



**University of
Zurich**^{UZH}

Department of Geography

Effectiveness of interactive data visualizations for
the public: a case study of refugee data

Author

Gordon Bühler

09-729-229

MSc thesis GEO 511

Department of Geography, University of Zurich

Geographic Information Visualization and Analysis, GIScience Center

Supervisors

Dr. Arzu Çöltekin

Simon Sontowski

Faculty representative

Prof. Dr. Sara Irina Fabrikant

Date of submission: 31 January 2016

Abstract

The scientific field of cartography has a research gap: interactive maps are not well understood. Particularly, the effect interactive maps have on their users' performance. The aim of this thesis was to contribute to filling this research gap by testing an interactive map against a static version. For this purpose, two versions of a digital map were created: an interactive and a static version. These two maps were then tested against each other in a between-subject study design to test two hypotheses: 1) that a static maps performs better than an interactive map regarding usability, and 2) that an interactive map engages users more to explore the data. In the study, neither of the hypotheses could be confirmed. The visualizations themselves however are a main output of this thesis that have the potential to be further developed.

Contents

1	Introduction	4
2	Research questions	5
3	Literature review	6
3.1	Map design	6
3.1.1	Map design for the internet	6
3.1.2	Visualization of movement data	6
3.2	Interactive maps	8
3.3	Visualizations for the public	10
3.4	Measuring engagement	11
3.5	Migration on the Mediterranean Sea	12
4	Gallery of visualizations	14
4.1	Static visualizations	14
4.1.1	Charles Joseph Minard’s flow maps	14
4.1.2	Migration	14
4.2	Interactive visualizations	20
5	Methods	21
5.1	Visualizations	21
5.1.1	Basic map	23
5.1.2	Interactive elements	30
5.1.3	Static version	33
5.2	User study	35
5.2.1	Getting set up	35
5.2.2	Pre-questionnaire	35
5.2.3	Main questionnaire	36
5.2.4	Post-questionnaire	36
5.2.5	Recruitment of participants	36
5.3	Experimental design	37
5.3.1	Between-subject design	37
5.3.2	Research question 1: usability	38

5.3.3	Research question 2: engagement	40
5.3.4	Other variables	42
5.3.5	Statistics	42
6	Results	43
6.1	Research question 1: usability	46
6.1.1	Accuracy	46
6.1.2	Response time	46
6.1.3	Confidence and ease	47
6.1.4	System usability scale	47
6.2	Research question 2: engagement	49
6.2.1	Exploration	50
7	Discussion	51
7.1	Research question 1: usability	51
7.2	Research question 2: engagement	52
7.3	Limitations and outlook	54
7.3.1	User study	54
7.3.2	Visualization	54
8	Conclusion	56
	Bibliography	57
A	D3 code for the interactive visualization	62
B	Questionnaire of the online survey	83

1 Introduction

New tools are allowing journalists to go new ways. The increased availability of data, along with new tools to help journalists process and visualize them, has created a surge of visualizations in online media. This exciting development has been well received by the users of online media. Within these unsteady times for journalism, where sales are dropping, data journalism promises to have a future.

One specific type of digital visualization is the interactive map. As are other products of data driven journalism, interactive maps are becoming more frequent. At the same time, research is lagging behind. The effects of interactivity on map use are not well understood. The aim of this thesis is to investigate interactive maps scientifically. How does the element of interactivity change the way a map is perceived and its information content is processed? This thesis seeks to contribute to filling the research gap by evaluating interactive maps against static maps. Two dimensions are of concern: firstly, users' effectiveness and efficiency when working with interactive maps; and secondly, an interactive map's potential to engage the users to explore the underlying data. For this purpose, an interactive and static version of a digital map are created and evaluated in an online user study.

2 Research questions

The first research question is aimed at evaluating the efficiency and effectiveness of users when using an interactive map.

Research question 1

How does an interactive visualization compare to an informationally equivalent static visualization regarding *usability issues*, in the context of public communication in online media?

The second research question is aimed at evaluating motivation and interest for exploring an interactive map.

Research question 2

How does an interactive visualization *engage users to explore the visualized data*, compared to a static visualization, in the context of public communication in online media?

Based on previous literature (section 3), a hypothesis is created for each research question.

Hypothesis 1

Users are less effective and efficient extracting information from a visualization of the interactive type than the static type.

Hypothesis 2

The interactive visualization type engages users more to explore the data.

3 Literature review

3.1 Map design

3.1.1 Map design for the internet

Jenny et al. (2008) note that maps designed for the internet should be designed even more simple than printed maps, containing few colors and only relevant information. This is because of limitations of screen resolution and larger viewing distances, as well as the fact that users will generally spend less time viewing an online map. Jenny et al. define a set of design principles for online maps. These will be applied in section 5.1.

3.1.2 Visualization of movement data

Andrienko et al. (2008b) provide an overview over some basic concepts of movement data, some of which are relevant to the work in this thesis: trajectories, space, time, moving entities, and related phenomena. A *trajectory* is defined as “the path made by the moving entity through the space where it moves” (Andrienko et al., 2008b, 18). Since a trajectory never happens instantly in a single moment in time, time is inherently linked to it. Typically, trajectories of different moving entities are analysed and compared. *Space* is defined as “a set consisting of locations or places” (Andrienko et al., 2008b, 20). Typically, this will be the physical geographic space. Two locations in space have a distance, and since space is continuous, there is an infinite number of locations between them and around them. For analytical purposes however, it can be useful to reduce the number of locations in space to a defined set. *Time* is defined as “a continuous set with a linear ordering and distances between the elements, where the elements are moments or positions in time” (Andrienko et al., 2008b, 23). In between and around any two moments, there is an infinite set of moments for which there are no data. Data for these moments can only be estimated by interpolation. Moments in time may be aggregated in intervals. An important characteristic of time is that it often includes cycles. These cycles can be natural, like the seasonal cycle of the year, or human-made, like the typical weekly work cycle. The *Moving entities* that are described in the data may themselves

have specific characteristics. For people, this might be their age or their means of movement. Finally, *related phenomena* have an influence on moving entities. These phenomena could e.g. include weather events or oil prices.

Andrienko et al. (2008a) identify some methods and techniques for analysing and visualizing movement data. “In a strict sense, visualization is representation of data in a visual form [...] For this purpose, items of data are translated into graphical features” (Andrienko et al., 2008a, 376). Relevant for the work in this thesis is the visualization of movements of multiple entities. This requires the aggregation of data. A resulting visualization can be either a discrete or continuous flow map. Discrete flow maps show movements as bands. The width of these bands is proportional to the magnitude of the flow that is visualized. Continuous flow maps show movements as vector fields or streamlines. Static maps can only show one moment in time or one aggregated time interval. To show a change over a period of time, the map must be iterated as multiples. However, the number of multiples that can feasibly be presented to a user is limited. Concluding this overview of visualization techniques, Andrienko et al. (2008a) point out that the map as a visualization medium is limited. To convey additional information, additional graphs, like histograms, may be shown alongside.

Doantam et al. (2005, 219) define three main characteristics of good hand-drawn flow maps: “intelligent distortion of positions, merging of edges that share destinations, and intelligent edge routing”. These techniques prevent the overlap of flow lines with each other and with origin and destination points, thus reducing visual clutter. Doantam et al. implement these principles into an automated environment with a layout adjustment algorithm. The goal is to automatically create maps with the visual quality of hand-drawn maps like the ones by Charles Joseph Minard (Minard, 1865, 1869). Similarly, Boyandin et al. (2010) explore a prototype of a flow map in which flows that have similar paths are automatically bundled together to follow the same path. A limitation in their implementation is that bundled flows are not aggregated, but overlap. As a result, the width of the bundles does not convey any information about the total magnitude of the flows in the bundle.

Jenny et al. (2016) test three design principles for origin-destination flow maps in a user study. They find that 1) curved lines score better than straight lines, 2) arrows score better than tapered flows to indicate direction, and 3) flows between nodes score better than flows between areas, in both user preference and user performance.

The merging and branching of edges in a flow map gives it the characteristics of a Sankey diagram that is overlaid onto a map. The Sankey diagram type

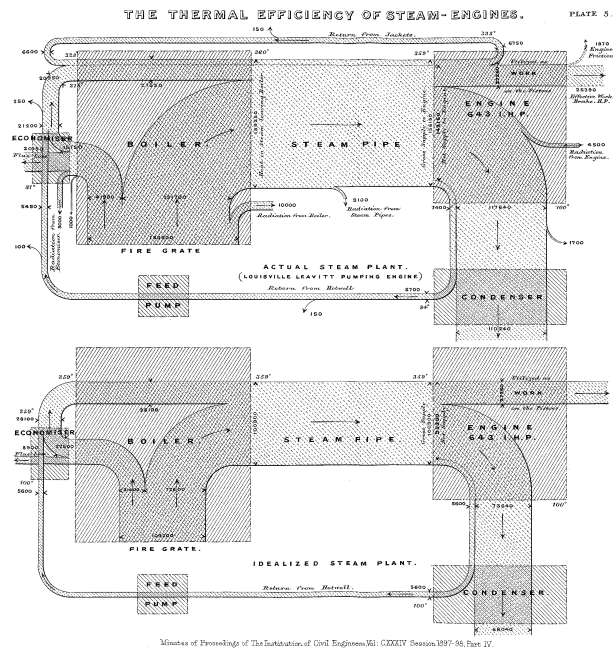


Figure 3.1: Sankey’s original diagram of 1898, showing energy flows in an actual (top) and ideal (bottom) steam engine (Sankey, 1898, Plate 5)

was created in 1898 by Matthew Henry Phineas Riall Sankey, who sought to illustrate inefficiencies in a steam engine system (Schmidt, 2008a). Fig. 3.1 shows Sankey’s original diagram of 1898. It effectively conveys the energy flows entering and leaving each part of the system, as well as energy losses. The Sankey diagram type was quickly adopted internationally by the engineering industry to show system energy flows, and later on the applications for the diagram were expanded, illustrating material flows in production systems (Schmidt, 2008a). The Sankey diagram is a tool not mainly to be used by experts, but by “an expanded circle of individuals who are not familiar with the details of the systems and to whom the crucial connections must be explained despite this” (Schmidt, 2008b). While a flow map can be seen as a Sankey diagram overlaid onto a map, Minard actually created his flow maps before Sankey had published his diagram (Minard, 1865, 1869). Minard’s flow maps are examined in more detail in section 4.1.1.

3.2 Interactive maps

Roth (2013) states that cartographic research is lagging behind technological developments in the field of interactive maps. He outlines a series of open re-

search fields and questions, some of which are relevant to this project. Maps allow people to generate new geographic insights, and interaction enhances visual thinking even more. Traditionally these two qualities (generation of insights and interaction) were exclusive to the exploration phase of research. Later on, in the presentation phase, static maps would then be produced for transferring information to the public. Increasingly, interactive maps that allow exploration have been aimed at the public, including in journalism. In this setting, visual thinking and visual communication are combined in one product. Here Roth raises the question whether the goal of interactivity is still the generation of insights, or something entirely different. Roth then raises the issue of interface complexity, stating that too much interactive freedom decreases productivity when working with a map. He then concedes that interactivity can be justified, e.g. to give the user a sense of control and thus motivation to study the map. Finally of relevance to this research project, Roth identifies characteristics of map users that might influence how they cope with interactive maps: ability, expertise, and motivation. Similarly to Roth with interactivity, Hegarty et al. (2009) find that realism and animation reduce effectiveness.

The study done in this thesis seeks to contribute to research in interactive maps by testing the effectiveness of interactive maps compared to static maps, as well as the effect of interactivity on user engagement.

Three research examples will be introduced here in which the effectiveness of interactive maps was tested. Albeit in all cases there was no direct comparison to static visualizations, the studies show possibilities and difficulties of interactive map usability testing. Andrienko et al. (2002) tested five different interactive tools for exploring and manipulating maps. They tested the effectiveness of CommonGIS implementations, a project which aimed to provide GIS tools for a broad community of users. Two test rounds were conducted on-site, and an additional round was conducted as an online study. The performance of the internet test users was much worse than that of on-site test users. Andrienko et al. identify four possible reasons: 1) on-site test users were introduced to the tools before testing, while internet test users had to rely on short illustrated explanations, 2) the internet test users had lower cartographic expertise, 3) most of the internet test users were required to take the tests, rather than taking them out of their own motivation, 4) explanations were in English, while the internet test users were native German speakers. The authors found that many of the online test users skipped some or all of the instructions and preferred a learning by doing approach. The second research example is by Çöltekin et al. (2009) who conducted an on-site user study of interactive map interfaces. They used a

combination of traditional usability engineering methods (satisfaction, efficiency and effectiveness, SEE) and eye tracking. The interfaces tested were Natlas and carto.net. Designer interviews showed that both tools were designed for average internet users with little or no cartographic expertise. Çöltekin et al.'s main finding that is of relevance to this research project is that users hardly ever consulted instructions when stuck, which is in accordance with Andrienko et al.'s findings. A third study by Poplin (2015) evaluated the usability of PPGIS for Wilhelmsburg. This online tool aims to allow people with no cartographic experience to add information to a map of Wilhelmsburg, the area of Hamburg where they live, by drawing on it with a pencil tool. The study was carried out on-site in two phases. The first phase had students with GIS experience as test subjects, while in the second phase the subjects were inhabitants of Wilhelmsburg. The second group is of relevance here. Contrary to the other two studies presented here, most participants read the provided instructions in the beginning or at least when they had problems using the interface.

3.3 Visualizations for the public

Zambrano and Engelhardt (2008) list some popular iconic uses of diagrams in recent years to illustrate how visual representations of data can grab the attention of large audiences in different media including online, making a profound impact on the public. They make the point that powerful visualizations give a broad public access to data, thus having an impact on public education and potentially ultimately leading to social change.

Segel and Heer (2010) develop a taxonomy of visualization types used in online media. In doing so, they focus on the degree of freedom that is given to the reader. They identify three types of visualization structure, each with a different balance between author-driven stories and reader-driven stories. The first type, mostly author-driven, is called the Martini Glass visualization structure. The reader is at first strongly constrained in the degree of interactive freedom by a narrative that is intended by the author. Once this narrative is completed, the interactive freedom increases like the shape of a Martini glass opening up from the stem, giving the reader more paths to follow in the narrative. The second type is called Interactive Slideshow. The reader is guided by the author along a typical slideshow, but enjoys the freedom to interact on each slide and then move on. The third type, mostly reader-driven, is called the Drill-Down Story. The reader is presented with a visual overview of the data, and is then able to explore different aspects and details at his own will, setting the stories to be

told in the process.

