical world might and include roads, place names, coastlines buildings and bodies of water or other physical objects (Cammack, 2007; Kraak & Ormeling, 2011).

3.4.1.2 Map elements

The first essential consideration of map design is that the map must communicate a clear purpose. In the case where mapping, for example, demographic diversity, the demographers are providing meaning to numeric data and this needs to be clearly portrayed by the map's theme. The map's title should include the area, the common denominator of the theme and the time. The purpose of the map needs to be clear to the target audience and they must be kept in mind during the design process (Kraak & Ormeling, 2011; G. N. Peterson, 2015; Robinson, 1995).

Other essential cartographic elements for print maps are north arrows, insets, scale bars, graticules and legends. However, this paper suggests that, because web maps are a different environment to print maps, the cartographer should feel free to make a judgement call as to whether these elements are truly required or if they could negatively impact effective communication. For example, a legend might not be required if the web map uses interactive pop-ups to describe a symbol. Symbols, lines and text should be visually distinct and readable and good maps will establish a clear visual hierarchy by contrasting the theme from the basemap. The map should only convey one or two phenomena as trying to include too many variables can be detrimental to communication. Additionally, data sources should be appropriately attributed (Kraak & Ormeling, 2011; G. N. Peterson, 2015; Robinson, 1995).

It can be useful to display two or more maps simultaneously, and can show a single phenomenon across different areas or times or a single area with multiple phenomena or times. Where two or more maps are being compared, the theme or the area (and scale) must remain constant when the opposing element changes (Ormeling, 2007).

3.4.1.3 Scaling and generalisation

It is not merely sufficient to simply display a basemap as it would be on a paper map; web maps can be zoomed, and the level of detail should be appropriate for the particular scale being displayed. At larger scales (zoomed in), a map could include features such as local roads and buildings, and the shape of rivers might be represented using polygons. As scale decreases (zooming out), these features become less relevant and cause the map to become cluttered and unreadable. To overcome this, web maps should employ g*eneralisation* – the elimination, simplification, displacement and enlargement of cartographic features – to achieve consistency in the level of detail across all zoom levels. For example, road detail could be decreased to show only highways, streams might be removed and rivers only displayed as simple lines (Gaffuri, 2011; Kraak & Ormeling, 2011; Miller, 2007; G. N. Peterson, 2015).

3.4.1.4 Projections

The standard projection used by web mapping software is Web Mercator and this is often the only option available. While Web Mercator preserves direction and shape, and results in a squareness which is useful for tiling purposes, it can produce noticeable distortion at a continental level and nearer the poles (G. N. Peterson, 2015). This effect may be noticed when comparing the shape of areas on small-scale overviews to the shape of those same areas when viewed at larger scales.

3.4.2 Cartographic communication

Like print maps, the primary function of (most) web maps will be the *communication* of spatially related information. In the case of communicating demographic information – among others – this means using some manner of *thematic map* to portray the spatial occurrence of a phenomenon. Maps communicate using the principles of *cartographic grammar* and *cartographic convention*. General reference maps, such as basemaps and topographical maps use cartographic convention, for example, the forest is green, the water is blue and north is up. By comparison, thematic maps employ the broader rules of cartographic grammar (Kraak & Ormeling, 2011). Substantial research on the neurophysiological, psychological and cognitive backgrounds of cartographic grammar underpins cartographic literature and has informed the general understanding of how humans see and interpret maps. The interested reader should look to works such as MacEachren (1995) for an explanation of this topic.

3.4.2.1 Map types

Map data can be qualitative or quantitative, and quantitative data can be absolute or proportional. Before choosing a map type, spatial data should be assessed to determine its nature, length and range. It is also useful to categorise quantitative data as either *spatially extensive* or *intensive*. Spatially intensive data is calculated using a spatial variable, for example, population density, where the denominator is the actual area of the aggregation unit. Conversely, spatially extensive data is not derived from a spatial variable and an example of this is absolute population counts (Buckley, 2013; Haining, 2003). The data's scale, and whether it is spatially intensive or extensive, should inform the choice of map type.

Choropleth maps use discrete, stepped colours to represent quantitative differences between areas. Typically, choropleths are used to visualise statistics which have been aggregated by counting areas, such as census units. This aggregation yields maps which tend to convey homogeneity in the distribution of the statistic, and it is important that choropleths be used with caution when mapping spatially extensive data. Choropleth boundaries may be arbitrary, and may not necessarily represent the distribution of the phenomenon, or may be subject to the modifiable area unit problem and ecological fallacies. Despite their drawbacks, choropleths are the most commonly used map for visualising socio-economic phenomena due to their ease of construction and interpretation. **Chorochromatic** maps are similar to choropleths but use colour to represent qualitative, rather than quantitative data. Choropleths should use differences in lightness or saturation to represent different values, while chorochromatic maps should only use changes in hue (Buckley, 2013; Haining, 2003; Kraak & Ormeling, 2011; Meirelles, 2013).

Cartogram maps are a clever alternative to choropleths where the sizes of areas are changed relative to the occurrence of the phenomenon being mapped, in order to overcome the issue of false ordinal importance mentioned above (Kraak & Ormeling, 2011; Meirelles, 2013) but can distort the size and shape of areas so much that they should be treated more as a supplementary tool. **Dasymetric maps** use additional topographic information to create areas which better represent the true distribution of the phenomenon being mapped, in order to reduce perceptions of homogeneity (Kraak & Ormeling, 2011; Meirelles, 2013).

Dot density maps use dots to convey the density of a phenomenon in a particular area. However, the use of dots can lead to a false perception of absolute location when used with aggregated data (Kraak & Ormeling, 2011; Meirelles, 2013).

Proportional symbol maps use geometric shapes of varying size to convey difference in value. The use of simple geometric shapes as symbols is important; the more complex the shape, the more difficult it is to perceive a difference in size. The range of values is important and the minimum and maximum values should be shown in the legend (Kraak & Ormeling, 2011; Meirelles, 2013).

Isopleth maps use *isolines* to partition intervals of a continuous phenomenon where there is a continuous distribution in all directions. While isopleths are not as intuitive as choropleths, they are appropriate for comparing different phenomena and are therefore more effective than choropleths because they can be used to demonstrate trends and the direction of change (Kraak & Ormeling, 2011; Meirelles, 2013).

Statistical surface maps are useful as dramatic three-dimensional visualisations of the peaks and troughs of a phenomenon. However, the rendered image may become overly complex and background areas, including basemaps, can be obliterated by peaks in the foreground (Kraak & Ormeling, 2011).

3.4.2.2 Colour and symbology

Once the correct map type is determined, the way the map is symbolised needs to be considered. Convention is important for map interpretation, and conventions for topographic features are well defined, e.g. water is blue, vegetation is green, and hospitals are marked with red crosses. Thematic maps do not necessarily have such conventions so cartographers must look to broader guidelines for colour and symbol choices. Bertin (as cited in Kraak & Ormeling, 2011) lists six graphic variables which can be exploited by cartographers for symbolising differences in qualitative and quantitative values:

Qualitative	Quantitative
• Hue • Shape • Textural orientation	• Symbol size • Lightness • Texture

Table 2: Cartographic variables for representing qualitative and quantitative differences.

Kraak and Ormeling augment Bertin's list with *saturation*, *arrangement* and *focus* and, in the context of a web application, *blur*, *focus*, *opacity*, *shadow* and *shading*. The hierarchy of information will translate into a visual hierarchy, and this visual hierarchy will determine which variables are used for each feature, e.g. the colour of the sea should not dominate a map explaining strictly land-based phenomena (Kraak & Ormeling, 2011; Miller, 2007; G. N. Peterson, 2015).

I mention above that choropleths should vary in lightness and chorochromatic maps should vary in hue. This is because changes in lightness or shade are perceived as quantitative differences, e.g. pink implies a lower density than red, while changes in hue are perceived as qualitative differences, provided all hues have a similar lightness. Along with hue and lightness, care needs to be taken with colour saturation; highly saturated colours tend to dominate an image so these should be reserved for smaller areas. The maximum number of colours which should be used varies by source but, as a guide, eight to ten individual colours and five shades of the same hue are discernible at a glance (Kraak & Ormeling, 2011; G. N. Peterson, 2015).

3.4.2.3 Spatio-temporal relationships

Static maps can only display a single snapshot in time, making it difficult to represent spatio-temporal change. While change can be represented using a series of static maps, as would be required if the maps were printed, a major benefit of on-screen cartography is that map themes themselves can be *animated,* which offers a very expressive way to represent changes in location and value (Kraak, 2007; Kraak & Ormeling, 2011). Where animations are provided, the user should be able to control the display with controls for play, pause, fast-forward, rewind at a minimum. While the rules for static maps are well defined, animation opens up a new realm of cartographic possibility, the effects of which are not yet fully understood and are currently being researched (Kraak, 2007; Kraak & Ormeling, 2011).

3.4.3 Map interactivity

In an interactive environment, users become the map *producers*; they want to assemble the pieces themselves and they may want to spend considerable time customising and exploring (Buchroithner & Gartner, 2013; Cartwright & Peterson, 2007). Cartwright and Peterson claim this is due to how humans learn and structure knowledge because interaction is fundamental in how we form knowledge. With an estimate of more than half of the educated population not being competent with maps, interactive multimedia cartography may help improve knowledge acquisition amongst these users (M. P. Peterson, 2007a).

Miller (2007) talks of the "hypermedia paradigm" by which users of web maps are now conditioned to expect extra information to be available by interacting with the map. To be effective, any interactive map elements must balance their cartographic symbology and iconology with the need to depict that the element is, in fact, interactive; web maps should actively promote interaction by indicating which features are interactive, such as the cursor changing on hover. The map object should be manually navigable by the user with pan and zoom tools, possibly alongside links that zoom to predefined areas.