It is worthwhile to mention some current undergoing transformations in the journalism world. The journalistic environment is undergoing profound changes. The shift of advertisement revenue towards the internet, away from traditional printed newspapers, has forced newspaper editors and journalists to adapt new strategies (Picard, 2011, 2015). This way, the internet has changed the relationship between journalists and the public. The public is no longer just receiving and spreading, but also creating and producing news stories: anyone can be a journalist (Singer, 2011). But the internet does not only threaten the existence of journalists. It also provides them with new opportunities. The emergence of extensive data online, as well as the tools to analyse it, allows journalists to write new types of stories (Howard, 2014; Segel and Heer, 2010). One type of visualization that has emerged in online data journalism is the interactive map. It provides readers with the possibility to explore data on their own (Roth, 2013). Journalists have always made visualizations for the public, but new tools are opening up new opportunities.

In taking the new paths of data driven journalism, data journalists take care that they make sure to follow journalistic standards – in order to gain legitimacy both in the eyes of the public (Singer, 2011) and of traditionally working journalists (Coddington, 2015). Standards and recommendations for designing data visualizations exist in the journalism world (Gray et al., 2012), but there is a lack of systematic scientific research to underlie them.

3.4 Measuring engagement

Engagement is the focus of research question 2 of this thesis. The problem of how to measure engagement has been the focus of previous research.

O’Brien and Toms (2010) present a user engagement scale (UES) to measure users’ engagement with an interactive system like a software or website. They state that the success of technologies that are designed for end consumers depends not only on the technologies’ efficiency or whether they provide satisfaction for the user. The way a technology makes its users feel, and manages to engage them and provide them with an experience, is equally important. O’Brien and Toms’ engagement scale contains six dimensions: Perceived Usability, Aesthetics, Focused Attention, Felt Involvement, Novelty, and Endurability. With this large number of dimensions, the scale contains a high degree of complexity. This is confirmed by the statement that “attributes of engagement are highly intertwined, a complex interplay of user-system interaction variables” (O’Brien

and Toms, 2010, 50). O’Brien and Toms (2013) also note that the UES is not consistent across various user studies.

Schaufeli et al. (2006) present a questionnaire for measuring work engagement — defined as “a positive, fulfilling work-related state of mind” (Schaufeli et al., 2006, 702). Their so-called Utrecht Work Engagement Scale (UWES) has only three dimensions: vigour, dedication, and absorption. The UWES was first developed with 17 items (UWES-17) and later shortened to just 9 items (UWES-9). The shorter scale proved to be consistent with the earlier, more extensive scale. The UWES is a contribution to the field of occupational health psychology. Work engagement according to Schaufeli et al. is the antipode to burnout. Accordingly, they found that the UWES is negatively correlated with burnout. The UWES is specifically designed to measure a long-term state of mind that is not directly related to a single event or situation. However, it shares some questions of similar type with O’Brien and Toms’ UES.

Engagement and user experience are certainly a factor in the rise of data journalism. The research done in this study can contribute to a better knowledge of user engagement in online news features.

3.5 Migration on the Mediterranean Sea

Heller and Pezzani (2016) give an overview over the development of assistance and non-assistance to migrants in distress on the Mediterranean Sea. They mark two pairs of incidents as turning points in the practices of EU member states: the twin shipwrecks of October 2013, and those of April 2015. Before October 2013, the Mediterranean coastal states followed a practice of non-assistance toward migrant boats, justifying this practice by at the same time illegalising this kind of migration. In two shipwrecks in October 2013, more than 500 people died, creating an outcry in the European public. This led to a change of practice by the Mediterranean coastal states and to the start of the operation Mare Nostrum by Italy. The acclaimed goal of this operation was to save lives. Military vessels were intercepting migrant boats on a large scale and in a wide radius before they would shipwreck. This new practice, however, did not result in viewer deaths, since human traffickers reacted by using bigger, unseaworthy boats. Operation Mare Nostrum ended in November 2014 and was replaced with the Triton operation. Contrary to Mare Nostrum, which had the goal of saving lives, the aim of Triton explicitly was to protect the borders of Europe. The radius of action of military vessels was contained to the European coasts, maintaining a passive role in rescue operations. This left the task of rescu-

ing migrants in distress to commercial ships whose crew were not trained or equipped for this task. In two incidents in April 2015, 1200 migrants died on the Mediterranean while the crew of commercial ships were in the process of attempting a rescue. Following these incidents, rescue operations by commercial ships decreased dramatically. This gap was filled in part by an expansion of the Triton operation and a new operation EUNAVFOR MED (later renamed Sophia), aimed at combatting human trafficking. In large part however, rescue operations were now coordinated and carried out by non-governmental organisations and networks that were established for this cause. Heller and Pezzani in their summary focus on the practices of EU governments concerning the assistance of boat migrants. While the number of deaths on the Mediterranean Sea can also be related to the number of migrants (with the number of refugees peaking due to the Arab Spring and the Syrian crisis), these practices play a fundamental role in the development of the death toll. They should be communicated along with migrant flows and casualties on the Mediterranean Sea to enable a better understanding.

4 Gallery of visualizations

4.1 Static visualizations

4.1.1 Charles Joseph Minard's flow maps

Charles Joseph Minard's 1869 map of the march of Napoleon's army on Moscow (fig. 4.1) is his best known work (Friendly, 2002). It shows the number of Napoleon's soldiers as they march into Russia in the winter of 1812/1813 and then retreat, their rapidly diminishing numbers illustrated by the width of the flow line. The visualization portrays distance and direction truthfully, without having an actual geographic map that it is overlaid on. The low temperatures of the Russian winter that took their brutal toll on the French Army are graphed alongside the line of retreat.

Another example of Minard's work is his 1864 map of French wine exports (fig. 4.2). Flow lines are merged where they follow the same direction. Because it shows exports by sea, all flow lines are drawn outside of landmasses. In order to have enough space for the flow lines, land areas are distorted to make sea straits wider.

A modern example for a Minard-style flow map is by Rosamund Pierce (fig. 4.3). It shows the global coal trade. Like in Minard's wine map, flows are only shown outside of landmasses and aggregated, yet not merged, where they follow the same direction, thus flowing parallel alongside. Landmasses are not distorted, but the flow lines are lead generously around areas where there would not be enough space.

Flow maps like the ones by Charles Joseph Minard show how the abstract nature of a Sankey flow diagram can be combined with a map by showing some geographic properties of flows, while abstracting others to make the map more readable and to emphasize specific parts of the portrayed information.

4.1.2 Migration

The visualizations presented in this section show either migration flows of the current refugee crisis or missing migrants in the Mediterranean Sea. As such, their content is very similar to that of the visualizations produced for this thesis.

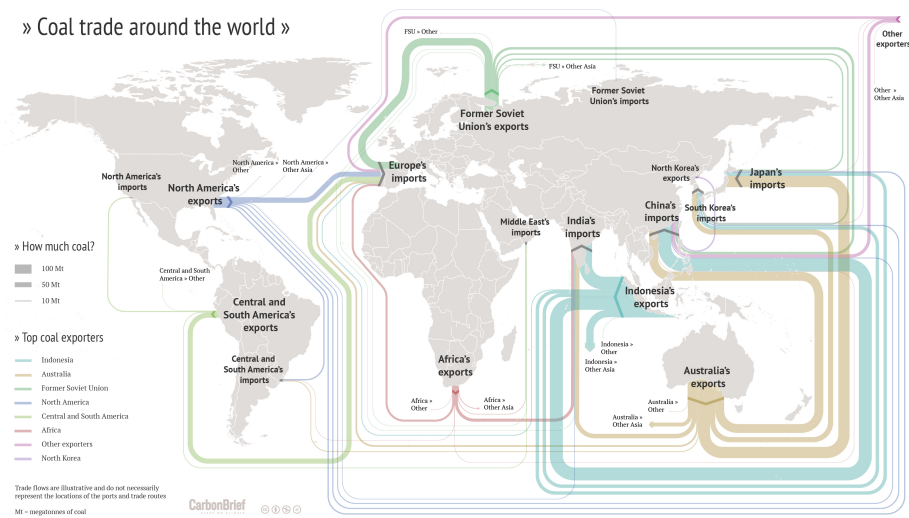


Figure 4.3: Flow map of coal exports around the world showing the top exporters (Pearce, 2016)

The Business Insider published a map in 2015 titled “Europe’s refugee crisis” (fig. 4.4). It is a simple flow map of refugee routes into Europe. The countries of the base map are colour-coded according to their EU/Schengen status. Unweighted flow line arrows in red show the migrant routes as they criss-cross into Europe.

Le Monde Diplomatique’s “Festung Europa 2016” (fig. 4.5) shifts the focus away from the flows a bit. Refugee flows are only indicated as they move toward Europe. The main content of the map are the various obstacles that refugees face when trying to get to Europe. By focussing on displaying only the relevant information, the map gives a lot of insight without being cluttered.

The Corriere della Sera in 2016 published a visually compelling, yet unsettling map of deaths in the Mediterranean Sea (fig. 4.6). The authors used data by the IOM (the same data source is used for the visualization in this project). On a basic topographic map with few labels and the Mediterranean in pastel blue, each migrant that died or went missing is portrayed by a black or red dot respectively. Dots are aggregated into clusters. The total number of deaths per region of the Mediterranean Sea is shown with a dashed circle. Additional graphs along the edge of the map show related information.



Figure 4.4: Europe's Refugee Crisis (BI Graphics, 2015)

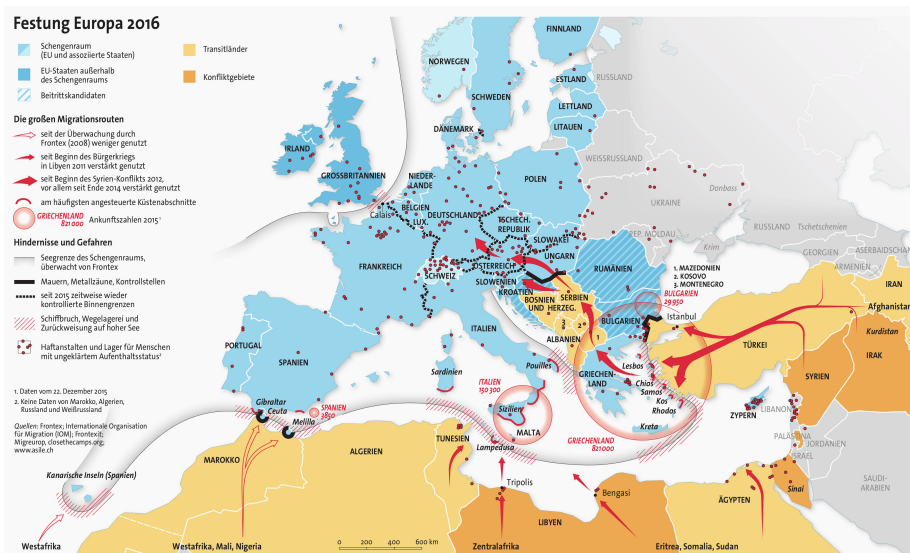


Figure 4.5: Festung Europa 2016 (Le Monde Diplomatique, 2016)

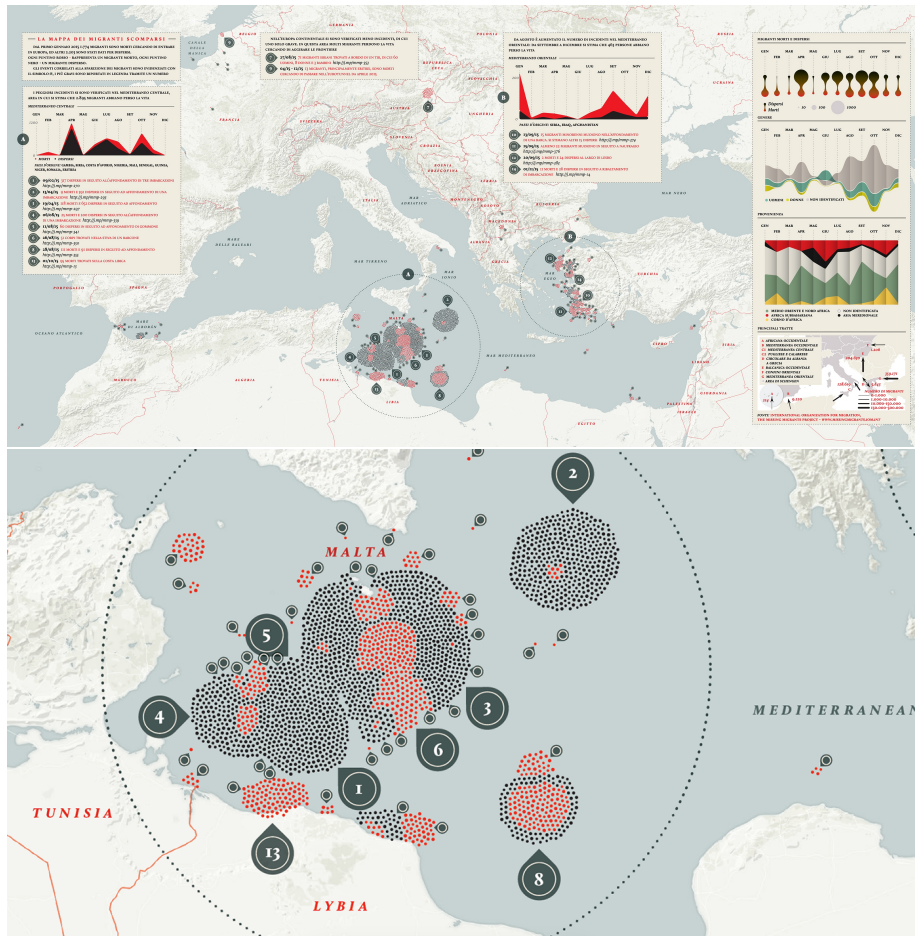


Figure 4.6: The missing migrants map (full map and detail) (Pellegrini and Mauri, 2016)

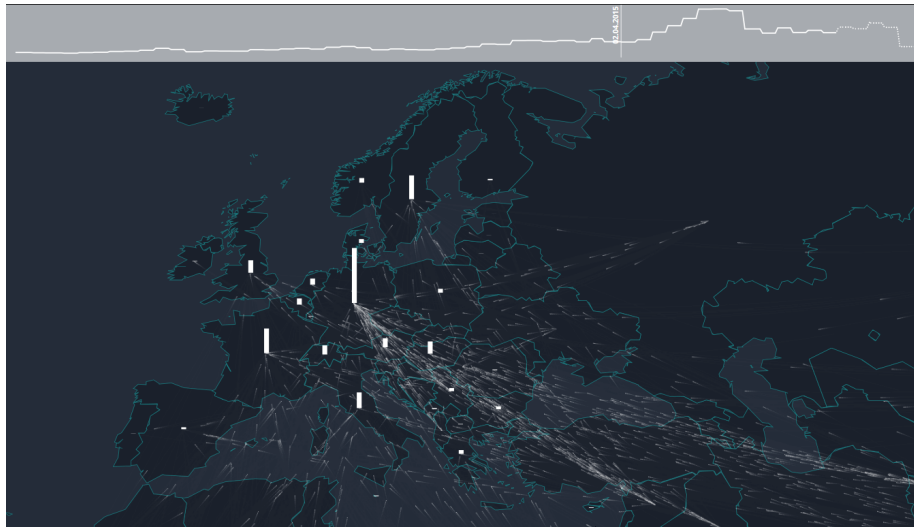


Figure 4.7: The flow towards Europe (Saarinen and Ojala, 2016)

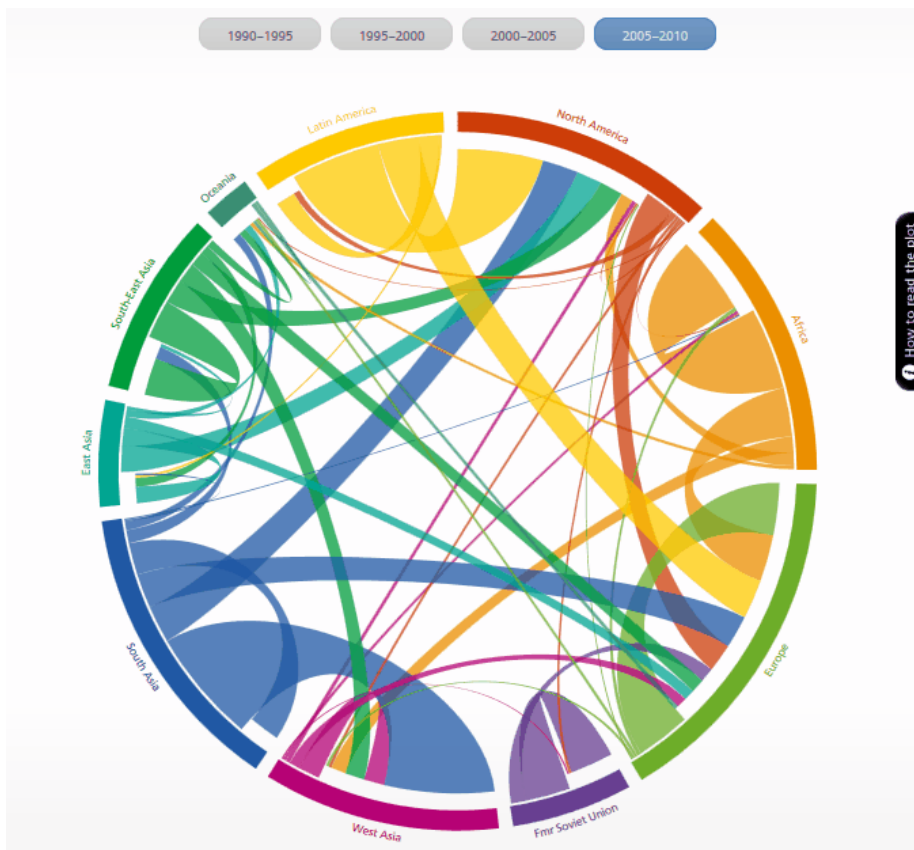


Figure 4.8: Circular migration plot showing worldwide migration (Sander et al., 2014)

4.2 Interactive visualizations

Two interactive visualizations are presented here that have the focus on migration. The first one, “The flow toward Europe” (fig. 4.7) by the Finnish visualization company Lucyfy, shows a continuous flow of white dots, representing migrants, streaming into Europe. A graph above the map shows the cumulation of migrants and allows the user interaction with the time displayed. Hovering over a country focuses on the flows to and from that country. A second example is Sander et al.’s global scale migration plot (4.8). It is unique in that it shows the regions of the Earth as segments of a circular plot. Clicking on the segment of a continent explodes it, showing the continent’s individual continents.

5 Methods

To answer the research questions outlined in section 2, a user study was conducted. As a case study, a visualization was designed in both an interactive and a static version. It shows migration data on the Mediterranean Sea, as well as incidents with dead or missing migrants, in the year 2015. A user study was then conducted online in which participants completed a number of task using the visualizations prepared for this purpose.

5.1 Visualizations

The visualizations designed for the user study are one of the main outputs of the project, besides the answers to the research questions. While they were designed to serve as the object in a user study, the goal was to design compelling visualizations that might have the potential to be developed further beyond the scope of the study. Accordingly, a substantial part of the resources for the project went into their design. An interactive visualization was first designed using the JavaScript library D3.js 4.0 (Data-Driven Documents, henceforth referred to as ‘D3’). D3 was created as a tool for implementing interactive visualizations in online environments (Bostock et al., 2011). Instances of the interactive visualizations were then exported as SVG files¹. The SVG files were then edited with Adobe Illustrator to create the static visualizations.

There are two main data sources for the visualizations. Data on migration flows was provided by the UN Refugee Agency (UNHCR, 2016b,c). Data on incidents with dead and missing migrants was provided by the International Organization for Migration’s Missing Migrants Project (IOM Missing Migrants Project, 2016). Further data sources were Natural Earth (2012) for the country geometries as well as a report by the House of Lords European Union Committee (European Union Committee, 2016, Figure 2) for the boundaries of international search operation areas.

A key consideration when creating the visualizations was obtaining informational equivalence between the interactive and static versions. Larkin and

¹using the New York Times SVG Crowbar bookmarklet for Google Chrome, <https://github.com/NYTimes/svg-crowbar>

Simon (1987, 67) define informational equivalence as obtained “if all of the information in the one is also inferable from the other, and vice versa”. In order to achieve this, the total time frame to be visualized was divided into the same number of smaller individual time frames for both the interactive and the static visualization types. For each individual time frame, the same information was visualized in both visualization types using the same map symbols. In the interactive version, users could select a time frame by clicking on a button. In the static version, users could scroll between individual map multiples for each time frame. In order to keep the layout clear and uncluttered (Andrienko et al., 2008a), the number of individual time frames was limited to four quarters.

The overall time frame to be visualized was determined by two factors: the goal of providing a conclusive narrative for the users, and data availability. Of the two factors, the first was the defining one, while the second was technical. Building on Heller and Pezzani (2016), a time frame with a conclusive narrative would be aligned with the shifts in European border control practices. The phases identified by Heller and Pezzani are: the non-assistance of migrants on the Mediterranean Sea (pre-October 2013); the Italian search and rescue operation Mare Nostrum (October 2013 to November 2014); and the EU border control missions Triton and Sophia (post-November 2014). The seasonality of the migration on the Mediterranean Sea can also provide a frame for a narrative, building on the concept by Andrienko et al. (2008b) of time cycles. Fig. 5.1 shows arrivals to Europe across the Mediterranean Sea for the years 2014 and 2015. While the rigid time frame of the calendar year is feasible for the seasonal migration flow in 2014, the flow of the year 2015 peaks late in the year in October and reaches its lowest point only in the first months of 2016, suggesting the frame should be widened beyond the calendar year. A general focus on the year 2015 provides an interesting narrative for the users of the visualization: migration numbers are higher than in 2014 by several factors (UNHCR, 2016a), and the international border control practices undergo some remarkable shifts (Heller and Pezzani, 2016). The exact borders of the time frame were eventually determined by data availability. The provision of data by the UNHCR on their online portal² is unfortunately not entirely systematic. The need to combine data for both Greece (UNHCR, 2016b) and Italy (UNHCR, 2016c), broken down to the desired levels for origin and destination areas, resulted in the calendar year of 2015 as the overall time frame. This time frame of 12 months also ensures that the four individual time frames, as described in the previous paragraph, are not too large and thus do not obscure seasonal change. The calendar year

²<http://data.unhcr.org/mediterranean/documents.php>

is also a time frame that most people should be very familiar with and should thus seem natural to the users of the visualization.

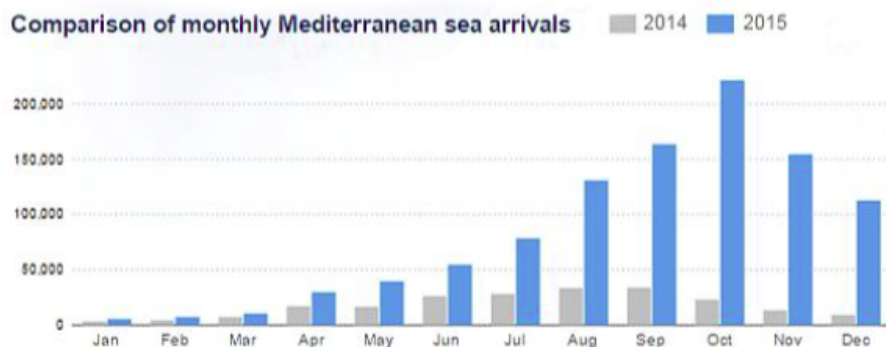


Figure 5.1: Mediterranean Sea arrivals 2014 and 2015 (crop of UNHCR, 2016a)

The spatial frame for the visualization was determined based on the magnitude of migration flows in the Mediterranean Sea. Mediterranean arrivals in 2015 to Spain, the third most important destination country next to Greece and Italy, were fewer by a factor of 40 compared to Italy and by a factor of more than 200 compared to Greece. Given the magnitude of these numbers, an exclusive focus on Greece and Italy was deemed suitable.

Two e-books provided the main resource for learning the D3 skills necessary for realizing this project: “Interactive Data Visualization for the Web” by Scott Murray (2013) and “D3 Tips and Tricks” by Malcolm Maclean (2015). Further resources were contributions by various users on Stack Overflow³ and Stack Exchange⁴. Where code snippets were directly adapted, they are referenced as comments in the source code found in appendix A. Note: whenever code lines are mentioned in this section, they refer to the source code of the visualization.

5.1.1 Basic map

The visualization has four basic elements types: a base map, migration flows, incident bubbles, and search operation areas. This section deals with the design and creation of each element type. Following that, some guidelines given by Jenny et al. (2008) for designing online maps are taken into account and applied. Even if their publication is from almost almost 10 years ago, it still provides a helpful resource. The sections that follow after that deal with the interactive

³stackoverflow.com

⁴stackexchange.com

elements that were implemented and with the adaptation of the visualization as a static map.

Base map

The base map consists of light grey country shapes with a very faint outline on top of a light blue background for the water bodies. Attention was given to obtain a good contrast between the two colours, while at the same time choosing colours with little saturation that visually remain in the background (code lines 28, 402, 419).

The country geometries were obtained from Natural Earth (2012) as a shapefile. In order to reduce rendering time, the geometries were reduced to only contain the desired countries using QGIS⁵. The same software was used to export the shapefiles as a JSON file for loading in D3 (code line 379).

Migration flows

Data on migration flows on the Mediterranean Sea are obtained from the UN-HCR (2016b,c). The data were aggregated in Excel⁶ to create the origin and destination regions, and then aggregated into the four time frames. They were then exported as CSV files for import in D3. Migration flows are displayed by flow lines similar to a Sankey diagram. There is one flow diagram for each destination country (Italy and Greece). The middle part of each flow line shows the country total for the number of migrants, while branches on each side show the number of migrants per origin and destination region. This is illustrated in figure 5.2. A dark grey colour was chosen to obtain an unobtrusive colour scheme while providing enough contrast with the background (code line 799). The region branches are in a slightly lighter grey and transparent to emphasize the middle part of the flow (code line 831). The paths of the flow lines do not follow the actual migrant routes — they are highly generalized. On the origin side of the flows, the lines branch out symmetrically with only little regard for the actual geographic locations of the origin regions. This allows for a clear distinguishability between the lines and prevents overlap. The lines then flow together to show the total number of migrants per country. The path of this line is hypothetical and follows the general direction of the migrant flow. On the destination side of the flows, the lines branch out again, ending in the actual geographic region of the migration flow's destination. Coordinates for the

⁵QGIS Desktop 2.16.0

⁶Microsoft Excel 2013

flow lines were prepared manually in a separate file and then joined with the migration flow data (code line 655).

Jenny et al. (2016) find that symmetrically curved lines perform best for origin-destination flow maps. Line paths are created using the SVG path mini-language. For the middle part of the flow lines, a straight line is drawn (code line 793). Curved lines are then added on each side as the region lines branch out (code line 839). To obtain a symmetric curve, a quadratic Bézier curve is drawn. The control point for the Bézier curve is generated dynamically from the coordinate data. The control point lies in the extension of the middle part of the flow line, so that the curve stems from the straight line with no sharp bend. The distance of the control point to the start of the curved line (where the branch leaves the middle part of the flow line) is proportional to the distance between the start and end points of the branch. To calculate the position of the control point for each branch curve, trigonometric functions are used (code line 613).

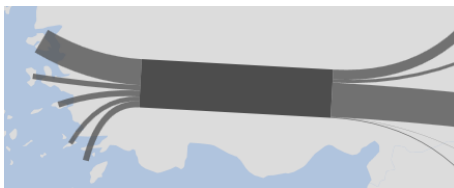


Figure 5.2: Detail of the visualization showing only flow lines

One challenge was to line the branch curves up next to each other, rather than having them overlaid, as was a limitation in Boyandin et al. (2010). In order for the branch curves to neatly line up, the starting point of each branch curve needs to be shifted along a line perpendicular to the middle part of the flow line (code line 711). The distance of the shift for each branch curve depends on the sum of the line width of all previous curves. To calculate the correct starting points for the curves, the same trigonometric functions are used as for calculating the positions of the control points for the Bézier curves (code line 613).

Since all flows go in the same direction, no graphic indication of direction was used. However, the direction of the flows was pointed out linguistically (sections 5.1.2 and 5.1.3).

Incident bubbles

For each incident where migrants died or went missing, a bubble is shown on the map (fig. 5.3). The area of each bubble is proportional to the combined

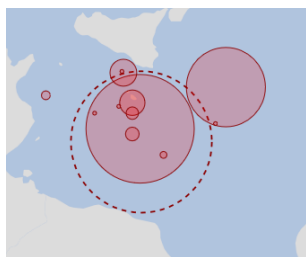


Figure 5.3: Detail of the visualization showing only incident bubbles

number of dead and missing people, using mathematical scaling (Slocum et al., 2010, 305). The bubbles have a red color, making them stand out in the color scheme of the map. To enable visibility for underlying bubbles when bubbles overlap (Slocum et al., 2010, 316), they are transparent with a solid red stroke (code line 515). The bubbles are sorted before drawing so that smaller bubbles are drawn on top (code line 483). To show the total number of migrants per migrant route (central and western Mediterranean), circles with no fill and a dashed stroke are drawn (code line 596).

The positions of the bubbles on the map are directly taken from the data by the IOM Missing Migrants Project (2016). While not having done a systematic check on the validity of the coordinates, some cases were encountered where the positions were clearly off. One common occurrence is a location description that includes the phrase “unspecified location between [...]”. Incidents with such a location description still have precise coordinates in the source data. This kind of uncertainty was not accounted for when designing the map. Another occasional occurrence is when the coordinates do not match the location description. In one case, an incident with the description “30 miles off the coast of Libya” is actually placed some 120 miles off Libya. By placing a bubble precisely on the map, a false sense of precision is given to the map users. However, the main goal of the incident bubbles is to give users an overview of the general pattern of incidents. For this purpose, the bubbles work well, since they show the spatial pattern, as well as patterns regarding the number and size of incidents. Another limitation of placing the bubbles on the locations given in the coordinates, besides the false sense of precision, is the overlapping of bubbles (fig. 5.4). The overlap prevents the identification of individual bubbles. This issue could have been resolved by using collision detection force functions in D3⁷ for the interactive version or by moving the bubbles manually for the static version. Neither of the options were put into practice, mainly due to scarce time resources in the project.