3.5 Human-computer interaction

Human-computer interaction (HCI) as a medley of several separate disciplines such as psychology, cognitive science, ergonomics, sociology, computer science, engineering, business and graphic design, to name a few. A fundamental principle of HCI is that computers and software should be designed in a user-centric way and with the understanding that users have specific tasks in mind and want to achieve those tasks in ways that are seamless with respect to their lives and workflows outside the computer. HCI is not just about pretty interfaces; it is integral to the design of systems as a whole (Dix et al., 2005; Taylor & Lauriault, 2007).

The science of HCI is broken down into input-output channels between humans and computers. A typical consumer computer has a monitor and speakers (output) and a keyboard and mouse (input). Conversely, humans have eyes and ears (input) and limbs (output). Although many other inputs and outputs are possible, this demonstrates the input-output channels of monitor-to-eyes, speakers-to-ears and limbs-to-keyboard-and-mouse (Dix et al., 2005).

The usability of a system is governed by how well the user can interact with it. In web mapping, pointing devices (mouses, trackpads, etc.) play the most important role, allowing the user to move the cursor, operate controls to query the data, pan and zoom map objects and interact with points or areas of interest. Fitts' Law (cited in Dix et al., 2005) is a formula which describes human accuracy when reaching for targets at increasing distances, such as moving a mouse to a button. Fitts' Law is an important tool for analysing user interface design as it shows that relatively smaller targets at relatively larger distances are increasingly difficult to reach, meaning commonly used controls should be grouped near each other or, if they are distant, targets should be of a sufficient size to hit them easily.

3.5.1 Design principles

Dix et al. (2005) discuss design principles and guidelines for applying HCI in practice without needing to understand the various disciplines which underpin them. The principles often overlap and may enhance or interfere with each other, meaning they need to be applied with careful thought. A distilled list of HCI principles from Dix et al. is below.

Learnability:

- **Predictability:** No surprises. Past interaction guides future interaction as recognition is more powerful than recall. All operations a user can (or cannot) perform should be visible.
- **Synthesisability:** Users should be able to determine that change has occurred as a result of an operation. This helps build a mental model of how the system operates.
- **Familiarity:** The user should be able to apply knowledge of other systems (computer or not) in the operation of the current system.
- **Consistency:** Similar to convention in cartography. Red-coloured feedback means ‘bad’, green-coloured feedback means ‘good’. Arrow keys perform directional operations.

Flexibility:

- **Dialogue initiative:** Interaction can be initiated by the system or by the user. User pre-emptive dialogue, where the user initiates action, is preferred over system pre-emptive, where system initiates action. This gives the user a feeling that they are in control and have the ability to perform actions as they need.
- **Multi-threading:** Allow multiple tasks to occur at the same time. Windowing systems are the most common example of this in practice. Modern web browsers allow this to occur seamlessly in a web context without having to build this into the product.
- **Task migratability:** Task execution can be transferred from the user to the system and vice versa. Mundane actions, like spell checking, can be performed by the system, but it would be unwise to allow the system to always autocorrect.
- **Substitutivity:** Allowing flexibility in the way a user interacts, e.g. when entering values either “2.5 cm” or “1 inch” are acceptable.
- **Customisability:** User-initiated customisability is *adaptability,* and system-initiated customisability is called *adaptivity*. Users should be able

to customise the way they use the system, such as moving toolbars, while the system might automatically adjust itself if it detects an advanced user over a novice. Adaptivity is likely to interfere with predictability if it is not done well.

Robustness:

- **Observability:** The user can determine the internal state of the system by its output.
    - **Browsability:** The state of the system can be explored via the interface
    - **Defaults:** Suggested input can quickly be perceived as correct or not and reduces physical action. May evolve adaptively.
    - **Reachability:** Ability to navigate through observable states.
- **Persistence:** System outputs should represent system state persistently. For example, a user may hear a 'beep' indicating a new email has arrived but will soon forget. A persistent visual notification will serve as an additional reminder.
- **Recoverability:** Humans make mistakes, and a system should allow recovery from these. Undo, redo, back, forward commands are examples of this.
- **Responsiveness:** Users, particularly on the web, are impatient, and the performance of the system is key to usability. Response time should be fast or at least indicate that the system has not stalled.
- **Task conformance:** Does the system perform all the required tasks? Does it fulfil the user's requirements and expectations based on their understanding of the tasks?

3.5.2 Evaluating usability

Beyond the HCI fundamentals above, there are already established approaches to how the usability of software is evaluated. Komarkova et al. (2007) have developed 138 guidelines for evaluating the usability of web-based GIS.

However, this list focusses in detail on the usability of the GIS itself rather than the cartographic communication of the output. Travis (2016) develops a more general approach with a comprehensive 247-point list of guidelines for evaluating usability, task orientation, navigation, data entry, content quality, layout and design, user support and fault tolerance, specifically for web applications. Travis' guidelines, as well as more fundamental heuristics from Travis and more broad heuristics from Nielsen (1995) have been adapted in the HCI factors in Table 4.

3.6 Software technology

The free and open source movement has become a cultural revolution, and this is not limited to software; we now have open source cars, electronics, and even firearms. The decision between significant capital expenditure in more a traditional commercial software versus using a competing and ostensibly free alternative is not obvious, and organisations initially approached this choice in a "reluctant and uncomfortable manner" (Fitzgerald et al., 2011, p. vii).

The terminology used to describe software distribution models can be a source of confusion. This is not an exhaustive list, but it does serve to generalise the most common paradigms of software distribution:

Open source software means the product's source code has been made publicly available. Open source software is usually non-commercial — although there are exceptions to this — and it is not necessarily free software (see definition below) (Fitzgerald et al., 2011; 'gnu.org', n.d.; Open Source Initiative, n.d.).

Proprietary software is the opposite of open source and is sometimes referred to as closed source (not to be confused with commercial software). Proprietary software may be freeware, commercial, or a mix of both (Esri, 2011; Fitzgerald et al., 2011).

Free software does not necessarily mean unpaid or non-commercial but rather free as in liberty. the freedom to use the software as you wish, to study how the software works, to change it and redistribute your changes to give the whole community the opportunity to benefit from your changes ('gnu.org', n.d.). Free software can be considered a subset of open source software. However, not all open source software is free software but, because most is, the two terms almost describe the same thing and are often used interchangeably (Fitzgerald et al., 2011).

Not to be confused with free software, **freeware** is proprietary software distributed without direct cost. Although it may still be commercialised indirectly, freeware cannot be classified as commercial software. Freeware can generally be considered a subset of proprietary software (Esri, 2011; Fitzgerald et al., 2011).

Software is classed as **commercial** if its purpose is to generate revenue directly and it can be free software (as defined above), open source, proprietary or a mixture (Esri, 2011; Fitzgerald et al., 2011).

Commercial software is sold directly under some form of licence, the terms of which may include the length of time the software may be used or the maximum number of concurrent users. Whatever the terms, commercial software incurs direct financial costs. Non-commercial software, such as free and open source and even freeware, may still be commercialised indirectly, even if the software itself is non-commercial, through channels such as user support or advertising.

3.6.1 The pros and cons of open source software

With all the options available, it can be difficult to choose which software to use, and this problem is not unique to GIS and web mapping. It would be unwise to discount any particular software product purely on its distribution model, but rather a more holistic approach is required.

Pros:

- No initial capital expense
- Full customisability

- Agility
- Openness
- Faster bug and security fixes
- Code quality
- Trust
- Teamwork
- Visible source code

Cons:

- Reduced support
- Risk to development continuity

(Fitzgerald et al., 2011; IDEA, 2011; Param, 2013; Synopsys, 2013).

Each product needs to be chosen on the merits of functionality, compatibility, convenience and support. The shape of a software stack or workflow may incorporate a mix of distribution models, depending on requirements. There is no perfect one-size-fits-all solution and use cases must be analysed on a case-by-case basis.

3.6.2 Evaluating software technology

Web applications operate using a client-server model where the client software, usually a web browser, communicates with a remote web server over a network, usually the Internet. While the browser can be thought of as client software, it should really be considered as a platform on which web applications can be run and that these web applications themselves are the real client software. The client and server software together form a web application's software stack, and the two are implemented to work together as one cohesive unit. However, the two components are still very much separate, being executed in the browser and server respectively. Because of this separation, it is useful to consider the two separately when evaluating a product.

For the purposes of evaluating web applications, it is often not possible for an end user to identify the exact software being used. The reasons for this include:

- JavaScript code may be obfuscated and minified which makes reverse engineering difficult.
- Software for Adobe Flash or Microsoft Silverlight plugins is distributed in binary.
- The server software might not identify itself through HTTP headers or other means.

For these reasons, it is not appropriate or fair to score a web application on the particular software used. However, in many cases, there is sufficient evidence to positively identify the software used and, in these cases, this can be commented on. Beyond the specifics of the software, web applications can be scored in terms such as performance and whether the client software requires any plugins in order to run in a browser.

The performance of web applications is measured in terms of the responsiveness of the graphical user interface to user interactions and the responsiveness of the server to requests for data. However, the source of performance issues is not always obvious and can be affected minute-to-minute by factors such as server load and network speed. Performance can be enhanced using a variety of techniques, such as content delivery networks, and caching on the server and in the browser and the use of these techniques should be taken into account when scoring.

The use of runtime environment plugins such as Adobe Flash and Microsoft require users to install extra software in order to use the web application. This task may not be achievable by the user due to their level of skill, IT infrastructure constraints or browser incompatibility.

This is not intended to form a comprehensive review of the software technology used by an web application as this is not necessarily achievable from an end-user perspective. Rather, it is intended to cause readers, evaluators, developers, etc.

to take pause and consider whether the software used is fit for purpose and performing well. The guidelines suggested are ‘talking points’ which might form a framework to which further detail can be attached in this area.

3.7 Method

The development of this method and the scoring guidelines has been a reflexive and reflective process. Initial scoring guidelines were written as a checklist of specific criteria but this was quickly determined as too restrictive to be effective with the myriad and varied web applications which might be subject to evaluation. Further development saw the guidelines being rewritten as more open-ended questions which are divided into groups under demography, cartography, HCI, and software technology, e.g. ‘user interface design’ is a group under HCI. The resulting guidelines are shown in Table 4.