⁷<https://github.com/d3/d3-force#collision>

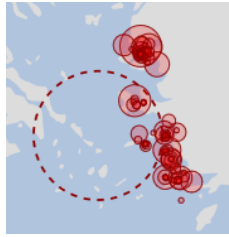


Figure 5.4: Detail of the visualization showing overlapping bubbles

Search operation areas

The geographic extent of the Triton and Sophia search operation areas was adopted from a map by the UK House of Lords European Union Committee (2016). The areas were digitalized manually in QGIS⁸ and then loaded in the visualization as a JSON file. To emphasize the area while making sure that it would not interfere with the distinguishability of migration flow lines and incident bubbles, a dashed fill with a high transparency was chosen (code line 430).

Applying Jenny et al.’s design guidelines for digital maps

Choosing an appropriate viewing technology (Jenny et al., 2008, 33). Web browser compatibility is important to consider when designing web visualizations for the public. Users want to use their standard browser. Jenny et al. point out that SVG is a problem, since it is not widely supported. Current modern browsers however do support SVG⁹. The D3 wiki states that it is tested against current versions of Firefox, Chrome, Safari, Opera, IE9+, Android and iOS¹⁰. In the development of the visualization, it was continuously tested on Firefox, Edge, and Chrome¹¹. CSS styling rules (code line 7) did not always work for Firefox, so styles were directly applied as element attributes in the D3 code.

Considering transmission speed (Jenny et al., 2008, 34). Transmission speed is an issue when working with large amounts of data. The amount of data used in this visualization is relatively small. Nevertheless, effort was put into making the data source files as small as possible by pre-filtering them to only include relevant entries.

Designing for varying screens sizes and resolutions (Jenny et al., 2008, 35).

⁸QGIS Desktop 2.16.0

⁹<http://caniuse.com/svg>

¹⁰<https://github.com/d3/d3/wiki#supported-environments>

¹¹Mozilla Firefox 50.1.0, Microsoft Edge 38.14393.0.0, Google Chrome 55.0.2883.87

Screen size is a relevant issue today more than ever, with mobile devices like smartphones making up a big share of internet use¹². This could be seen in the user study of this thesis: of the 294 people that started the study, 54% (158 people) used a mobile device. It was not possible however to design the visualization in a way that it was displayable on the small screens of mobile devices. Furthermore, users needed a mouse as an input device. Therefore, participants were told at the very beginning of the study to display the visualization on a computer, not a mobile device. In order to adapt to varying screen sizes and resolutions, the visualization was made responsive to always fill the entire width of the browser window (code line 46). The aspect ratio of the visualization was chosen to fit a 16:9 screen, which is the most common aspect ration on current computers¹³. On a more “quadratic” aspect ratio like 4:3, the visualization would just be scaled smaller and still be viewable without being cropped, since it adapts to the width of the screen.

Increasing legibility of map elements: anti-aliasing (Jenny et al., 2008, 37). Anti-aliasing is the smoothing of jagged edges. Since the SVG elements of D3 are rendered directly in the browser, the browser takes care of anti-aliasing. However, as Jenny et al. (2008) state, anti-aliasing makes SVG rendering time-consuming. This did make a noticeable difference specially for the country geometries, where the SVG rendering time could significantly be reduced by loading only the desired countries.

Making signatures and symbols readable: dimensions and distances (Jenny et al., 2008, 39). For incidents, the bubble symbols are scaled such that the smallest possible bubble (1 migrant) is easily distinguishable from the background. For flow lines, this was not entirely possible, since the differences in magnitude between different flows are that significant. By grouping origin and destination regions together, some of the thinnest flow lines could be avoided.

Designing text: screen typography (Jenny et al., 2008, 41). Since space on the screen is limited, some of the font is quite small (depending of course on the screen size). Legibility was tested on a 12.5 in laptop screen. To increase legibility, a sans serif font was used.

Dealing with inconsistent color display (Jenny et al., 2008, 44). Since a very reduced colour scheme was used, discernibility between colours is not an issue. In order to account for deuteranopia, Color Oracle¹⁴ was used.

¹²<http://gs.statcounter.com/comparison-market-share/desktop-mobile-tablet>

¹³<http://gs.statcounter.com/screen-resolution-stats/desktop>

¹⁴Color Oracle version 1.2.1

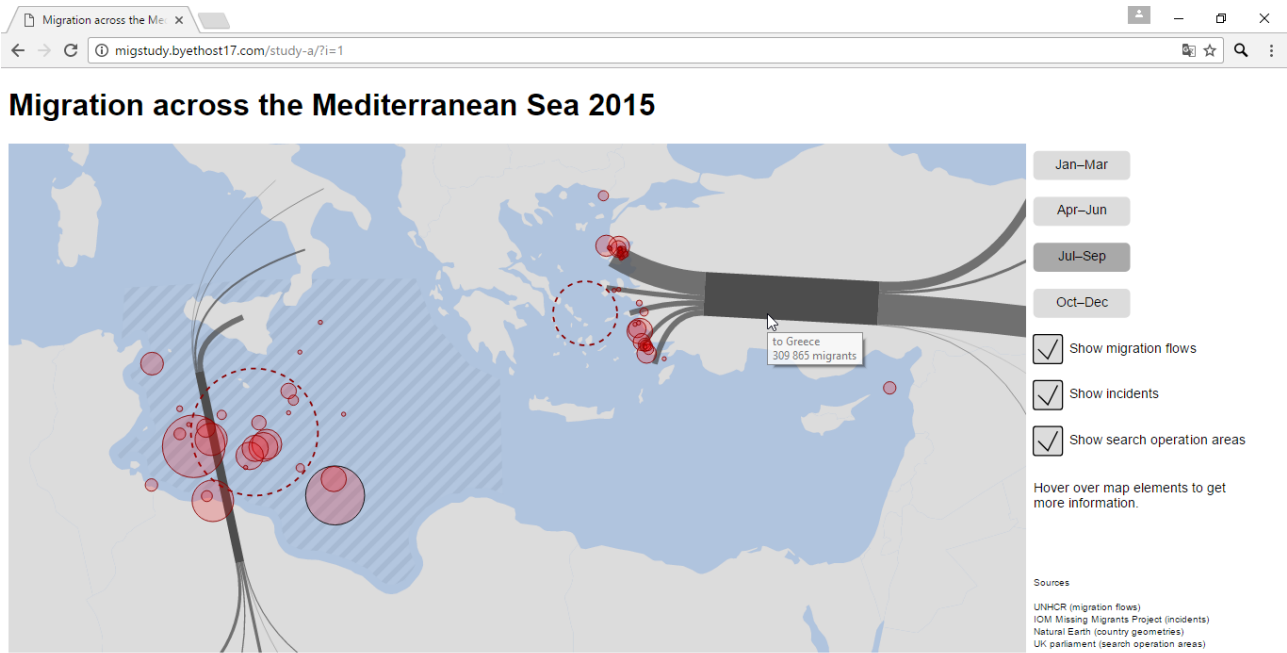


Figure 5.5: Interactive version of the visualization, showing a tooltip on hovering over an element, as displayed in in a web browser

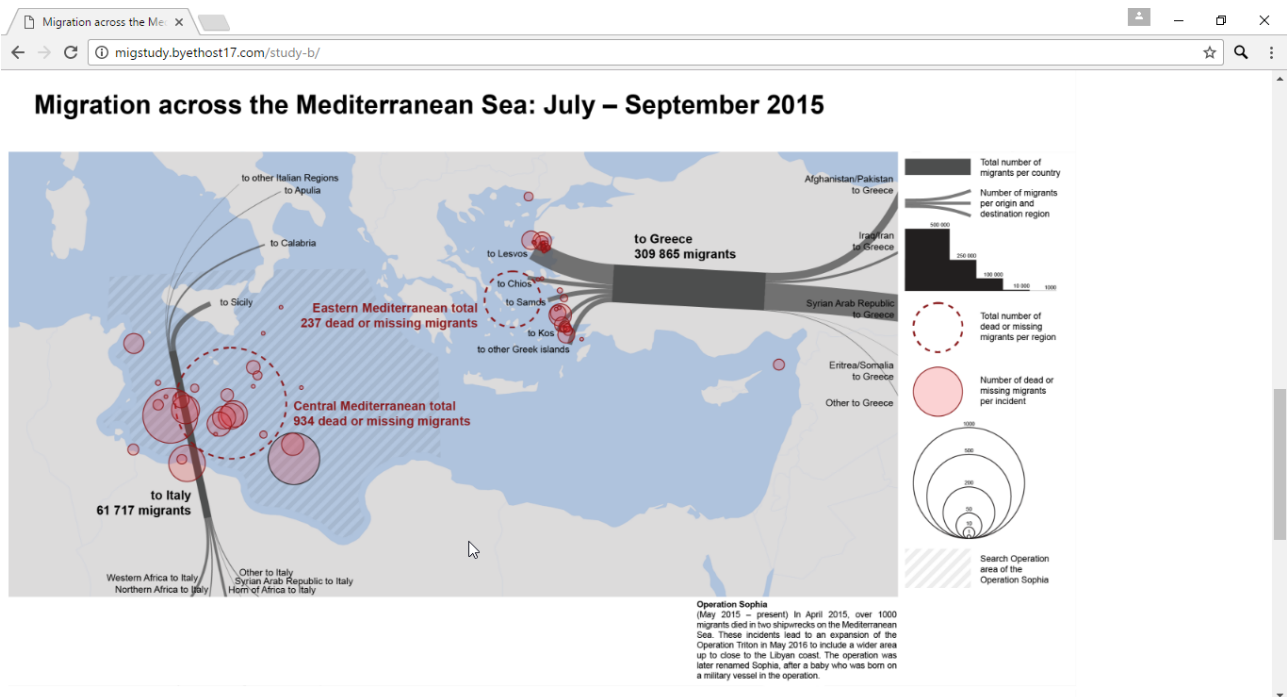


Figure 5.6: Static version of the visualization, as displayed in a web browser

5.1.2 Interactive elements

Interactive elements in the visualization are tooltips, time period buttons, and visibility checkboxes. Whilst creating the interactive version, one focus was to build a visualization that does not need any lengthy instructions and can be learned by using it. A screenshot of the interactive version of the visualization is shown in figure 5.5¹⁵.

Tooltips

Tooltips are text elements that appear when the user hovers the mouse pointer over an element in the visualization. The tooltips implemented in the visualization are a built-in functionality by web browsers. These are easy to implement, yet have some drawbacks (Murray, 2013)¹⁶. They are non-customizable, so they are not very pretty to look at. Also, they do not show up immediately after hovering over an element, but only after 1–2 seconds. To make users aware of the tooltip functionality, an indication was added at the side of the map (fig. 5.7). Also, most elements that have tooltip functionality implemented also react to the mouse pointer in other ways, e.g. by changing their colour or opacity. This effect is immediate and helps users become aware of interactive functionality connected to the element. There are tooltips for all of the four basic element types of the map.

Base map: For the base map, tooltips only exist for the two destination countries Italy and Greece. On hovering the mouse pointer over them, they turn dark grey (code line 388) and show the country name as a tooltip (code line 404).

Flow lines: When hovering the mouse pointer over any flow line, a tooltip is shown with two lines of text (code lines 805, 855). The first line contains the name of the destination country (for the middle part of the flow) or the name of the origin or destination region (for the outer branches of the flow). For the lines of the origin regions, the tooltips show the name of the origin region as well as the destination country. Since the flow lines themselves have no indication of direction, the tooltips contain a linguistic indication of direction (e.g., “to Calabria” or “Iraq/Iran to Greece”). The second line of the tooltip contains the number of migrants for this particular flow. Example tooltips (oblique indicates line break): “to Greece/62 763 migrants”, “Western Africa to Italy/15 577 migrants”. Some of the flow lines are very thin and hovering over them

¹⁵At the time of the publication of this thesis, the interactive visualization can be viewed under the following link: <http://migstudy.byethost17.com/study-a>

¹⁶http://chimera.labs.oreilly.com/books/1230000000345/ch10.html#_default_browser_tooltips

would be very frustrating and near impossible. To facilitate this, ghost flow lines are added over every flow line (code lines 802, 834). These ghost lines have their opacity set to 0 and have a greater width than the original line, greatly facilitating the use of tooltips.

Incident bubbles: The tooltips for the two types of incident bubble differ slightly. For the dashed circles showing the total number of dead or missing migrants per migrant route, the tooltips are made up of two lines of text. The first line contains the designation of the migrant route, and the second line contains the number of migrants (code line 603). Example tooltip (oblique indicates line break): “Central Mediterranean total/1272 dead or missing migrants”. For the transparent bubbles showing the number of dead or missing migrants for individual incidents, the tooltips are made up of three lines of text. The first line contains the number of dead or missing migrants, the second line contains a description of the location, and the third line contains the date that the incident occurred (code line 541). Example tooltip (oblique indicates line break): “46 dead or missing migrants/Unspecified location south of Sicily, Italy/04.05.2015”. The incident bubbles also show a direct reaction to hovering over them, changing their fill from transparent to solid and getting a wider stroke (code line 523). This makes it possible to identify the individual bubble for which the tooltip is being displayed, and so makes it possible to see through some of the clutter caused by overlapping bubbles (fig. 5.4).

Search operation areas: The search operation areas have the most comprehensive tooltips of all element types. They contain a short info text about the Triton or Sophia search operations (code line 450). Also while hovering over the area of the search operation, its opacity changes from very faint to fully solid (code line 434).

Tooltips for the search operation areas:

Operation Triton (October 2014 – May 2015)

After two shipwrecks in October 2013, which left a total of over 500 migrants dead, the Italian authorities started the Operation Mare Nostrum, which was conducted by the Italian Navy. The goal of Mare Nostrum was to save migrants from dying in incidents on the Mediterranean Sea.

The Operation Triton replaced Mare Nostrum in October 2014. Triton is conducted by Frontex (the EU’s border security agency). The main goal

of Triton is to secure the European borders against illegal immigration. Compared to Mare Nostrum, the operational area of Triton was reduced and limited to areas near the Italian coast.

Operation Sophia (May 2015 – present)

In April 2015, over 1000 migrants died in two shipwrecks on the Mediterranean Sea. These incidents lead to an expansion of the Operation Triton in May 2016 to include a wider area up to close to the Libyan coast. The operation was later renamed Sophia, after a baby who was born on a military vessel in the operation.