Howells (2011) and Komarkova, Sedlak, Novak, Musilova, and Slavikova (2011) evaluate web applications using systems where each guideline (or heuristic) is assigned a score in a narrow range three to four values. This same approach has been adapted in this research, but instead of assigning a score to each individual guideline, scores are applied to each group of guidelines in order to allow more flexibility; however, this has led to a somewhat subjective method of applying scores. This is reflected in the scoring model; the narrow band of zero to three recognises that the subjective evaluation process is applied easiest if a web application can be scored in terms of “low”, “medium” or “high” compliance, as shown in Table 3. A score of zero can be applied where the web application makes no attempt to comply with any of the guidelines. If a web application does not strictly meet the guidelines for a particular evaluation factor, but shows an ingenious or effective implementation in other ways related to that factor, then the evaluator is at liberty to raise the score. Conversely, if many guidelines are met but the implementation is of poor quality then the evaluator is free to lower the score. The evaluator should not feel compelled to award a lower or higher score based strictly on the guidelines if they deem the factor to be met or unmet in the web application’s wider context.

Score	Compliance	Criteria
0	No attempt	The web application does not attempt to comply and no guidelines are met.
1	Low	An attempt at compliance but few guidelines are met.
2	Medium	Moderate compliance and some guidelines are met.
3	High	Excellent compliance and most or all guidelines are met.

Table 3: Score interpretations for individual evaluation factors.

This subjective method of scoring does, of course, mean that the opinion of the evaluator affects the way scores are applied, both in terms of personal bias and experience and knowledge in the evaluation factors. This would make it problematic to compare results between different evaluators. If this comparability were necessary, it would be recommended to follow the example of Komarkova et al. (2007) where evaluators first rank the importance they placed on each factor and used this to weight their scores.

For consistency, all evaluations were undertaken by the author using Google Chrome on a 2011 MacBook Pro with a 2.4GHz Intel Core i5 processor and 16GB RAM, viewed on a 24-inch 1080p monitor. Google Chrome's *incognito mode* was used to ensure a "clean slate" for each evaluation.

3.7.1 Scoring guidelines

Demography
Data procurement: • Can the user find and select data with a high degree of flexibility? • Are there multiple characteristics variables, particularly those that account for diversity, available in multiple levels of conceptual granularity? • Is the data available for multiple spatial scales? • Are there any useful, pre-built analyses such as ratios, proportions and high-level indices? • Is the data available for a number of temporal scales? • Can the data be previewed as a table? • Is the data available for download in a common data format? • Is there metadata available, including warnings or caveats for any limitations? • Can the user upload their own data? • Are graphs and other useful visual devices used to enhance the interpretation of the map and the data?
Cartography
Cartographic communication: • Is the most suitable map type being used for the information being communicated? • Does the theme's colour scheme reflect the theme of the information and is the colour progression correct for the type of data? • Are all symbols readable and representative of the phenomenon?

• Is there a visual hierarchy between foreground and background (theme and basemap) which is also harmonious with the visual hierarchy of the web application? • Are multiple maps used to compare multiple areas, times or phenomena? • Is animation used to demonstrate change over time? • Is the user able to customise the cartographic output? • Do the tools for customisation make it easier to follow cartographic best practice than to break it? Map interactivity: • Can the map be panned and zoomed (unless it is designed to show only a fixed extent)? • Is extra information available by interacting with the map and does the map actively encourage interaction? • Do animations have controls for play, pause, forward, backwards, etc.?
Human-computer interaction
System functionality: • Is the purpose of the web application clear and does it conform to the task is designed for? • Does the web application mimic real-world workflows and use familiar language? • Does the web application automate as many tasks as possible and are common or repetitive actions easy to perform? • Is the web application free from unnecessary features, complications, confusions or distractions? • Is there a low barrier to entry? • Is the web application free from errors and broken functionality?

User interface design:

- Are there any substantial violations of Fitts' Law?
- Is the user interface legible and coherent?
- Are the controls and layout consistent with convention among contemporary systems?
- Does the user interface avoid horizontal scrolling (with the exception of map panning)?
- Are all active elements (controls) visibly distinct from passive elements and do they visually invite interactivity? Conversely, are all passive elements obviously non-interactive?
- Is the amount of user input (clicking, scrolling, typing) required to perform a task kept to a reasonable minimum?
- Is the web application's status visible through confirmations and progress bars?
- Do potentially dangerous actions require user confirmation?

User input:

- Do inputs contain default values as appropriate?
- Are all inputs labelled concisely and unambiguously?
- Are all inputs sized according to the length of data they are expected to contain?
- Are inputs grouped and ordered in a logical and natural way?
- Are pre-defined options provided as appropriate using drop-down menus, radio buttons and checkboxes?

User support and recovery:

- Is there a comprehensive user guide, possibly supplemented with context-sensitive help?
- Are all error messages, warnings and notices coherent and do they guide the user to a solution?

• Is the user's work recoverable in the event of a system or user error? • Are there undo and redo controls?
Software technology
Client-side software: • What is the distribution and licencing model of the software[3]? • Are any plugins required for the software to run? • Does the software make efficient use of resources such as CPU, RAM and network? • Is the software using the latest or a near-latest version? • Is the software suited to the task with no apparent issues relating to the choice of software? • Are there any relevant critical reviews of this software? Server-side technology: • What is the distribution and licencing model of the software[3]? • Does the server respond to most individual requests (not necessarily whole page loads within two seconds)? • Does the system make use of client- and server-side caching techniques? • Is the software using the latest or a near-latest version? • Is the software suited to the task with no apparent issues relating to the choice of software? • Are there any relevant critical reviews of this software?

Table 4: List of evaluation guidelines.

[3] Comment only for reference; should not affect the score as there is no right or wrong method.

3.7.2 Chosen web applications

All five web applications are provided by government organisations with the common objective of distributing statistical data and visualisations. This was to ensure some uniformity in purpose between web applications. Screenshots of the chosen web applications are shown in Figure 5.

3.7.2.1 Australian Bureau of Statistics' "TableBuilder"

http://www.abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder

The primary purpose of this system is to make census data available for viewing and download. However, this web application was chosen for evaluation because graph and map builders are also provided for visualising selected data. The "Basic" version is free while the "Pro" version is paid with the advantage being access to more complex data. The workflow is very data-oriented, and the tools for compiling tables are impressive. Tables, graphs and maps can be exported in a variety of formats although exporting a map in PDF format does not work. Mapping functionality is provided by Esri's ArcGIS for Server and ArcGIS API for JavaScript, which are designed to be used in tandem. The level of functionality of the system as a whole is impressive, and the freely available data covers a comprehensive range of census statistics.

3.7.2.2 Statistics New Zealand's "StatsMaps"

http://www.stats.govt.nz/statsmaps/home.aspx

StatsMaps is a free web application which provides access to thematic maps of certain New Zealand statistics and also provides access to the latest census area unit and electorate boundaries. The data is pre-categorised by themes such as ethnicity or age, and there are no options to customise which data is used; e.g. it is not possible to create a map showing the percentage of just Maori males. While this makes for a very simple approach which may serve the majority of use cases, it lacks the flexibility required by serious analysts[4] and was chosen primarily to demonstrate the limitations of a New Zealand-specific example. Mapping functionality is provided by Esri's ArcGIS for Server and ArcGIS API for JavaScript.

3.7.2.3 Statistics Canada's "GeoSearch"

http://geodepot.statcan.gc.ca/GeoSearch2011-GeoRecherche2011/GeoSearch2011-GeoRecherche2011.jsp?lang=E&otherLang=F

This interactive web mapping application is designed for finding places in Canada, locating them using a map and viewing basic information about those places. It has two versions available: one each for the 2011 and 2006 censuses. The basic tabular data is reasonably comprehensive, but the selection of thematic maps is limited by comparison. This appears to be a system which has been custom built for Statistics Canada and does not appear to use any particular client or server software which can be readily identified.

3.7.2.4 United States Census Bureau's "Census Data Mapper"

http://www.census.gov/geo/maps-data/maps/datamapper.html

This tool is intended to allow users to view, save and print United States county-based demographic maps from the 2010 Census. It was chosen for evaluation as it is a very simple tool with very limited data options but it is of interest to this paper as it does allow the user to customise some aspects of the cartographic output. Also of interest is the use of Adobe Flash for the user interface.

3.7.2.5 Australian Government's "NationalMap"

https://nationalmap.gov.au/

[4] Statistics New Zealand has an excellent tool called NZ.Stat for accessing pure data but it does not make use of maps. If NZ.Stat and StatsMaps were combined they would be an extremely powerful tool.

NationalMap is an initiative of the Australian Government's Department of Communications and Arts (now managed by the Department of the Prime Minister and Cabinet) for viewing spatial data across multiple agencies of both the Federal and state governments. NationalMap is of particular interest to this paper because it is an entirely open source project which itself leverages several other open source projects, most importantly TerriaJS, which is the framework the web application is built on. TerriaJS is entirely client-based software built using HTML and JavaScript, and data is accessed directly from the source agency. However, TerriaJS does have a small server component that runs on NodeJS for accessing data from services which the client software cannot access directly for various reasons.

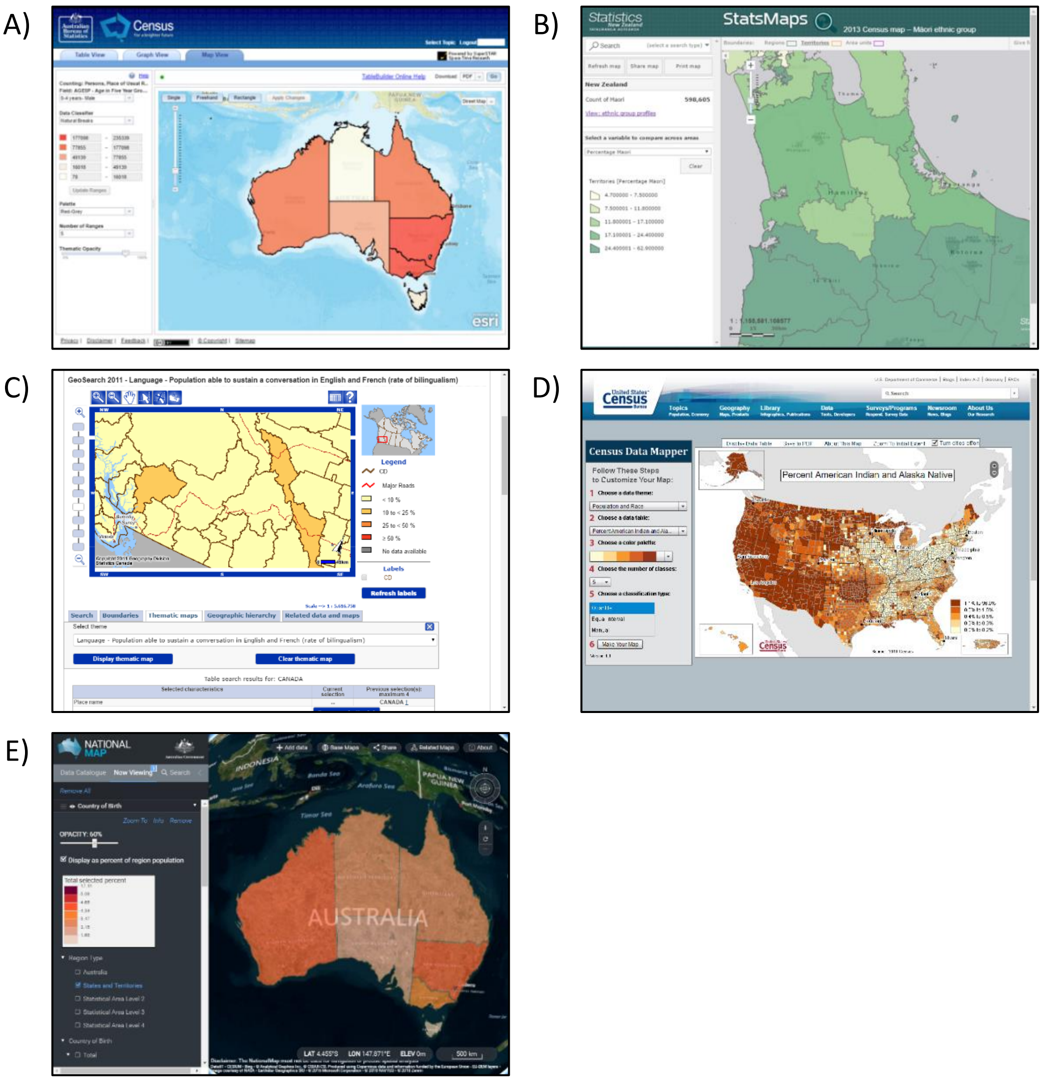

Figure 5: Screenshots of the five evaluated web applications: A) Australian Bureau of Statistics' "TableBuilder" showing a choropleth with options for customising the cartography (Australian Bureau of Statistics, n.d.), B) Statistics New Zealand's "StatsMaps" showing the percentage of Maori population by

territorial authority (Australian Bureau of Statistics, n.d.), C) Statistics Canada's "GeoSearch" showing a choropleth of the percentage of the population able to speak both English and French by Census Divisions (Statistics Canada, 2012), D) United States Census Bureau's "Census Data Mapper" showing a choropleth of the percentage of American Indian and Alaska Natives (United States Census Bureau, n.d.), E) the Australian Government's "NationalMap" showing a choropleth, in 3D mode, of the percentage of the population born outside Australia by state (Australian Government, n.d.).

3.8 Results

	Australian Bureau of Statistics' "TableBuilder"	**Statistics New Zealand's "StatsMaps"**	**Statistics Canada's "GeoSearch"**	**United States Census Bureau's "Census Data Mapper"**	**Australian Government's "NationalMap"**
Demography					
Data procurement	3/3	1/3	1/3	1/3	1/3
Total for demography	**100%**	**33%**	**33%**	**33%**	**33%**
Cartography					
Essential map components	3/3	2/3	2/3	0/3	3/3
Cartographic communication	2/3	2/3	2/3	2/3	3/3
Map interactivity	3/3	2/3	1/3	2/3	2/3
Total for cartography	**89%**	**67%**	**56%**	**44%**	**89%**
Human-computer interaction					
System functionality	2/3	2/3	1/3	2/3	2/3
User interface design	2/3	2/3	1/3	2/3	3/3
User data input	3/3	3/3	3/3	3/3	3/3
User support and error recovery	2/3	3/3	2/3	2/3	2/3
Total for HCI	**75%**	**83%**	**58%**	**75%**	**83%**
Software technology					
Client-side software	3/3	3/3	1/3	1/3	3/3
Server-side software	3/3	3/3	1/3	2/3	3/3
Total for software technology	**100%**	**100%**	**33%**	**50%**	**100%**

Table 5: Summary of evaluation results for all web applications.

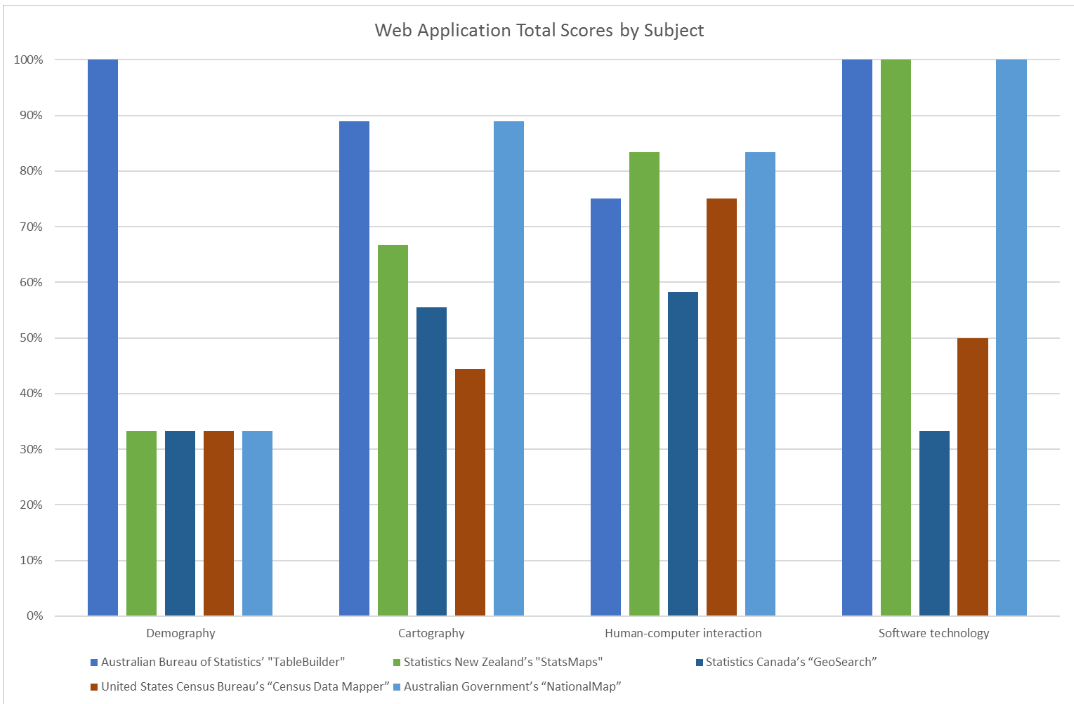

Figure 6: Bar graph of evaluation results for all web applications.

3.9 Discussion and Conclusion

This paper has attempted to answer the question: 'what factors are important for the design of web mapping applications that provide visualisations of demographic diversity?' using a high-level review of four themes which are most relevant to demographic web map development – demography, cartography, human-computer interaction and software technology. These themes and how they intersect have informed the development of a set of guidelines for evaluating demographic web mapping applications. Due to the hugely varied nature of web applications, the guidelines were kept broad so they would remain relevant across a range of paradigms. These guidelines were then applied to five different web applications, which deliver demographic data using web maps.

By reviewing demography, cartography, human-computer interaction, it is clear that there are two significant gaps. Firstly, there is a need for a web application that allows users to upload their own data and automatically map this data with a higher degree of sophistication to what is currently available. Secondly, there is a need for cartographic communication to be given a higher priority when evaluating web maps. It is hoped that this work will encourage future researchers and web mapping developers to pay more attention to the way cartography is implemented beyond just the usability of web maps. Web maps offer the benefits of user interaction and animation for information discovery but these were not harnessed by the evaluated web applications. While this paper assessed the needs for mapping population diversity, but more broadly the outcome is that it shouts "what about the *content* of the map?"

3.9.1 Demography

Because data analysis will likely need to be performed using third-party software, the tasks of data selection and data visualisation are separated and, therefore, do not necessarily need to be accomplished in the same web application. Because of this, a web application for the purpose of generating visualisations of demographic data should provide the functionality to upload customised data after it has been analysed.

From the web application evaluations, only the Australian Bureau of Statistics' "TableBuilder" was satisfactory in delivering comprehensive statistics and allowing for custom data to be imported. The other web applications scored lower against the guidelines as they only provide fixed and inflexible data selections or only cover a narrow range of selectable variables. This is probably due to these web applications aiming to provide maps to a wide audience and not necessarily targeted to demographic analysts.

3.9.2 Cartography

The principles of print cartography still mostly apply to web maps but the availability of interaction and animation, and the dynamicity of on-screen maps, bring new challenges and opportunities. Map interactivity is significant to effective knowledge transfer and should be harnessed, given this opportunity. Existing studies of web maps have focussed on the human-computer usability of maps rather than the cartographic usability and this research looked to address this by establishing a set of guidelines that take cartographic communication into account.