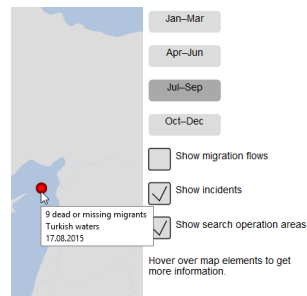


Figure 5.7: Detail of the interactive visualization showing interactive elements

Time period buttons

Perhaps the most important interactive feature of the visualization is the ability to change the time period that is displayed by clicking a button. There is a button for each of the four time periods. Clicking it changes the data source for incident bubbles, flow lines and search operation areas to the corresponding time period and then updates all map elements (code line 157). The buttons themselves are held in a colour scheme that matches the map (fig. 5.7). Using a loop function to create the buttons CSS, styling rules are can be avoided (they caused problems, see section 5.1.1) while still avoiding duplicate code (code line 121).

Visibility checkboxes

All map elements of each element type (flow lines, incident bubbles, and search operation areas) can be unchecked, causing them to disappear from the map. The idea behind this is that the user can turn off distracting elements in order to

better be able to focus on relevant elements. For each element type, a checkbox is created using a loop function (code line 192) in a colour scheme that matches the map and the time period buttons (fig. 5.7). For each checkbox, there is a binary variable that registers whether it is ticked or not. Clicking a checkbox changes the binary variable and calls the update function (code line 248). In the update function (code line 300), first all of the map elements are created and the ticks in the checkboxes are made visible, regardless of the state of the binary variable. Then, for each of the element types a function is called that checks the state of the variable, and if its state is false, removes the elements of the corresponding element type and sets the opacity of the tick in the box to 0 (code line 270).

5.1.3 Static version

In order to create the static version of the visualization, instances of the four time periods in the interactive version were exported as SVG files¹⁷. The files were then adapted in Adobe Illustrator¹⁸. Since in the static version there are no tooltips to give information about the character and scale of the features that are represented in the map, labels and a legend were added. A screenshot of the static version of the visualization is shown in fig. 5.6¹⁹.

Legend

The legend (fig. 5.8) shows the map symbol types for incident bubbles, flow lines, and search operation areas, as well as (for the former two) their scale (Slocum et al., 2010, 194-197, 314-315). The info text about the search operation area is placed below the map and is the same as in the tooltips in the interactive version.

Labels

Each part of the flow line is labelled. The labels match the first line of the tooltips in the interactive version, also containing a linguistic indication of direction. This makes the map slightly cluttered, but ensures information equivalence between the two map types according to Larkin and Simon (1987). For the total country flows, the exact number of migrants is also shown on the map.

¹⁷using the New York Times SVG Crowbar bookmarklet for Google Chrome, <https://github.com/NYTimes/svg-crowbar>

¹⁸Adobe Illustrator CS6 16.0.3

¹⁹At the time of the publication of this thesis, the static visualization can be viewed under the following link: <http://migstudy.byethost17.com/study-b>

Thus, the labels for the country flows match the tooltips in the interactive version. The second type of label is for the circles showing total number of dead or missing migrants per migration route. These labels also contain the total number of dead or missing migrants, thus making the labels exactly match the tooltips in the interactive version. The label colour gives an indication of the element type it describes.

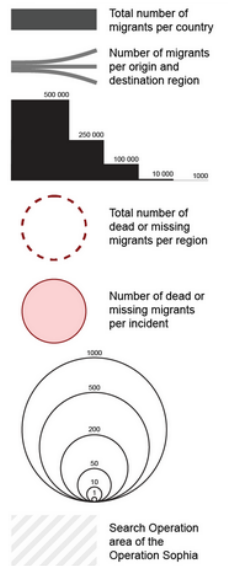


Figure 5.8: Detail of the static visualization showing the legend

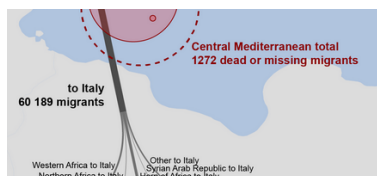


Figure 5.9: Detail of the static visualization showing labels

Presentation

Because the maps are too complex to be viewed as small multiples at the same time, they were designed to be displayed at full screen. To view the different maps, the user needs to scroll vertically between them. Arguably, this scrolling can be seen as a form of interaction, making the visualization not 100 percent static.

5.2 User study

A user study was conducted online on the survey portal LamaPoll²⁰. The survey consisted of three parts: a pre-questionnaire, a main questionnaire where participants completed a number of tasks using a visualization of either the interactive or the static visualization type, and a post-questionnaire. Before the actual questionnaire started, participants were directed in getting set up.

This section gives an overview over the structure of the questionnaire. Operationalisation of concepts and variables is explained in more detail in section 5.3. The user study ran from 2.1.–10.1.2017. Before it was distributed to participants, it was piloted by two individual people who gave a number of comments that lead to minor changes in the survey design.

5.2.1 Getting set up

For technical reasons, it was not possible to integrate the survey questionnaire and the visualization onto the same web page. It was thus necessary to instruct participants on getting set up (Pages 1–4 of the survey in appendix B). As mentioned earlier in section 5.1.1, the visualization is not suited to be displayed on small screens. Participants were instructed to either open the visualization and the survey in separate browser windows or tabs, or to open the visualization on a suitably large screen and the survey on a mobile device. For this purpose, participants could also close the survey and start it again on a different device. Overall, 54% of the participants that started the survey opened it on a mobile device. Of the 91 participants that completed the survey to the end, only 21% were on a mobile device, suggesting that most participants favoured the option to have the visualization and questionnaire on the same device.

5.2.2 Pre-questionnaire

The pre-questionnaire ranges from pages 5–11 of the survey (appendix B). *Getting used to the map*: Participants were first told to take their time to look at the visualization, without focussing on the content, to get a technical understanding of how the map works. *Engagement*: This was followed by two items to assess engagement. *Exploration*: The participants then had the opportunity to explore the content of the map, taking notes of anything they found interesting. They were told to take as many note as they like and move on as soon as they want to. This is a learning-by-doing approach as encouraged by Andrienko et al. (2002)

²⁰<https://www.lamapoll.de/>

that should engage the user more than reading instructions. *Preference*: Participants were asked about their perceived effect of interactive elements. *Personal interest and experience in relevant fields*: Participants scored their interest and experience in the fields of maps, data visualization, migration, and online news stories.

5.2.3 Main questionnaire

In the main questionnaire (pages 12–28 of the survey in appendix B), participants worked with the visualization to complete a number of tasks. Details about this can be found in section 5.3.

5.2.4 Post-questionnaire

The post-questionnaire ranges from pages 29–37 of the survey (appendix B). *Exploration*: After having worked with the visualization, participants are for a second time given the chance to take notes and give comments. *Engagement*: Four items assess engagement to detect change from before working with the visualization. *Preference*: The same items as in the pre-questionnaire allow the detection of change in preference. *SUS*: For a general measure of usability, the system usability scale is used. *Opinion*: Revisiting the items from the pre-test allows the detection of change. *Demographics*: Participants state their age, gender and level of education. *Language problems*: Since English is not the first language for most participants, they are asked whether they had trouble understanding some of the questions.

5.2.5 Recruitment of participants

Participants were mainly recruited in the personal network of the author of this thesis, mainly using the messaging platform WhatsApp²¹. People were also encouraged to themselves share the link to the survey in order to achieve a greater number of participants. The survey was also shared privately on the online social media platform Facebook²² and posted in three Facebook groups: *I'm interested in Migration!*²³, a group with about 3000 members with an interest in migration; *Sharing is Caring Universities of Zurich/ETH (UZH/ETH)*²⁴, a group with about 10 000 members who are students in Zurich; and *EGEA* -

²¹<https://www.whatsapp.com>

²²<https://www.facebook.com>

²³<https://www.facebook.com/groups/182579205255346>

²⁴<https://www.facebook.com/groups/sharingiscaringunizurich>

*Official Group*²⁵, a group with about 4000 members who are all members of the European Geography Association. In the group about migration, which has members from all over the world, the text distributed with the link to the survey contained the note that the study is aimed at people living in Europe, and that a basic understanding of the refugee crisis in the Mediterranean Sea is required. Recruitment was done over several days, starting on 2.1.2017. Participants took the survey until 10.1.2017. Of the 294 people that started the survey, 109 (37%) reached the survey page via a direct link from Facebook. This includes the aforementioned groups, but also private sharing by the author of the survey and others. Nothing can be said about the referral pages of the remaining participants, since they used a direct link.

91 participants completed the survey, which translates to a completion rate of 31%. 148 participants left the survey on the page “Getting set up” (Pages 3-4 of the survey in appendix B) or before. Of the 146 participants that started the actual survey (starting on page 5 of the survey in appendix B), 62% completed the survey to the end.

5.3 Experimental design

5.3.1 Between-subject design

The study was conducted online. While this provides less control over the environment in which the participants take the survey, it provides a realistic setting for the context of public communication in online media, as defined in the research question.

For the study, a between-subject design was used. The main reason for this was to allow participants to get engaged with a particular visualization type as thoroughly as possible. Switching between visualization type would reduce this experience. A within-subject design where participants would use one visualization type after the other was not an option either, since the time for the participant to take the survey needed to be minimised, particularly since they were taking the survey online and unsupervised.

Participants were assigned to one of the two groups (interactive and static visualization) based on the day of the month they were born.

²⁵<https://www.facebook.com/groups/egeagroup>

5.3.2 Research question 1: usability

Accuracy

Knapp (1995) defines four visual operators that can be used to define tasks that make use of geographic visualizations. The operators can directly be understood as general task types. For each of the four visual operators, tasks were designed. The task types *identify*, *locate*, and *compare* are analyzed overall, since they allow for a rating of answers as either correct or incorrect. For questions where a quantity was asked, answers in a range of $\pm 25\%$ of the exact answer were counted as correct. Where no answer was given, the answer was rated as incorrect. However, it was possible that a participant did not get to see a specific question, if it was a follow-up question to a question that they did not answer. In that case, the question was omitted for this participant. For the visual operator *associate*, no accuracy could be measured, since the character of the tasks did not allow for a rating as correct or incorrect. Tasks for the visual operators were:

Identify

- Identify the meaning of map symbols
- Determine the size of a flow or an incident bubble

Locate

- Determine the geographical location of a map element
- Determine the temporal location of a map element

Compare

- Find the biggest occurrence of a map symbol type
- Rank several map elements by size

Associate

- Judge the influence of factors influencing the number of dead or missing migrants
- Judge the relationship between origin and destination regions

Confidence and ease

Confidence was measured after each question on a five point Likert scale. However, it was recorded only for questions of task types where answers can be rated right or wrong (*identify*, *locate*, and *compare*, corresponding to accuracy), because the answers to tasks of the visual operator *associate* were given on a five

point Likert scale, implying confidence. A concept similar to confidence that was measured across all task types is ease. Analogous to confidence, survey participants gave a rating on a Likert scale of the ease after each question. The two concepts are no doubt related, yet their values are reported independently alongside each other.

Response time

Response time was measured throughout the survey. It was recorded individually for each page that the participant viewed. This includes the time the participant needed to answer the follow-up questions about confidence and ease. Since answering these follow-up questions was required, this factor was the same across all participants. Where more than one question was displayed on the same page, the response time assigned to each individual question is the mean of the total number of questions on the page that were answered.

According to Field (2013, 317) reaction time is “notoriously messy”. The same can be assumed of response time in a study setup where participants are unsupervised and can be distracted at any time. In order to remove outliers, a standard deviation based rule was used. For each individual question, values that were more than 2.5 standard deviations from the mean were removed. This process was repeated iteratively until there were no values more than 2.5 standard deviations from the mean. For most of the questions, this took 1–2 iterations, however for some questions it took up to 7 iterations. The time was then averaged for each participant across all questions for which there was a value. For this mean, the standard deviation rule was again applied. Only one outlier was removed.

Response time was analyzed across all task types and regardless of whether the participants had answered correctly or not. This allowed the inclusion of the task type *associate* in the analysis, and it is also justifiable with the fact that participants who answered a question did spend time on it, regardless of whether the answer was correct. However, questions that were not answered were excluded, for two reasons: firstly, participants are likely to only have glanced at the question without spending too much time on it; and secondly, the follow-up questions for confidence and ease were not shown if a question was not answered, which would be an automatic shortcut.

System usability scale

The system usability scale (Brooke, 1996) is a standardized tool that provides a simple, yet reliable method for assessing subjective usability as perceived by

the users of a system. The scale consists of 10 items (5 positive and 5 negative). For each item, the users score their agreement on a Likert scale. The individual scores are then computed into a single score across all users, allowing a direct comparison between differing systems.

5.3.3 Research question 2: engagement

Engagement

Based on O'Brien and Toms (2010) and Schaufeli et al. (2006), items were created to measure engagement before and after the participants worked with the visualization. Because of the insecurities concerning the robustness of the user engagement scales (UES) by O'Brien and Toms (see section 3.4), Schaufeli et al.'s Utrecht Work Engagement Scale (UWES) was used as a base. Because it provides more items that could be applied from the general setting of the UWES to the required situational setting, the earlier UWES-17 with a greater number of items was used. For the pre-test, only two items were created as a baseline measurement, since any concepts of engagement applying to the visualization can not be used yet. For the post-test, for each of the three dimensions vigour, dedication, and absorption, at least one item was created. Because absorption was perceived to be the most important in the setting of getting engaged by an online visualization, two items of this dimension were included.

The items created by Schaufeli et al. are all formulated in positive form. This roots in the basic concept of the scale, since it was meant to counter the sole negative focus in previous work in occupational health psychology. For this study, the items are balanced by reformulating some of them into negative form. Also, the items were reformulated to match the specific situational requirements of the study.

The items that are used in this study to measure engagement are presented here. For reference, for each item the original item by Schaufeli et al. (2006, 714) that it is based on is listed. To illustrate the similar character of the UWES and UES, an item by O'Brien and Toms (2013, 1100) is also included where applicable.

Pre-Test:

- At this moment, I am tired and low on energy.
 - Schaufeli et al.: At my work, I feel bursting with energy (vigour)
- I am excited to see what waits for me in this study.

- Schaufeli et al.: I am enthusiastic about my job (dedication)
- O'Brien and Toms: I felt interested in my searching tasks (novelty)

Post-Test:

- The visualization helped to fully immerse myself in the topic.
 - Schaufeli et al.: I am immersed in my work (absorption)
 - O'Brien and Toms: I was absorbed in my searching task (focused attention)
- At this moment, I am tired and low on energy.
 - Schaufeli et al.: At my work, I feel bursting with energy (vigour)
- The tasks felt tedious and boring for me.
 - Schaufeli et al.: Time flies when I am working (absorption)
 - O'Brien and Toms: The time I spent searching just slipped away (focused attention)
- Working with the visualization felt meaningful to me.
 - Schaufeli et al.: I find the work that I do full of meaning and purpose (dedication)
 - O'Brien and Toms: I felt involved in the searching tasks (felt involvement)

To analyze the effect of the visualization type on engagement, statistical methods are used. Additionally, an engagement score is proposed, inspired by the system usability scale by Brooke (1996). Analogous to the SUS, the scale position of each item is computed into a score contribution ranging from 0 to 4 (taking account of whether items are formulated in positive or negative form). The score contributions of the post-test items are then summed with a weight of 5. This results in an engagement score ranging from 0 to 80. When the items of the pre-test have a high score, a penalty is taken from the engagement score, because there was no positive change between the pre-test and the post-test. The score contributions of the pre-test items are given a weight of 2.5 and then subtracted from the score. The maximum penalty applied to the score this way is 20, resulting in a score ranging from -20 to 80. By adding 20 to the overall score, the final engagement scale ranges from 0 to 100. Like the SUS, the proposed engagement scale has the potential to measure the effect of a system on the user's engagement in a quick and easy way.