The choropleth map is still dominant but there were examples where choropleths were used to visualise absolute values, which can be misleading. There was a noticeable lack of interactivity in terms of being able to click on map elements for further information. Some web applications offer the ability to customise the output of the choropleths, particularly the colour palette and the number of classes and method of classification, but these options are reasonably limited.

It is disappointing, but not surprising, that more sophisticated map types are not utilised more. This is likely due to choropleths being the most flexible and easily created map type, especially for computers. The creation of arbitrary sophisticated maps with user-selected data requires manual intervention using GIS. In a web context; this crosses from web maps into web-based GIS, requiring the user to be skilled with these systems. Sophisticated maps also require input from skilled cartographers. Choropleths should not be considered a default or catch-all solution for mapping population data, but this paper recognises that it is often the only practical way to create maps automatically and dynamically.

3.9.3 Human-computer interaction

This paper reviewed human-computer interaction with a two-fold perspective. Firstly, previous studies provide lists of existing usability evaluation guidelines, underpinned by the design principles from the more fundamental HCI literature, which was adapted to produce a condensed set of guidelines, more appropriate for evaluating web mapping applications. Secondly, research into the methods of usability studies revealed a framework for evaluating and scoring web applications, which this research has adapted for its own evaluations.

In general, the reviewed web applications performed reasonably well in terms of usability. Statistics Canada's "GeoSearch" was found to be difficult to use because the interface was unintuitive, but the reasons why were not immediately obvious. It is acknowledged that this is a subjective result and more in-depth usability testing using specific guidelines would likely help to formally identify the exact issues which make it difficult to use.

In terms of system design, it is interesting to note, in hindsight, that the evaluation guidelines for demography are at odds with the HCI guidelines for task conformity. The HCI guidelines evaluate whether the web application conforms to the task it was *designed* to do, not what the demography guidelines *expect* it to do. In the former, most of the web applications scored moderately well.

3.9.4 Software technology

The review of software technology examined the difference between open source and proprietary software distribution models and what they might mean for web mapping. This was particularly focussed on licencing models and why open source software is not necessarily any better or worse than proprietary software.

Esri's software appears to be popular, with three of the five web applications using ArcGIS for Server to serve their geospatial data. In these cases, the client-side software also used either the ArcGIS API for JavaScript or the ArcGIS API for Flex. In most cases, these implementations performed well, with the exception of the United State Census Bureau's "Census Data Mapper" which was extremely limited in functionality. Statistics Canada's "GeoSearch" appears to be entirely custom-built.

Of particular interest was the discovery of the Australian Government's "NationalMap" and the open source software that it is built entirely upon. This software showed ingenuity and imagination, which was not apparent in the other web applications evaluated. This helps confirm that open source initiatives can and do produce exemplary software solutions.

3.9.5 Limitations

The evaluations were deliberately limited to a small sample of government-run web applications that provide access to national statistics and are not representative of the global situation. As this research developed it became obvious that the functions of providing data and visualising that data are fundamentally divided by the demographer's analysis needing to be performed outside the context of the web application. This is reflected by the reality that government-run web applications focus on either providing raw data or on showing selected pre-processed visualisations.

In developing the guidelines, it was difficult to find a balance between general and specific rules. Specific guidelines decrease in validity as the broadness in the range of evaluated web applications increases. The guidelines rely on intuition and it is firmly acknowledged that this intuition leads to significant subjectivity in scoring.

3.9.6 Future Research

Future studies should aim to formally define guidelines for cartographic communication using web maps, paying closer attention to cartographic grammar and convention over usability. The guidelines in this paper could serve as a starting point for developing more specific rules to reduce the level of subjectivity and need for cartographic expertise. These future rules could also be used in research to establish the automation of cartography which currently require human input to be successful. This, in turn, could lead to better web mapping tools, not just for demographers and mapping population diversity, but for analysts from all disciplines, without the need for specialised cartographic knowledge and expertise. This is moving beyond web mapping and into web-based GIS but is part of the overall movement to bring GIS and mapping to a wider community of professionals.

4 Discussion and Conclusion

The reality of web application development is that the technology is often built first and attention is initially placed on answering the question: ‘how will we get this to work?’ before asking the arguably more important question: ‘how will it communicate?’. This issue is compounded by a dearth of tools for automating the process of building sophisticated maps beyond the humble choropleth. Cartographic theory is being outpaced by GIS and web technology (Harrower & Fabrikant, 2008), but this is an opening for experimentation and trial and error, in order for new ideas to emerge. These new ideas should be tested and compared for their effectiveness at communicating geographic information in a multimedia environment.

This thesis has attempted to answer the question: ‘what factors are important for the design of web mapping applications that provide visualisations of demographic diversity?’ using a high-level review of four themes which are most relevant to demographic web map development – demography, cartography, human-computer interaction and software technology. These themes and how they intersect have informed the development of a set of guidelines for evaluating demographic web mapping applications. Due to the hugely varied nature of web applications, the guidelines were kept broad so they would remain relevant across a range of paradigms. These guidelines were then applied to five different web applications which deliver demographic data using web maps.

4.1 Demography

A review of the discipline of demography and how population diversity is quantitatively measured and compared was undertaken. Through this, it has been determined who demographers are and what they want. In particular, it is noted that demographers want to stay close to the data in order to perform their own analyses (Van Dalen & Henkens, 2012). There can be no one-size-fits-all strategy for accessing data, and basic off-the-shelf or pre-processed data will not always be adequate for direct rendering to a web map for most demographer’s needs without undertaking further custom analysis. It is not necessarily the job of the agency providing the data to also provide this pre-processed data as the generation of this is the role of demographers, and they are not necessarily employed by those agencies. Analysing and processing data and turning this into useful information is what demographers do, and their research questions can be highly varied and unique. This results in a split between the tasks data selection and data visualisation and, therefore, these functions do not necessarily need to be accomplished in the same web application. Because of this, a web application for the purpose of generating visualisations of demographic data should provide the functionality to upload customised data after it has been analysed using third-party software such as Microsoft Excel.

From the web application evaluations, only the Australian Bureau of Statistics' "TableBuilder" was satisfactory in delivering comprehensive statistics and allowing for custom data to be imported. The other web applications were scored lower against the guidelines as they only provide fixed and inflexible data selections or only cover a narrow range of selectable variables. This is probably due to these web applications aiming to provide maps to a wide audience and not necessarily targeted to demographic analysts.

4.2 Cartography

This thesis established the essential map elements and, in particular, how maps communicate. The most common thematic map types and symbologies were discussed. The principles of print cartography still apply to web maps but the availability of interaction and animation, and the dynamicity of on-screen maps, bring new challenges and opportunities. Map interactivity is significant to effective knowledge transfer and should be harnessed, given this opportunity (Buchroithner & Gartner, 2013; Cartwright & Peterson, 2007). Existing studies of web maps have focussed on the human-computer usability of maps rather than the cartographic usability and this thesis looked to address this by establishing a set of guidelines that look at map usability and cartographic communication with more balance.

Findings from the web application evaluations show that the choropleth map is still dominant. There were examples where choropleths were used to visualise absolute values, which can be misleading. Most web applications evaluated showed an appropriate level of topography in their basemaps and made good use of panning and zooming. However, there was a noticeable lack of interactivity in terms of being able to click on map elements for further information. Some web applications offer the ability to customise the output of the choropleths, particularly the colour palette and the number of classes and method of classification, but these options are reasonably limited.

It is disappointing, but not surprising, that more sophisticated map types, such as dot density maps and statistical surfaces, are not utilised more, particularity where the user is also the map producer. This is likely due to choropleths being the most flexible and easily created map type, especially for computers. The creation of arbitrary sophisticated maps with user-selected data requires manual intervention using GIS. In a web context, this crosses from web maps into web-based GIS, requiring the user to be skilled with these systems. Sophisticated maps also require input from skilled cartographers who are able to make informed decisions about map types, symbologies and colour. As simple as choropleths may seem, they have their nuances and pitfalls, particularly regarding spatially extensive data, and this may also require input from cartographers.

Choropleths should not be considered a default or catch-all solution for mapping population data, but this thesis recognises that it is often the only practical way to create maps automatically and dynamically, based on user input. It is also recognised that the web applications evaluated are not meant to be fully-fledged geographic information systems and that the functionality they provide will naturally be limited.

4.3 Human-computer interaction

This thesis reviewed human-computer interaction with a two-fold perspective. Firstly, previous studies provide lists of existing usability evaluation guidelines, both broad and specific, which were produced for evaluating software and web applications in general, with one study focusing on the usability of web-based GIS (Komarkova et al., 2007). These guidelines are underpinned by the design principles from the more fundamental HCI literature, which is concerned with the sciences behind the human factors required for software to achieve its goals. The literature was adapted to produce a condensed and more specific set of guidelines, appropriate for evaluating the usability of web applications that map demographic information. Secondly, research into the methods of usability studies revealed a framework for evaluating and scoring web applications, which this thesis has adapted for its own evaluations.

In general, the reviewed web applications performed reasonably well, and equally, in terms of usability. However, Statistics Canada's "GeoSearch" was found to be difficult to use because the interface was unintuitive, but the reasons why were not immediately obvious. It is acknowledged that this is a subjective result and more in-depth usability testing using specific guidelines would likely help to formally identify the exact issues which make it difficult to use. This gap highlights that the guidelines in this thesis are useful for identifying whether a major issue exists or not, but are less useful in determining root causes.

In terms of system design, it is interesting to note, in hindsight, that the evaluation guidelines for demography are at odds with the HCI guidelines for task conformity. For demography, this thesis evaluates whether a web application is suitable for procuring and visualising demographic data, but most of the web applications evaluated scored poorly in this area. However, the HCI guidelines evaluate whether the web application conforms to the task it was *designed* to do, not what the demography guidelines *expect* it to do. In the former, most of the web applications scored moderately well.

4.4 Software technology

The review of software technology examined the difference between open source and proprietary software distribution models and what they might mean for web mapping. This was particularly focussed on licencing models and why open source software is not necessarily any better or worse than proprietary software. Esri, a supplier of proprietary GIS and web mapping software, advocates for an approach that selects technology from a variety of sources as suited to the specific needs of the implementation (Esri, 2011).

It is difficult to evaluate a web application based on the client and server software technology chosen, particularly as an end user. Firstly, there is not necessarily a right or wrong answer and it is sometimes impossible to determine what software is being used as an outsider looking in if the software does not publicly identify itself. Secondly, it was not necessary to review the fitness for purpose of the software as this was already addressed by the demography and human-computer interaction guidelines. Therefore, the guidelines for evaluating software technology were developed from more of a ‘general discussion’ point of view with emphasis on apparent performance. This has meant that the software technology evaluations of this thesis are more subjective than the other themes.

Esri’s software appears to be popular, with three of the five web applications using ArcGIS for Server to serve their geospatial data. In these cases, the client-side software also used either the ArcGIS API for JavaScript or the ArcGIS API for Flex. In most cases, these implementations performed well, with the exception of the United State Census Bureau’s “Census Data Mapper” which was extremely limited in functionality. Statistics Canada’s “GeoSearch” appears to be entirely custom-built and this does not appear to have served them well, particularly because the software was difficult to use (see HCI) and had poor response times. However, it is acknowledged that response time may be due to a number of factors beyond the software itself.

Of particular interest was the discovery of the Australian Government's "NationalMap" and the open source software that it is built entirely upon. This software showed ingenuity and imagination, which was not apparent in the other web applications evaluated. This helps confirm that open source initiatives can and do produce exemplary software solutions.

4.5 Limitations

It was not possible to comprehensively review the four themes individually in the scope of this thesis. This was a comparative exercise that only touched the surface of each discipline without going into the intricacies.

The evaluated web applications were limited to government-run projects that provide access to national statistics and the sample size is small. While this was a deliberate choice, it has excluded the web mapping work by other organisations or individuals and the results are not representative of the global situation. In particular, the initial choice of web applications to evaluate was driven by an assumption that providers of national statistics would also provide analysis and visualisations of those statistics. However, as this thesis developed it became obvious that the functions of providing data and visualising that data are fundamentally divided by the demographer's analysis needing to be performed outside the context of the web application. This is reflected by the reality that government-run web applications focus on either providing raw data or on showing selected pre-processed visualisations.

In developing the guidelines, it was difficult to find a balance between general and specific rules. Specific guidelines decrease in validity as the broadness in the range of evaluated web applications increases. As this study focussed on government statistics web applications, its range was limited, but widening this range will have likely led to issues in consistently applying the guidelines. Broad guidelines are useful for rapidly evaluating a system but do require a certain amount of expertise from the evaluator on the theme being evaluated. As mentioned above, the guidelines are useful at rapidly determining if there is a problem but do not necessarily pinpoint the root cause of the issue. The guidelines rely on intuition and it is firmly acknowledged that this intuition leads to significant subjectivity in scoring.

4.6 Contribution to knowledge

This thesis considered and assembled a wide range of subject matter to make the most important information available to those either using or creating web maps for visualising demographic diversity. Web maps are, by nature, an intersection of different disciplines and some understanding each of these disciplines adds value to the design process and eventual product.

By reviewing demography, cartography, human-computer interaction, it is clear that there are two significant gaps. Firstly, there is a need for a web application which allows users to upload their own data and automatically map this data with a higher degree of sophistication to what is currently available. An idea for a prototype is made below.

Secondly, there is a need for cartographic communication to be given a higher priority when evaluating web maps. It is hoped that this thesis will encourage future studies and web mapping developers to pay more attention to the way cartography is implemented beyond just the usability of web maps. Web maps offer the benefits of user interaction and animation for information discovery but these were not harnessed by the evaluated web applications. While this thesis assessed the needs for mapping population diversity, but more broadly the outcome is that it shouts “what about the *content* of the map?”

4.7 Future Research

Future studies should aim to formally define guidelines for cartographic communication using web maps, paying closer attention to cartographic grammar and convention over usability. The guidelines in this thesis could serve as a starting point for developing more specific rules to reduce the level of subjectivity and need for cartographic expertise. These future rules could also be used in research to establish the automation of cartography, which currently requires human input to be successful. This, in turn, could lead to better web mapping tools, not just for demographers and mapping population diversity, but for analysts from all disciplines, without the need for specialised cartographic knowledge and expertise. This is moving beyond web mapping and into web-based GIS but is part of the overall movement to bring GIS and mapping to a wider community of professionals.

4.7.1 The shape of a potential web application for visualising demographic diversity

Future research could also take the form of developing a prototype web application that fulfils the requirements of demographers seeking to create visualisations of demographic diversity and performs well against the guidelines of this thesis. Specifically, the ability for users to upload data from their own analysis, such as the functionality provided by the Australian Bureau of Statistics' "TableBuilder" web application (Australian Bureau of Statistics, n.d.), and the use of more sophisticated and automated map types.

The idea of a web application which creates maps from user-provided data offers a balance between the needs of demographers to prepare their own data, using tools like Microsoft Excel, and their need to create visualisations from this data, without needing skills in GIS or cartography. It removes the need for the user to understand and manage the geometric spatial data but provides a simple way for them to apply their own numeric data to these geometries.

An example of a sophisticated map, beyond choropleths, could be a smoothed statistical surface to help overcome some of the spatial issues associated with aggregated data, such as that created by Luebbering et al. (2013). This provides a balance between data accuracy and the map's ability to communicate, and could be represented as an isopleth or as a three-dimensional model which could be panned and zoomed, thus benefiting from interactivity. The map's communication could be extended by animations that smoothly change the shape of the model to portray temporal change. Further, the use of a smoothed surface allows the effects of changing boundaries, such as updated meshblock borders in the New Zealand Census, to be elegantly incorporated into the visualisation. However, a problem of using three-dimensional maps for representing data is that they can cause confusion between terrain and phenomenon; the phenomenon's peaks and troughs could be misinterpreted as physical terrain, and so the cartography would need to steer the reader towards the theme rather than the topography. Using colours from nature, such as greens and blues, might cause the reader to interpret the theme as topography, while using less "natural" colours, such as reds and purples, would convey a more thematic objective. Furthermore, the underlying basemap would be obscured by the three-dimensional model, which diverges from the guideline on spatially orienting the user. A transparent theme might assist with this, along with place name labels attached to vertical lines, or even the original aggregation boundaries draped over the theme.

Further sophistication could incorporate a dasymetric mapping approach to further reduce the impact of aggregated population data, where the shapes of the choropleth polygons would be modified to better represent the distribution of the phenomenon being mapped. In order to do this, extra spatial data would be required, such as the boundaries of urban areas, to enable the aggregated boundaries to be remodelled to be more representative of the true geography of populations. For demographers' purposes, this might be considered a fundamental dataset for enabling the visualisation of spatio-temporal data, such as changing diversities. If a template for remodelled boundaries was available, data from custom analyses could be applied to this template in order to create sophisticated maps which represent the phenomenon more realistically and with more meaning to the map reader, even though the data has been abstracted away from the original aggregation units.

Bibliography

Arduino. (n.d.). Retrieved 17 March 2016, from https://www.arduino.cc/

Australian Bureau of Statistics. (n.d.). TableBuilder. Retrieved 20 June 2016, from http://www.abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder

Australian Government. (n.d.). NationalMap. Retrieved 20 June 2016, from https://nationalmap.gov.au/

Bakens, J., Nijkamp, P., & Poot, J. (2015). E pluribus prosperitas: on cultural diversity and economic development. In J. Bakens, P. Nijkamp, & J. Poot (Eds.), *The economics of cultural diversity* (pp. 1–13). Cheltenham, United Kingdom: Edward Elgar. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=11095044

Buchroithner, M. F., & Gartner, G. (2013). The New Face of Cartography. *GIM International*, *27*(6), 22–27.

Buckley, A. (2013). Understanding statistical data for mapping purposes. Retrieved 19 June 2016, from http://www.esri.com/esri-news/arcuser/winter-2013/understanding-statistical-data-for-mapping-purposes

Cammack, R. G. (2007). Cartographic approaches to web mapping services. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 441–453). Berlin: Springer.

Cartwright, W., & Peterson, M. P. (2007). Multimedia cartography. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 1–10). Berlin, Germany: Springer.

Cecconi, A., & Galanda, M. (2002). Adaptive zooming in web cartography. *Computer Graphics Forum*, *21*(4), 787–799.

Defense Distributed. (n.d.). https://defdist.org/. Retrieved 17 March 2016, from https://defdist.org/

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2005). *Human-computer interaction* (3rd ed.). New York, NY: Pearson/Prentice-Hall.

Esri. (2011). Open source technology and Esri. Retrieved 17 March 2016, from http://www.esri.com/news/arcnews/spring11articles/open-source-technology-and-esri.html

Fitzgerald, B., Kesan, J. P., & Russo, B. (2011). *Adopting open source software: A practical guide*. Cambridge, MA: MIT Press. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10509225

Gaffuri, J. (2011). Improving Web Mapping with Generalization. *Cartographica*, *46*(2), 83–91.

gnu.org. (n.d.). Retrieved 17 March 2016, from https://www.gnu.org/philosophy/free-sw.html

Haining, R. (2003). *Spatial data analysis: theory and practice*. West Nyack, United Kingdom: Cambridge University Press. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10062302

Harrower, M., & Fabrikant, S. (2008). The role of map animation for geographic visualisation. In M. Dodge, M. McDerby, & M. Turner (Eds.), *Geographic visualisation: Concepts, tools and applications*. Hoboken, NJ: Wiley.