Exploration

After having had a first look at the visualization and after having worked with the visualization solving the tasks, the participants had the chance to take notes and make comments. They were absolutely free to take as many or few notes as they like, or to move on immediately. The individual statements that were collected this way were grouped into four categories: content-related statements, statements that stemmed from a misinterpretation of the map, technical comments, and positive remarks.

Preference

Preference was assessed after the participants had had a first look at the map and after they had worked with the visualization to solve the tasks.

5.3.4 Other variables

Demographics

Participants stated their gender, age group, and education status. They also rated their personal interest and experience in four fields on a Likert scale: migration, maps, data visualization, and online news stories.

5.3.5 Statistics

Statistical analysis was carried out using SPSS²⁶. Where means were compared, an independent or paired-samples *t*-test was carried out, following guidelines by Field (2013). Each result of a *t*-test is reported with mean *M* and standard error *SE*, as well as the mean difference with its 95% confidence interval *CI*. Significance values are reported for a significance of $p = 0.05$. The effect size is reported for all results, including non-significant ones. As a measurement of effect size, Cohen's *d* is reported, since it is less prone to distortion by differing sample size compared to *r* or η_p^2 .

According to Field (2013), for sample sizes > 30 normality can be assumed. Nevertheless, before running the *t*-test, each variable was tested and confirmed for normality using a one-sample Kolmogorov–Smirnov test.

In an effort to remove outliers, a standard deviation based rule was used, removing values more than 2.5 standard deviations from the mean. Only a small number of individual values were removed this way. An exception are the values of response time, as described in section 5.3.2.

²⁶IBM SPSS Statistics Version 21

6 Results

Based on the pool of people from which the participants were recruited, it must be assumed that a large part of the participants did not speak English as their first language. Of the 91 participants that completed the survey, 7 participants were removed from the sample because they stated having problems understanding some of the questions. They had given an agreement score of 4 on a five point Likert scale for the statement “I had trouble understanding some of the questions”. This leaves a total of 84 participants that are included in the sample.

In order to achieve a random selection for the groups (evaluation types), the participants were asked on which day of the month they were born. Participants who entered a number smaller or equal to 15 were assigned to the interactive visualization group, while the remaining participants were assigned to the static visualization group. This resulted in 54 participants in the interactive visualization group and 37 participants in the static visualization group. It is unclear what resulted in this slightly unbalanced outcome of the group assignment. Statistical data on birthdays in Switzerland suggests no such imbalance in the general population (fig. 6.1). A possible reason came up in an informal conversation with one of the participants after she had taken the survey, who stated that she had mistakenly given the number of the month of the year instead of the number of the day of the month. This could be a result of priming (Schacter and Buckner, 1998) to a common survey question. This effect could not be analyzed, but if many participants experienced the same effect of priming, this would result in a shift in the number of participants toward the interactive visualization group.

41 participants were male, while 43 were female. While there were roughly the same number of men and women in the overall sample, they were quite unevenly distributed between the visualization types, as fig. 6.2 shows. Age distribution is shown in fig. 6.3. While the age distribution is heavily imbalanced between the age groups (the many young people are a direct result of the recruitment methods), it is almost perfectly even across the visualization types. The distribution of highest education degrees is shown in figure 6.4. Most participants were students, also a result of the recruitment method. Fig. 6.5

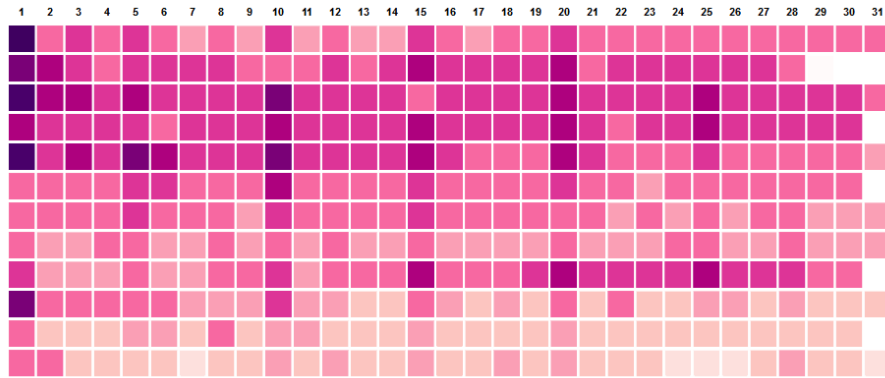


Figure 6.1: Frequency of birthdays in Switzerland, with darker colours indicating a higher frequency. The vertical axis shows the month of the year chronologically with January displayed on top, the horizontal axis shows the day of the month (Bauer, 2013)

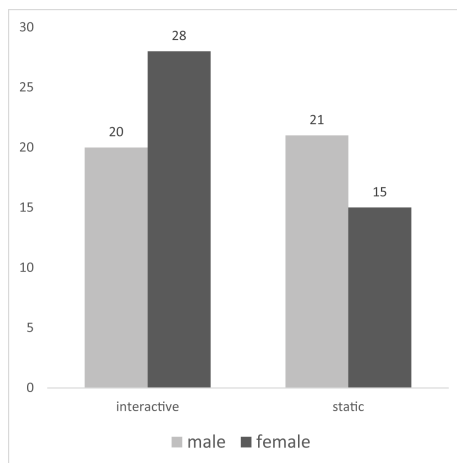


Figure 6.2: Gender distribution in the participant sample

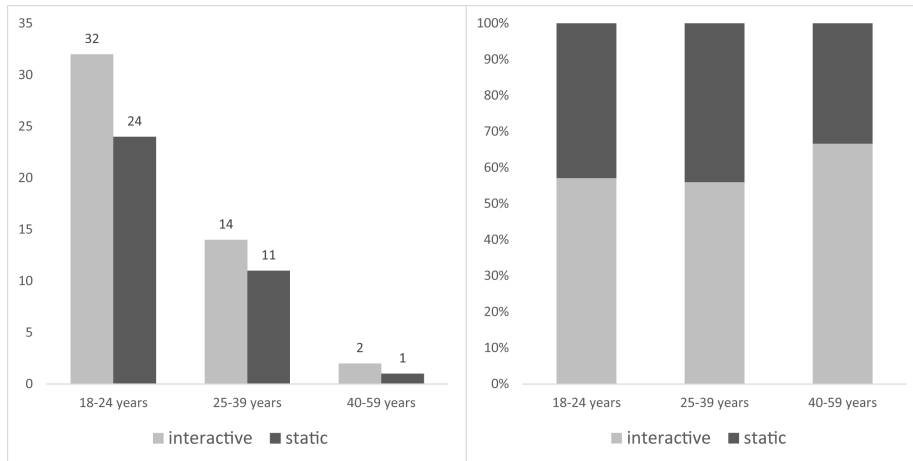


Figure 6.3: Age distribution in the participant sample (grouped and 100% stacked bar chart)

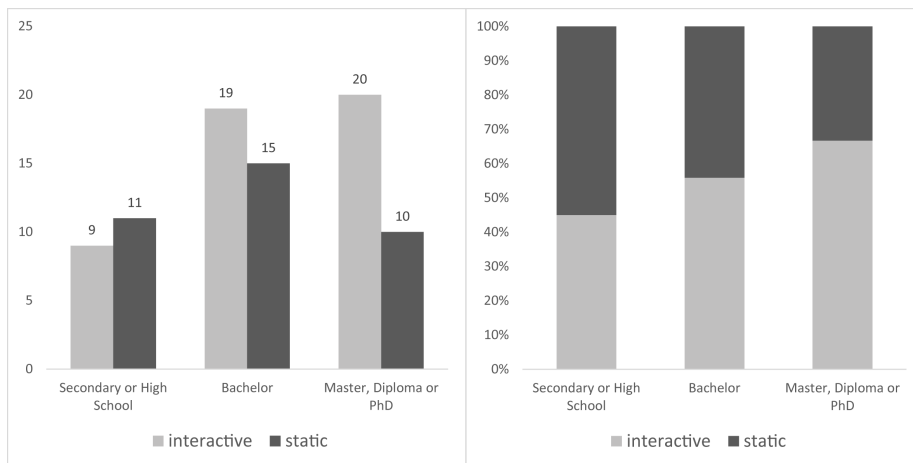


Figure 6.4: Distribution of highest education degrees in the participant sample (grouped and 100% stacked bar chart)

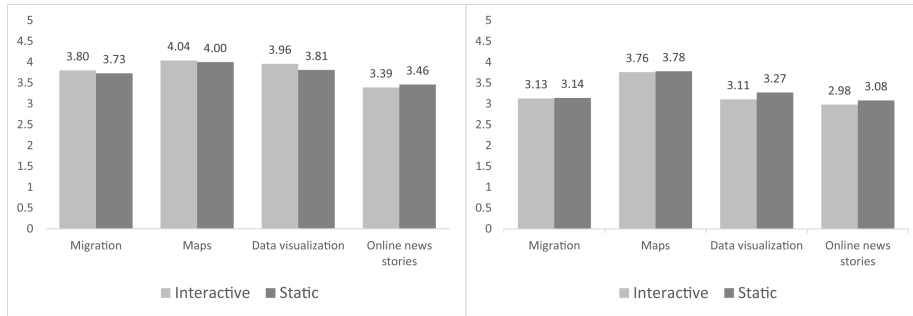


Figure 6.5: Self-declared interest (left) and experience (right) in four fields related to the study

shows self-declared interest and experience in four fields related to the study. There are no major differences between visualization types.

6.1 Research question 1: usability

6.1.1 Accuracy

Accuracy is measured in percent correct answers. There is a small difference in accuracy between the two visualization types. The mean accuracy (fig. 6.7A) for the interactive visualization type is $M = 79.52$, $SE = 1.77$; and it is $M = 82.00$, $SE = 1.84$ for the static visualization type. This difference of -2.48 , $CI[-0.63, 2.67]$, is not significant ($t(81) = -0.96$, $p = 0.34$). The effect size $d = -0.22$ is small.

Associate task type

Fig. 6.6 shows mean response scores for the questions of the task type associate.

6.1.2 Response time

Response time is measured in seconds. For the interactive visualization type, the mean (fig. 6.7B) is $M = 40.73$, $SE = 1.32$; for the static visualization type, the mean is $M = 40.95$, $SE = 1.56$. The mean difference is 0.22 , $CI[-3.85, 4.3]$ is not significant ($t(88) = 0.11$, $p = 0.91$). The effect size $d = 0.02$ is virtually inexistent.

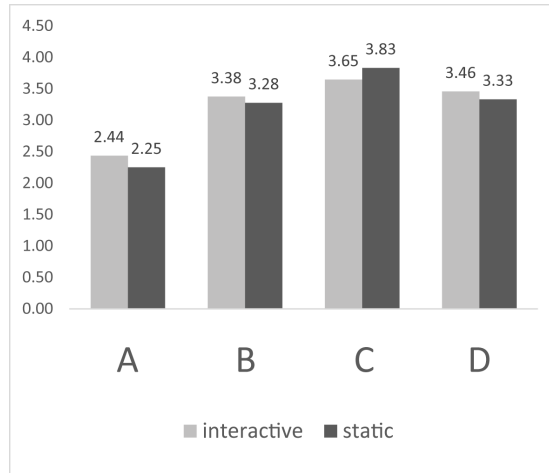


Figure 6.6: Associate task type: mean scores (Likert 1-5) for (A) The number of dead or missing migrants is a direct consequence of the total number of migrants. (B) The extent of the search operation areas had a strong effect on the number of dead or missing migrants. (C) The civil war in Syria was the main driver of migration to Europe across the Mediterranean Sea in 2015. (D) Most migrants that went to Lesbos came from Syria.

6.1.3 Confidence and ease

Confidence and ease are measured on a five point Likert scale. The mean confidence (fig. 6.7C) for the interactive visualization type is $M = 4.19$, $SE = 0.09$. The values for the static visualization type are $M = 4.26$, $SE = 0.07$. The mean difference is -0.06 , $CI[-0.31, 0.19]$. There is no significant effect ($t(82) = -0.50$, $p = 0.62$). The effect size $d = 0.14$ is small.

For the interactive visualization type, the mean ease (fig. 6.7D) is $M = 3.46$, $SE = 0.09$; and $M = 3.55$, $SE = 0.09$ for the static visualization type; with a mean difference of -0.10 , $CI[-0.35, 0.16]$. There is no significant effect ($t(88) = -0.75$, $p = 0.46$). The effect size $d = -0.18$ is small.

6.1.4 System usability scale

The interactive visualization type received an SUS score of 74.2, while the static visualization type was scored 72.0. According to Bangor et al. (2009), these are good scores.

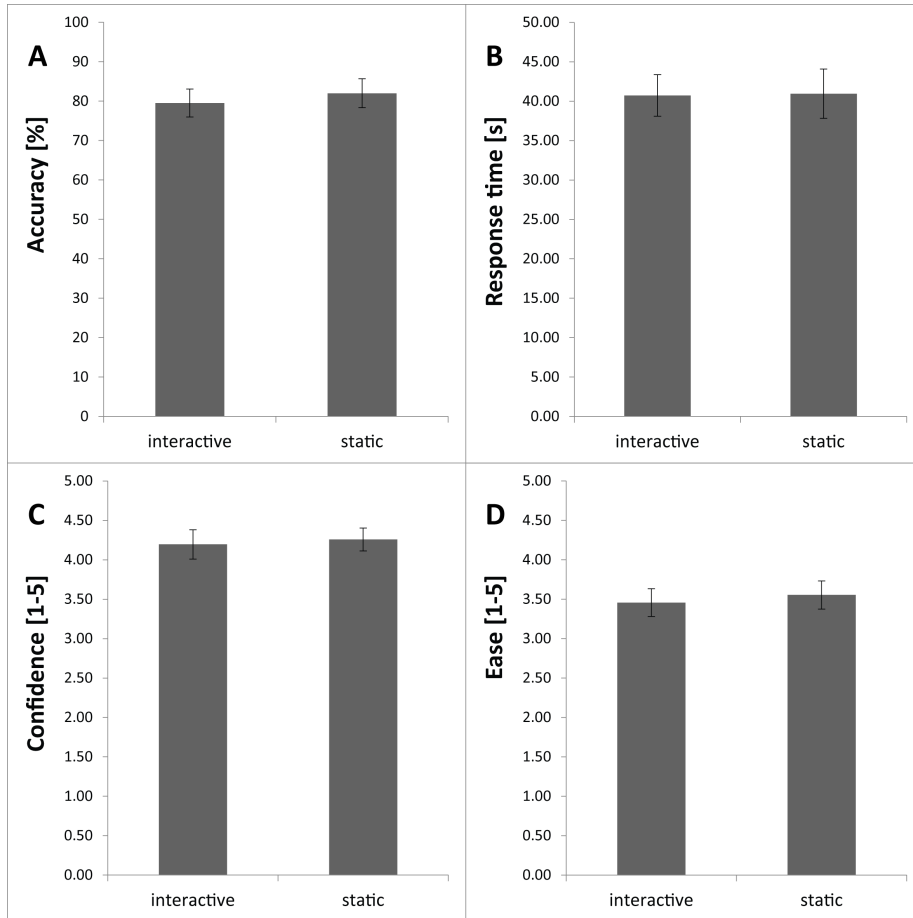


Figure 6.7: Mean and error bars ($\pm 2SE$) for (A) percent correct answers, (B) response time in seconds, and (C) confidence and (D) ease on a five point Likert scale

6.2 Research question 2: engagement

An independent sample t -test revealed no effect of visualization type on engagement after working with the visualization. The mean engagement (fig. 6.8A) for the interactive visualization type is $M = 2.64$, $SE = 0.08$; and $M = 2.59$, $SE = 0.08$ for the static visualization type. The mean difference is 0.05, $CI[-0.19, 0.29]$. It is not significant ($t(89) = 0.43$, $p = 0.67$).