Holdsworth, C. (2013). *Population and society*. Los Angeles, [Calif.]: Sage Publications.

Howells, L. (2011, December 16). A Guide To Heuristic Website Reviews. Retrieved 24 April 2016, from https://www.smashingmagazine.com/2011/12/a-guide-to-heuristic-website-reviews/

IDEA. (2011, July 22). Open source vs. proprietary software. Retrieved 17 March 2016, from http://www.idea.org/blog/2011/07/22/open-source-vs-proprietary-software/

Jackson, N. (2007). Population ageing in a nutshell: A phenomenon in four dimensions. *People and Place*, *15*(2), 12.

Jobst, M. (2007). Standards and open source for cartographic multimedia applications. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 415–426). Berlin, Germany: Springer.

Komarkova, J., Jedlicka, M., & Hub, M. (2010). Usability user testing of selected web-based GIS applications. *Technology*, *24*, 27.

Komarkova, J., Sedlak, P., Novak, M., Musilova, A., & Slavikova, V. (2011). Methods of usability evaluation of Web-based geographic information systems. *International Journal of Systems Applications, Engineering & Development*, *5*(1), 33–41.

Komarkova, J., Visek, O., & Novak, M. (2007). Heuristic evaluation of usability of geoweb sites. In J. M. Ware & G. E. Taylor (Eds.), *Web and wireless geographical information systems* (pp. 264–278). Berlin, Germany: Springer. Retrieved from http://link.springer.com.nz/chapter/10.1007/978-3-540-76925-5_20

Kraak, M.-J. (2007). Cartography and the use of animation. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 317–326). Berlin, Germany: Springer.

Kraak, M.-J., & Ormeling, F. (2011). *Cartography: visualization of spatial data* (3rd ed.). New York, NY: Guilford Press.

Luebbering, C. R., Kolivras, K. N., & Prisley, S. P. (2013). Visualizing linguistic diversity through cartography and GIS. *Professional Geographer*, *65*(4), 580–593. http://doi.org/10.1080/00330124.2013.825517

Lundquist, J. H., Anderton, D. L., & Yaukey, D. (2015). *Demography: the study of human population* (4th ed.). Long Grove, IL: Waveland Press.

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design*. New York, NY: Guilford Press.

MacEachren, A. M., Crawford, S., Akella, M., & Lengerich, G. (2008). Design and implementation of a model, web-based, GIS-enabled cancer atlas. *Cartographic Journal*, *45*(4), 246–260. http://doi.org/10.1179/174327708X347755

MacKenzie, I. S. (2013). *Human-computer interaction: an empirical research perspective*. Amsterdam, Netherlands: Morgan Kaufmann.

Meirelles, I. (2013). *Design for information: an introduction to the histories, theories, and best practices behind effective information visualizations*. Beverly, MA: Rockport Publishers.

Miller, S. (2007). Design of multimedia mapping products. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 89–104). Berlin, Germany: Springer.

Nielsen, J. (1995, January 1). 10 heuristics for user interface design. Retrieved 22 April 2016, from https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J., & Mack, R. L. (1994). *Usability inspection methods*. New York: Wiley.

Nijkamp, P., & Poot, J. (2015). Cultural diversity: a matter of measurement. In J. Bakens, P. Nijkamp, & J. Poot (Eds.), *The economics of cultural diversity* (pp. 17–51). Cheltenham, UK: Edward Elgar. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=11095044

Open Source Initiative. (n.d.). Open Source Initiative. Retrieved 17 March 2016, from https://opensource.org/

Ormeling, F. (2007). Map concepts in multimedia products. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 105–115). Berlin, Germany: Springer.

OSVehicle. (n.d.). Retrieved 17 March 2016, from https://www.osvehicle.com/

Param, S. (2013, August 7). Proprietary vs. open software - the pros and cons explored. Retrieved 17 March 2016, from http://www.3pillarglobal.com/insights/proprietary-vs-open-software-the-pros-and-cons-explored

Peterson, G. N. (2015). *GIS cartography: a guide to effective map design* (2nd ed.). Boca Raton, FL: CRC Press.

Peterson, M. P. (1995). *Interactive and animated cartography*. Englewood Cliffs, NJ: Prentice Hall.

Peterson, M. P. (2007a). Elements of multimedia cartography. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 63–73). Berlin, Germany: Springer.

Peterson, M. P. (2007b). The internet and multimedia cartography. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 35–50). Berlin, Germany: Springer.

Robinson, A. H. (1995). *Elements of cartography*. (J. L. Morrison, P. C. Muehrcke, A. J. Kimerling, & S. C. Guptill, Eds.) (6th ed.). New York, NY: Wiley.

Statistics Canada. (2012). GeoSearch. Retrieved 20 June 2016, from http://geodepot.statcan.gc.ca/GeoSearch2011-GeoRecherche2011/GeoSearch2011-GeoRecherche2011.jsp?lang=E&otherLang=F

Statistics New Zealand. (2015, October 29). Privacy, security, and confidentiality of information supplied to Statistics NZ. Retrieved 26 June 2016, from http://www.stats.govt.nz/about_us/who-we-are/policies-and-protocols/confidentiality-of-info-supplied-to-snz/safeguarding-confidentiality.aspx

Synopsys. (2013). Coverity Scan report finds open source software quality outpaces proprietary code for the first time. Retrieved 17 March 2016, from http://www.coverity.com/press-releases/coverity-scan-report-finds-open-source-software-quality-outpaces-proprietary-code-for-the-first-time/

Taylor, D. R. F., & Lauriault, T. P. (2007). Future directions for multimedia cartography. In W. Cartwright, M. P. Peterson, & G. Gartner (Eds.), *Multimedia cartography* (pp. 505–522). Berlin, Germany: Springer.

Travis, D. (2016, April 12). 247 web usability guidelines. Retrieved 13 April 2016, from http://www.userfocus.co.uk/resources/guidelines.html

United States Census Bureau. (n.d.). Census Data Mapper. Retrieved 20 June 2016, from http://www.census.gov/geo/maps-data/maps/datamapper.html

Van Dalen, H. P., & Henkens, K. (2012). What is on a demographer's mind? A worldwide survey. *Demographic Research*, *26*(16), 363–408. http://doi.org/10.4054/DemRes.2012.26.16

Weinstein, J. A., & Pillai, V. K. (2016). *Demography: the science of population* (2nd ed.). Lanham, MD: Rowman & Littlefield.

Appendix: Full results with comments

The results tables in this appendix contain the comments given for each factor. The summarised scores are available in Table 5 and as a bar graph in Figure 6.

Australian Bureau of Statistics' "TableBuilder"

Demography

Factor	Comments	Score
Data procurement	Databases are comprehensive and cover multiple statistics. Many spatial variables are available, but the data is only for the 2011 census. Custom data can also be imported, but this feature was not tested. Data can be selected with an impressive level of flexibility and complexity. Tables can be exported in several formats, including for Microsoft Excel, once they have been compiled. Data can be represented as absolute values or percentages. An impressive table builder. While the primary purpose of the system is to compile data tables, a good graph builder is also available along with a map builder if the data contains a spatial variable. There are warnings that cells with small values are unreliable and that some cells are randomly adjusted to protect privacy.	3/3
	Total Score	3/3

Cartography

Factor	Comments	Score
Essential map components	State boundaries are always displayed. Basemap topography contains natural and built features, scales well and has varying levels of detail. The legend forms an integral part of the controls for customising the map. A single web map is available if a spatial variable is selected. Regions can be included and excluded from the table and graph via the web map, but only once it is available.	3/3
Cartographic communication	Choropleth map type. Defaults are suitable, but some fine-tuning is necessary. Colour palettes are ranges of lightness and ranges of greyness, which is appropriate for the data I had selected. The theme's opacity is adjustable to establish a balance with the topography. Good selection of classification methods and the number of classes. Difficult to build an inappropriate map. No alternatives to choropleth map type. There is no mechanism for comparing change over time.	2/3
Map interactivity	Clicking regions of the map shows exact figures in a popup. The user can select and deselect regions using the web map, which also updates the data selection.	3/3
	Total score for Cartography	8/9

Human-computer interaction

Factor	Comments	Score
System functionality	The system's name is "TableBuilder" and it performs this task very well. A real-world workflow always starts with selecting raw data, which can then be used to create visualisations and this process is clearly followed. User registration requires name, email, address and telephone numbers. Some broken functionality, e.g. exporting a map in PDF format.	2/3
User interface design	The interface is built for complex functionality over looks. Some table controls are small, but this is to keep table dimensions small. No unusual layouts, controls or elements. Navigation is ordered by process of table-graph-map. Database and table selection has too few items per page, has no search function and requires a redundant mouse click to accept the selection although this is a minor detail. Some font sizes are too small, and some errors and warnings are not highlighted. User actions are acknowledged by loading indicators.	2/3
User data input	No issues noted.	3/3
User support and error recovery	A comprehensive online manual is available with one click. Some context-based help links in critical places. Data fields have links to a census dictionary for further explanations. No undo or redo functionality. Session	2/3

	timeout is too short and work in progress is not saved. Error messages are not visible enough.	
	Total Score	9/12

Software technology

Factor	**Comments**	**Score**
Client-side software	Mapping support is provided by version 2.3 (current is version 4) of Esri's ArcGIS API for JavaScript, which is proprietary but freely available and designed to work with ArcGIS for Server. Basemaps and polygons are provided pre-rendered by the server while the data elements are rendered in the browser, allowing theme changes to occur instantly. Being a JavaScript library, it does not require any other software to be installed by the user. CPU and RAM usage profiles are perfectly acceptable.	3/3
Server-side software	The website itself appears to be built using JavaServer Pages (not to be confused with JavaScript). Mapping support is provided by ArcGIS for Server (version not identifiable) which is proprietary and paid software. The server responds in less than one second, and there is evidence of browser caching being used.	3/3
	Total Score	6/6

Statistics New Zealand's "StatsMaps"