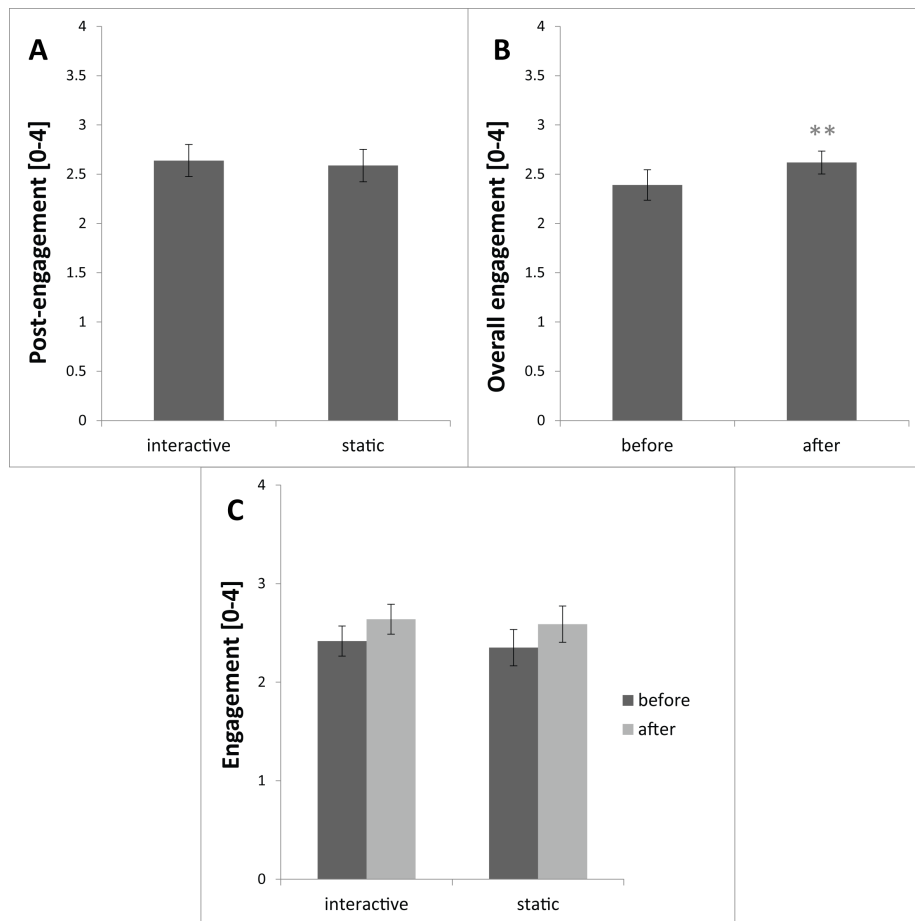


Figure 6.8: Mean and error bars ($\pm 2SE$) for (A) engagement after working with the visualizations, (B) engagement over both visualization types before and after working with the visualizations, and (C) a combined view

However, a paired-samples t -test reveals a significant difference across both visualization types between engagement scores before (pre) and after (post) working with the visualizations (fig. 6.8C). The pre engagement score has a mean of $M = 2.39$, $SE = 0.08$. The post engagement score has a mean of $M =$

2.62, $SE = 0.06$. The mean difference of -0.23, $CI[0.36, -0.10]$ is significant ($t(90) = -3.40$, $p = 0.001$). The effect size $d = -0.41$ is small to medium. Fig. 6.8C shows a combined view for visualization type and pre- and post-engagement,

For the proposed engagement scale, the score is similar for both versions of the visualization: 60.69 for the interactive version, and 60.00 for the static version.

6.2.1 Exploration

Participants made use of the opportunity to take notes and make comments, especially after they had seen the map for the first time. Users of the interactive version spent on average 188 seconds on the survey page where they could make comments. 77% of users of the interactive version gave at least one comment. Users of the static version spent on average 211 seconds on the comment page, with 81% of users making at least one comment.

The following table shows the number of individual statements per category:

	interactive	static
content	0.96	1.78
misunderstood	0.15	0.03
technical	0.54	0.30
positive	0.23	0.08
total	2.65	2.19

7 Discussion

7.1 Research question 1: usability

With the data collected in the user study, no statement could be made about the influence of interactivity on effectiveness. Results were non-significant for accuracy, response time, confidence, and ease. Differences between the visualization types for the dependant variables were so small that not even non-significant effects can be reported. These results suggest that the two map types that were used in the study are very similar with regard to their effectiveness and efficiency.

When the two visualization types were designed to meet informational equivalence, they were perhaps designed all too similarly. To recollect, Larkin and Simon (1987, 67) define informational equivalence as obtained “if all of the information in the one is also inferable from the other, and vice versa”. On the same occasion, they also give a definition for a more specific kind of informational equivalence. Computational equivalence is obtained “if they are informationally equivalent and, in addition, any inference that can be drawn easily and quickly from the information given explicitly in the one can also be drawn easily and quickly from the information given explicitly in the other, and vice versa”. It seems that the two visualization types are computationally equivalent. Hypothesis 1 was based partly on the fact that in interactive maps, part of the information is hidden at all times and it requires an action by the user, manipulating the visualization through the interface to reach the desired information (Roth, 2013). In the interactive map designed for this study, with just four time periods to select from, this interaction and manipulation is quite the opposite of complex. In fact, as stated earlier in section 5.1.3: in the static version of the map, users need to scroll to reach each multiple, which is a very similar degree of interaction to clicking a button, as is required in the interactive version.

The remaining interactive elements, i.e. the tooltips and visibility checkboxes, did not have the effect of making users perform worse with regard to effectivity and efficiency. Whether users used the visibility checkboxes at all when working with the maps was not recorded. The tooltips, even though they required a mouse movement to activate and took 1–2 seconds to load, did not slow the users down. The interactive effort it took to reach the hidden informa-

tion might have been compensated by the fact that once the information was obtained in the interactive version, it was exact. In the static version, a value had to be estimated using the legend as a reference. This led both to a slow response, because of the conscious effort required, and to inexact responses.

Even if unfortunately Hypothesis 1 can not be confirmed, the results show that the designed map works well in both visualization types. The SUS scores of 74.2 (interactive) and 72.0 (static) are good (Bangor et al., 2009). With an average confidence rating of over 4 out of 5 across all users and tasks, the same conclusion can be drawn that both maps work for the users.

Since the tasks of the type could not be analyzed statistically, they are just touched on briefly. The results are shown in fig. 6.6. The differences between the visualization types are not really relevant, but there are some findings of interest. Most people do not believe that higher migration numbers directly result in more casualties (fig:associateA, mean Likert scores 2.44/2.25). For the remaining associate questions, the mean Likert scores are neutral. 6.6B (mean Likert scores 3.38/3.28) is about the effect of search operation areas on deaths. This is an aspect that many people will probably have heard of for the first time. 6.6D (mean Likert scores 3.46/3.33) was sort of a trick question: it is not possible to say whether most migrants to Lesbos came from Syria (since the flow passes through the aggregated country). Users were not ready to make this bold statement and answered neutrally. The highest Likert scores (3.65/3.83) resulted from the connection between the Syrian civil war and overall migration numbers. This is surprising, since the statement is formulated in quite a bold way in that a) the majority of definitely migrants come from Europe and b) what caused them to come to Europe was definitely the civil war. The relatively high score (3.65/3.83) could be a product of previous knowledge and priming, and of impressively wide flow bars in the visualization.

7.2 Research question 2: engagement

No effect of the visualization type on engagement could be found. Hypothesis 2 could not be confirmed. The engagement score after working with the visualization is practically the same for both visualization types. It seems likely that the reason is the same as the one suggested in 7.1: the two visualizations are just too similar. The fact that no differences could be found is even more surprising than for research question 1. While effectiveness and efficiency are very technical characteristics of a visualization, engagement is a very subjective matter. That both visualization types should have the same effect on the way

the users feel is remarkable. Of course, this conclusion is not at all confirmed statistically — yet given that both visualization types score almost the same in all of the variables that were measured, including engagement, this is a feasible assumption. Specifically with regard to engagement, it is not the case that there was just no change recorded. Engagement analyzed over both visualization types changed significantly, getting higher after participants worked with the visualizations. This is a further indicator that the visualizations are quite well designed — not only do they score well in usability and confidence, but users seem to enjoy them as well.

Exploration. Users spent an average of more than three minutes to write comments and statements. While not clearly belonging into the engagement section of the discussion, some of the most common comments are summarized here.

Content. Content-related statements were the most common statements. Users of the static map took made more statements of this type than users of the interactive type (on average 1.78 static, 0.96 interactive). This was not analyzed statistically, but it is a further cue that the interactive visualization did not in fact engage users more. One commonly named process was the shift from the Libya-Italy route to the Turkey-Greece route. Occasionally, it was noted that the Libya-Italy route seems to follow a seasonal cycle, while the Turkey-Greece route just keeps on growing. This is of course due to the rigid calendar year time frame that does not cover the seasonal cycle of 2015, as stated in sectionsec:meth.vis. The pattern of incident was commonly spotted: Few large incidents near Italy, and many smaller incidents near Greece. This indicates that the goal of letting users discover the pattern was successful, even if the incident bubbles overlap dramatically. A connection between the search operation areas or the total number of migrants with the number of incidents was also sometimes made.

Misunderstood. By far the statement that was made the most that has to be called a misunderstanding is that there are no rescue missions (sometimes noted with outrage) in the Aegean Sea. The fact is that the Greece is responsible for its sea borders. There is no EU mission (European Commission, 2016), but that does not mean there is no coast guard.

Technical. Technical comments were definitely most common for the interactive visualization type. Many only discovered the hover/tooltip feature by chance. The feature was described on the side of the map (fig. 5.7), but many users apparently did not realize this. Apparently what Andrienko et al. (2002) and Çöltekin et al. (2009) found, that users do not read instructions, is true

even for short notes. Many users also missed a legend. In hindsight, a legend should have been included. The rationale for not including one was that all of the necessary information about the map elements can be retrieved through tooltips. One example that might have been cleared up by a legend is that people were sometimes confused about the two different kinds of circles. The learning-by-doing approach for the interactive map as outlined in section 5.2 worked well overall, but left some features unclear for some users. Technical issues that were mentioned by users of the static map are a cluttered layout (overlapping bubbles and too many labels on the map) and tiny font.

Positive. Positive comments mentioned the clear design and the excitement of seeing data visualized in a novel way.

7.3 Limitations and outlook

7.3.1 User study

Given that the study produced hardly any significant output and none of the research hypotheses could be confirmed, there are quite some limitations. It is difficult to say however how the study can be adapted to produce significant output. Since the degree of interactivity of the interactive map was quite limited, the most likely conclusion seems to be that the two map types were just too similar. The research questions should be analyzed in future research with a more sound visualization and study design.

The engagement scale that was proposed in this thesis as a quick and easy instrument to measure the effect of a system on its users' engagement should be tested in future research.

7.3.2 Visualization

The visualizations created for the user study, specifically the interactive version, have the potential to be developed further. The user should be given more control about the time variable, e.g. setting the size of the time frame and shifting it around continuously. Also, the clutter of the overlapping bubbles needs to be addressed. The collision detection forces that are a feature of D3 could be a possibility, clustering bubbles up alongside each other. Instead of displaying incidents as bubble, the approach by Pellegrini and Mauri (2016) is also a possibility. Since both these approaches disrupt the geographic pattern of the incidents, the user should be given the possibility to switch between incident display types. Besides the implementation of features to enhance the interactive

experience for the user, the data goal of providing more up-to-date data could be addressed, perhaps obtaining data in real time or at set intervals from the IOM and UNHCR data sources.

8 Conclusion

In this thesis, the influence of interactivity of a visualization, specifically a map, on usability and user engagement was tested. It was hypothesised that the interactive version would be less efficient and effective than a static visualization, while at the same time engaging users more to explore the data. None of the hypotheses could be confirmed. It is assumed that the degree of interactivity in the interactive version was too small to make a difference. One output of this thesis is however an interactive map that scores well for usability and confidence. This visualization can further be refined to allow more freedom of interaction and include a wider range of information.

Bibliography

- Andrienko, G., Andrienko, N., Kopanakis, I., Ligtenberg, A., and Wrobel, S. (2008a). Visual Analytics Methods for Movement Data. In Giannotti, F. and Pedreschi, D., editors, *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*, pages 375–410. Springer, Berlin, Heidelberg.
- Andrienko, N., Andrienko, G., Pelekis, N., and Spaccapietra, S. (2008b). Basic Concepts of Movement Data. In Giannotti, F. and Pedreschi, D., editors, *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*, pages 15–38. Springer, Berlin, Heidelberg.
- Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J., and Kretschmer, U. (2002). Testing the Usability of Interactive Maps in CommonGIS. *Cartography and Geographic Information Science*, 29(4):325–342.
- Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123.
- Bauer, D. (2013). *Häufigste Geburtstage in der Schweiz*. Tageswoche, 15.8.2013. Accessed from: http://www.tageswoche.ch/de/2013_33/leben/569747/mit-wie-vielen-anderen-in-der-schweiz-teilen-sie-den-geburtstag.htm (last accessed 20.1.2017).
- BI Graphics (2015). *Europe's Refugee Crisis*. Business Insider. Accessed from: <http://static.businessinsider.com/image/55f2f958bd86ef1c008b9ab9/image.jpg> (last accessed 23.1.2017).
- Bostock, M., Ogievetsky, V., and Heer, J. (2011). D3: Data-Driven Documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309.
- Boyandin, I., Bertini, E., and Lalanne, D. (2010). Using Flow Maps to Explore Migrations Over Time. *Geospatial Visual Analytics Workshop in conjunction with The 13th AGILE International Conference on Geographic Information Science*, page 4.
- Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7.