Demography

Factor	Comments	Score
Data procurement	Although there are many pre-built themes, the raw data is not accessible using StatsMaps, limiting its usefulness by serious demographers. Some maps have a slider to change between the 2001, 2006 and 2013 censuses. However, the automated choropleth classification ranges change between years making the maps incomparable with each other. Maps showing ethnicity do not address ethnic over-count, despite this being a characteristic of the ethnicity data from Statistics New Zealand. No other data visualisations are available.	1/3
	Total Score	1/3

Cartography

Factor	Comments	Score
Essential map components	Area unit boundaries are displayed depending on the zoom level. Basemap topography contains natural and built features, scales well and has varying levels of detail. The legend is obscure and uses too many decimal places. Areal borders feel too thick, though this may be intended to indicate that the areas are clickable. Cadastral information is shown at larger scales, and the basemap may also be switched to satellite imagery.	2/3
Cartographic communication	Choropleth map type with one colour palette. Colours change in lightness with a mild change in hue, but it is still appropriate. No options to customise the map's cartography.	2/3
Map interactivity	Clicking a map refines data to just that area. A map becomes thematic once a variable has been selected. The map can be panned using the mouse and zoomed in by double clicking. Zooming out requires using the zoom tool on the interface. I expected the scroll wheel to control zoom but, unusually, the mouse's scroll wheel pans the map vertically, and mouses equipped with a horizontal scroll wheel are handled as well.	2/3
	Total Score	6/9

Human-computer interaction

Factor	Comments	Score
System functionality	No user registration required. Each data theme has its own page, and there are many of these. Some broken functionality, particularly the sharing functions.	2/3
User interface design	The interface is simple to use due to minimal functionality. The tools to generate a choropleth map are not immediately obvious if the number of variables pushes this off the screen. The left sidebar contains all major controls but it is crowded and often requires scrolling; in particular, this obscures the legend. The page footer is unnecessarily tall, and this wastes valuable map space.	2/3
User data input	No issues noted.	3/3
User support and error recovery	An online manual is available with one click. Undo and redo commands are not present but are unnecessary due to the simplicity of the interface. Error messages, when they occur, are simple to understand.	3/3
	Total Score	10/12

Software technology

Factor	Comments	Score
Client-side software	Mapping support is provided by version 3.6 (current is version 4) of Esri's ArcGIS API for JavaScript, which is proprietary but freely available and designed to work with ArcGIS for Server. Being a JavaScript library, it does not require any other software to be installed by the user. CPU and RAM usage profiles are perfectly acceptable.	3/3
Server-side software	The website itself appears to be built using ASP.NET, while mapping support is provided by ArcGIS for Server (version not identifiable) which is proprietary and paid software. There do not appear to be any issues with performance.	3/3
	Total Score	6/6

Statistics Canada's "GeoSearch"

Demography

Factor	**Comments**	**Score**
Data procurement	There are only limited datasets available for creating thematic maps, mostly covering variables about spoken languages. Like the rest of Statistics Canada's website, there are excellent and very clear footnotes explaining the anomalies and features of the data. However, there are no other visualisations aside from the simple choropleth maps.	1/3
	Total Score	1/3

Cartography

Factor	Comments	Score
Essential map components	Topographic information includes coastlines, rivers, lakes and major roads and has multiple levels of detail based on zoom. The projection used is not specifically identified, but it is appropriate for a region approaching the North Pole. The legend is good and includes both topographic and thematic layers. Maps are used for navigating to data for an area and for describing the location of an area. An inset map is also provided which shows the location of an area within Canada. The main web map is also used for displaying thematic data.	2/3
Cartographic communication	Choropleth maps with adequate colour palette in four classes. No customisation of the thematic cartography is available.	2/3
Map interactivity	The map cannot be panned and zoomed using a mouse unless those specific controls are selected, making the map difficult to navigate. Changing the extent of the map causes the entire page to reload rather than loading just the map data in the background.	1/3
	Total Score	5/9

Human-computer interaction

Factor	Comments	Score
System functionality	The application seems confused about what its primary task is. The functionality to select an area, select a theme and view tabular data is not cohesive. There are frequent errors, timeouts and inconsistencies with the way the system behaves. While the page as a whole can be printed, maps and data cannot be exported.	1/3
User interface design	Controls are spread about the interface but are of a sufficient size so as to not break Fitts' Law. The overall design is unintuitive, and I spent longer than I ought to have in establishing how it works.	1/3
User data input	No issues noted.	3/3
User support and error recovery	Comprehensive help is available. There are no undo or redo actions as such, but up to four previously selected areas can be easily restored to the map.	2/3
	Total Score	7/12

Software technology

Factor	Comments	Score
Client-side software	Appears to be custom built using plain HTML and JavaScript. The web map is rendered as a GIF image that is generated by the server and loaded directly into an HTML image tag. Tools to pan and zoom simply reload the entire page.	1/3
Server-side software	Appears to be custom built using JavaServer Pages. I cannot identify any particular software used. The server is very slow to respond, and I encountered critical errors (HTTP 500) over several days of using the application.	1/3
	Total Score	2/6

United States Census Bureau’s “Census Data Mapper”

Demography

Factor	Comments	Score
Data procurement	The data is only available in three pre-built themes: “Age and Sex”, “Population and Race” and “Family and Housing”. Within these themes, there are only limited options, such as percent male or female or percent aged over 65. This is not suitable for serious demographic analysis. There is no further information available about the data directly through this application. Limited to the 2010 Census.	1/3
	Total Score	1/3

Cartography

Factor	Comments	Score
Essential map components	No topographic information is present at all. Although county boundaries are present, it is impossible to relate counties to the physical world.	0/3
Cartographic communication	Choropleth map type with appropriate colour palettes. Within the confines of a choropleth, the cartography can be customised with a variety of appropriate colour palettes. The number of classes is selectable from three to five and can use quantile, equal interval or manual classification methods.	2/3
Map interactivity	The map can be zoomed using scroll wheel but, unfortunately, this also scrolls the entire page in the web browser. Hovering over areas shows the name and the exact value of the data for that area. Data tables can be shown as a popup, containing a tabular version of what the map is displaying. Clicking a row in the table moves the map to that area but the area is not highlighted for identification when this happens.	2/3
	Total Score	4/9

Human-computer interaction

Factor	**Comments**	**Score**
System functionality	In order to search for an area by name, the user must first open the table viewer, which is unintuitive. There is functionality to export maps in PDF format, but the resulting print map includes zoom buttons and is low resolution. However, this product is only designed to be a simple tool and it does perform this task. No user registration is required.	2/3
User interface design	The interface does not scale with the browser window, leaving a very small map. The interface is simple, uncluttered and mostly easy to use.	2/3
User data input	No issues noted.	3/3
User support and error recovery	Very limited documentation is available by clicking "About This Map". However, for the simplicity of the interface, this is probably acceptable.	2/3
	Total Score	9/12

Software technology

Factor	Comments	Score
Client-side software	This web application is based on Adobe Flash, meaning I am unable to identify any particular software products used. The user must be able to install Adobe Flash, which may be problematic in some circumstances. Because this client software communicates with ArcGIS for Server, it may use the ArcGIS API for Flex, but this minor detail would require confirmation from the Bureau.	1/3
Server-side software	The geospatial component of the server stack is ArcGIS for Server version 10.31 which is paid proprietary software. Because the client software is based on Adobe Flash, the server returns data in AMF format, which contains exact drawing instructions and takes several seconds to load each map.	2/3
	Total Score	3/6

Australian Government's "NationalMap"

Demography

Factor	**Comments**	**Score**
Data procurement	There is a large variety of datasets available from multiple government agencies, some of which is available for multiple time periods. However, none of the data is available in its raw or tabular form using this application, making it unsuitable for serious demographers. No other visualisations or interpretations are available, but this is not the purpose of this application.	1/3
	Total Score	1/3

Cartography

Factor	Comments	Score
Essential map components	Several basemaps are available, including topographic maps and satellite imagery. Of particular note is the ability to view the map in 3D with the thematic layers draped over the terrain. Data is presented in thematic layers over a topographic basemap. Multiple layers can be shown at the same time and clicking a location shows a popup of information for all layers at that location.	3/3
Cartographic communication	There are a variety of map types in use, including choropleths, heat maps and point symbols, depending on the data selected. The map type and symbology is pre-configured based on the type of data. There are many datasets available in the catalogue and the few I tried all had appropriate map types and symbologies. Aside from changing the opacity of each layer and changing the order of the layers, there are no options for changing map types or symbologies.	3/3
Map interactivity	The map can be panned and zoomed intuitively using a mouse. Clicking on a location shows exact figures from all layers at that location. It is not immediately obvious that the map can be clicked on as there are no user-notification stimuli.	2/3
	Total Score	8/9

Human-computer interaction

Factor	Comments	Score
System functionality	The application performs its tasks well and is easy and intuitive to use. There are no obvious bugs or errors, and no user registration is required. What lets this application down is the lack of ability to print or export maps.	2/3
User interface design	The interface is clean, modern and visually pleasing. The contrast is light on dark, but this does not affect legibility. The interface has just enough tools to complete the tasks it is designed to do and is uncluttered and well organised.	3/3
User data input	No issues noted.	3/3
User support and error recovery	Comprehensive help is available, and the interface design is simple enough to not warrant undo and redo functionality.	2/3
	Total Score	10/12

Software technology

Factor	Comments	Score
Client-side software	Built entirely from free and open source components, particularly TerriaJS. 3D basemaps are produced using Cesium, and thematic maps are rendered using Leaflet, both also free and open source. Manipulating a map in 3D was computationally expensive and was slow on my hardware, but this is otherwise a modern and robust client.	3/3
Server-side software	As the TerriaJS client software is designed to access data directly from various agencies using open architectures, there is little need for an active server stack beyond the serving of static assets. However, TerriaJS has a small server component for proxied access to data which the client software cannot access directly, such as if authentication is required.	3/3
	Total Score	6/6