- Coddington, M. (2015). Clarifying Journalism’s Quantitative Turn: A topology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3):331–348.
- Çöltekin, A., Heil, B., Garlandini, S., and Fabrikant, S. I. (2009). Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-movement Analysis. *Cartography and Geographic Information Science*, 36(1):5–17.
- Doantam, P., Xiao, L., Yeh, R., Hanrahan, P., and Winograd, T. (2005). Flow Map Layout. In *IEEE Symposium on Information Visualization (InfoVis)*, pages 219–224.
- European Commission (2016). EU Operations in the Mediterranean Sea.
- European Union Committee (2016). *Operation Sophia, the EU’s naval mission in the Mediterranean: an impossible challenge*. HL 144.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications.
- Friendly, M. (2002). Visions and Re-Visions of Charles Joseph Minard. *Journal of Educational and Behavioral Statistics*, 27(1):31–51.
- Gray, J., Bounegru, L., and Chambers, L., editors (2012). *The Data Journalism Handbook: How Journalists Can Use Data to Improve the News*. O’Reilly Media, Sebastopol.
- Hegarty, M., Smallman, H. S., Stull, A. T., and Canham, M. S. (2009). Naïve Cartography: How Intuitions about Display Configuration Can Hurt Performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3):171–186.
- Heller, C. and Pezzani, L. (2016). Ebbing And Flowing: The EU’s Shifting Practices Of (Non-)Assistance And Bordering In A Time Of Crisis. *Near Futures Online*, 1.
- Howard, A. B. (2014). *The Art and Science of Data-Driven Journalism*. Tow Center for Digital Journalism, New York.
- IOM Missing Migrants Project (2016). Incidents on the Mediterranean Sea. (received by e-mail 2.11.2016).

- Jenny, B., Jenny, H., and Räber, S. (2008). Map design for the Internet. In *International Perspectives on Maps and the Internet*, pages 31–48. Springer, Berlin, Heidelberg.
- Jenny, B., Stephen, D. M., Muehlenhaus, I., Marston, B. E., Sharma, R., Zhangd, E., and Jenny, H. (2016). Design principles for origin-destination flow maps. *Cartography and Geographic Information Science*.
- Knapp, L. (1995). A Task Analysis Approach to the Visualization of Geographic Data. In Nyerges, T. L., Karwan, M., Laurini, R., and Egenhofer, M. J., editors, *Cognitive Aspects of Human Computer Interaction for Geographic Information Systems*, pages 355–371. Springer Netherlands, Dordrecht.
- Larkin, J. H. and Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11(1):65–100.
- Le Monde Diplomatique (2016). *Festung Europa 2016*. Accessed from: https://monde-diplomatique.de/images/Festung_Europa_2016.jpg (last accessed 23.1.2017).
- Maclean, M. (2015). *D3 Tips and Tricks v3.x*. Leanpub. Accessed from: <https://leanpub.com/D3-Tips-and-Tricks> (last accessed 22.1.2017).
- Minard, C. J. (1865). *Carte figurative et approximative des quantités de vin français exportés par mer en 1864*. École Nationale des Ponts et Chaussées. Accessed from: https://upload.wikimedia.org/wikipedia/commons/6/62/Minard%E2%80%99s_map_of_French_wine_exports_for_1864.jpg (last accessed 22.1.2017).
- Minard, C. J. (1869). *Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813*. Regnier et Dourdet. Accessed from: <https://upload.wikimedia.org/wikipedia/commons/2/29/Minard.png> (last accessed 22.1.2017).
- Murray, S. (2013). *Interactive Data Visualization for the Web*. O'Reilly Media, Sebastopol. Accessed from: <http://chimera.labs.oreilly.com/books/1230000000345/index.html> (last accessed 22.1.2017).
- Natural Earth (2012). *1:50m Cultural Vectors: Admin 0 – Countries*. http://www.naturalearthdata.com/download/50m/cultural/ne_50m_admin_0_countries.zip (last accessed 18.7.2016).

- O'Brien, H. L. and Toms, E. G. (2010). The Development and Evaluation of a Survey to Measure User Engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69.
- O'Brien, H. L. and Toms, E. G. (2013). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing and Management*, 49(5):1092–1107.
- Pearce, R. (2016). *Flow map of coal exports around the world showing the top exporters*. Carbon Brief. Accessed from: <https://www.carbonbrief.org/wp-content/uploads/2016/05/coal-trade-around-the-world-1.jpg> (last accessed 22.1.2017).
- Pellegrini, V. and Mauri, M. (2016). *La mappa dei migranti scomparsi*. Corriere della Sera. Accessed from: https://mir-s3-cdn-cf.behance.net/project_modules/max_3840/954d9c34680727.56d98271cd2e7.jpg and https://mir-s3-cdn-cf.behance.net/project_modules/1400/ca1de834680727.56d9cb3ba294b.png (last accessed 22.1.2017).
- Picard, R. G. (2011). Certainty and uncertainty of change in journalism. In Lowrey, W. and Gade, P. J., editors, *Changing the News: The Forces Shaping Journalism in Uncertain Times*, pages vii–ix. Routledge, London.
- Picard, R. G. (2015). *Journalists' Perceptions of the Future of Journalistic Work*. Reuters Institute for the Study of Journalism, Oxford.
- Poplin, A. (2015). How user-friendly are online interactive maps? Survey based on experiments with heterogeneous users. *Cartography and Geographic Information Science*, 42(4):358–376.
- Roth, R. E. (2013). Interactive maps: What we know and what we need to know. *Journal of Spatial Information Science*, (6):59–115.
- Saarinen, V. and Ojala, J. (2016). *The flow towards Europe*. Lucify. Accessed from: <http://www.lucify.com/the-flow-towards-europe> (last accessed 23.1.2017).
- Sander, N., Abel, G., Bauer, R., and Schmidt, J. (2014). Visualising Migration Flow Data with Circular Plots. *Vienna Institute of Demography: Working Papers*, 02.
- Sankey, M. H. P. R. (1898). Introductory note on the thermal efficiency of steam-engines. *Minutes of Proceedings of the Institution of Civil Engineers*, 134:278–283.

- Schacter, D. L. and Buckner, R. L. (1998). Priming and the Brain. *Neuron*, 20:185–195.
- Schaufeli, W. B., Bakker, A. B., and Salanova, M. (2006). The Measurement of Short Questionnaire: A Cross-National Study. *Educational and Psychological Measurement*, 66(4):701–716.
- Schmidt, M. (2008a). The Sankey diagram in energy and material flow management: Part I: History. *Journal of Industrial Ecology*, 12(1):82–94.
- Schmidt, M. (2008b). The Sankey diagram in energy and material flow management: Part II: Methodology and current applications. *Journal of Industrial Ecology*, 12(2):173–185.
- Segel, E. and Heer, J. (2010). Narrative visualization: telling stories with data. In *IEEE transactions on visualization and computer graphics*, volume 16, pages 1139–1148.
- Singer, J. B. (2011). Journalism and Digital Technologies. In Lowrey, W. and Gade, P. J., editors, *Changing the News: The Forces Shaping Journalism in Uncertain Times*, pages 213–229. Routledge, London.
- Slocum, T. A., McMaster, R. B., Kessler, F. C., and Howard, H. H. (2010). *Thematic cartography and geovisualization*. Pearson Prentice Hall, Upper Saddle River, NJ, third edition.
- UNHCR (2016a). *Mediterranean Sea Arrivals: Regional Overview (Screenshot) 31 Dec*. <http://data.unhcr.org/mediterranean/download.php?id=461> (last accessed 5.1.2017).
- UNHCR (2016b). *Monthly Arrivals by Nationality to Greece, Italy and Spain – Jan 2015 to Jan 2016*. <http://data.unhcr.org/mediterranean/download.php?id=864> (last accessed 19.11.2016).
- UNHCR (2016c). *Monthly Sea Arrivals to Italy and Malta – Jan–Dec2015*. <http://data.unhcr.org/mediterranean/download.php?id=450> (last accessed 12.12.2016).
- Zambrano, R. N. and Engelhardt, Y. (2008). Diagrams for the masses: Raising public awareness – from Neurath to Gapminder and Google Earth. In *Proceedings of the 5th international conference on Diagrammatic Representation and Inference*, pages 282–292.

A D3 code for the interactive visualization

Note: the data source files used in this code are not provided with this thesis. On its own, the code is not executable. The code itself is commented throughout to facilitate its understanding.

```

1 <DOCTYPE html>
2 <html lang="en">
3 <head>
4 <meta charset="utf-8">
5 <title>Migration across the Mediterranean Sea 2015</title>
6 <script type="text/javascript" src="../d3/d3.js"></script>
7 <style type="text/css">
8   h1 {font-family: sans-serif;
9     font-size: 200%;}
10  text {
11    font-size: 5.5px;
12    font-family: "Helvetica Neue", Helvetica, Arial, sans-serif;
13  }
14  .overallContainer {
15    display: inline-block;
16    position: relative;
17    width: 100%;
18    padding-bottom: 40%; /* aspect ratio */
19    vertical-align: top;
20    overflow: hidden;
21  }
22  .mapContainer {
23    display: inline-block;
24    position: absolute;
25    width: 80%;
26    height: 6100%;
27    left: 0;
28    background-color: lightsteelblue;
29  }
30  .legendContainer {
31    display: inline-block;
32    position: absolute;
33    height: 100%;
34    width: 19%;
35    left: 80%;
36  }
37 </style>
38 </head>
39 <body>
40 <h1>Migration across the Mediterranean Sea 2015</h1>
41 <script type="text/javascript">
42
43 // The code for this visualization is based on a sample code for a choropleth map
44 //   by Scott Murray (2013) in his book "Interactive Data Visualization for the Web
45 //   ". The specific code can be found here:
46 // https://github.com/alignedleft/d3-book/blob/master/chapter_12/05_choropleth.html
47
48 // The svg is made responsive to dynamically adapt to window size.

```

```

47 // Adapted from an answer by user cminatti on stackoverflow:
48 // http://stackoverflow.com/questions/16265123/resize-svg-when-window-is-resized-in
    -d3-js
49 var overallContainer = d3.select("body")
50   .append("div")
51   .classed("overallContainer", true);
52
53 // Create SVG element for the map
54 var svg = overallContainer
55   .append("svg")
56   .attr("preserveAspectRatio", "xMinYMin meet")
57   .attr("viewBox", "0 0 600 400")
58   .classed("mapContainer", true);
59
60 // Create groups to place the map elements in. This ensures correct drawing order.
61 svg.append("g").attr("class", "triton");
62 svg.append("g").attr("class", "background");
63 svg.append("g").attr("class", "flowGroup");
64 svg.select(".flowGroup").append("g").attr("class", "destGroupGhost");
65 svg.select(".flowGroup").append("g").attr("class", "destGroup");
66 svg.select(".flowGroup").append("g").attr("class", "countryGroupGhost");
67 svg.select(".flowGroup").append("g").attr("class", "countryGroup");
68 svg.select(".flowGroup").append("g").attr("class", "origGroupGhost");
69 svg.select(".flowGroup").append("g").attr("class", "origGroup");
70 svg.append("g").attr("class", "circleGroup");
71
72 // Scale factors for flows and incident bubbles
73 var flowScale = 12000;
74 var incScale = 1.5;
75
76 // Variables for the data source files
77 var destSource = "arrivalsDest_1.csv";
78 var origSource = "arrivalsOrig_1.csv";
79 var incSource = "inc15_1.csv";
80 var triSource = "triton.json";
81
82 // The functions to create the flows and search areas need to be called initially
83 createFlows();
84 createTriton();
85
86 // Duration of the change transitions
87 var transitionDur = 800;
88
89 // Define map projection
90 var projection = d3.geoMercator()
91   .center([33, 31]).scale([1050]);
92
93 // Define path generator

```

```

94 var path = d3.geoPath()
95   .projection(projection);
96
97 // Create svg element for the legend/buttons
98 var legend = overallContainer
99   .append("svg")
100   .attr("preserveAspectRatio", "xMinYMin meet")
101   .attr("viewBox", "0 0 100 20")
102   .classed("legendContainer", true);
103 var buttonGroup = legend.append("g").attr("class", "buttonGroup");
104
105 // Create buttons to select the time frame to be displayed
106
107 var
108   buttonH = 12,
109   buttonW = 40,
110   buttonR = 2,
111   buttonDist = 7,
112   buttonX = buttonH/4,
113   buttonY = buttonX,
114   buttonPressed = "darkgrey"
115   buttonIdle = "gainsboro",
116   checked = "black",
117   unchecked = "white",
118   strokeW = 0.5
119   checkR = 1;
120
121 // The buttons are created in a loop function, using above variables as basic
    building blocks
122 var button = [];
123 for (var i=0; i<4; i++)
124 {
125   button[i] = buttonGroup.append("g").attr("class", "button"+(i+1)+"group")
126   // Define the buttons' dimensions and place them on the canvas
127   button[i]
128     .append("rect")
129     .attr("id", "brect"+(i+1))
130     .attr("width", buttonW)
131     .attr("height", buttonH)
132     .attr("fill", buttonIdle)
133     .attr("x", buttonX)
134     .attr("y", buttonY+i*(buttonH+buttonDist))
135     .attr("rx", buttonR)
136     .attr("ry", buttonR)
137     .classed("quarterButton", true);
138   // The first button should appear as pressed when the map is first loaded
139   button[0].selectAll("rect").attr("fill", buttonPressed);
140   // Add labels to all the buttons

```

```

141     button[i]
142         .append("text")
143         .text(i==0 ? "Jan-Mar" : i==1 ? "Apr-Jun" : i==2 ? "Jul-Sep" : "Oct-Dec")
144         .attr("x", buttonX+buttonW/2)
145         .attr("y", (buttonY+buttonH/2)+i*(buttonH+buttonDist))
146         .attr("text-anchor", "middle")
147         .attr("alignment-baseline", "middle")
148         .attr("pointer-events", "none");
149     };
150
151     // Implement the buttons' functionality (on click)
152     // This first function updates the buttons' appearance to match the correct status
153     // (pressed or idle)
154     function buttonsUpdate(b) {
155         d3.select(".buttonGroup").selectAll("rect").attr("fill", buttonIdle)
156         d3.select(b).attr("fill", buttonPressed)
157     };
158     // When button 1 is clicked, the data source is updated to the first time frame and
159     // all map elements are updated.
160     d3.select("#brect1").on("click", function() {
161         incSource = "inc15_1.csv";
162         destSource = "arrivalsDest_1.csv";
163         origSource = "arrivalsOrig_1.csv";
164         triSource = "triton.json";
165         buttonsUpdate("#brect1");
166         updateAll();
167     });
168     // When button 2 is clicked, the data source is updated to the second time frame,
169     // etc.
170     d3.select("#brect2").on("click", function() {
171         incSource = "inc15_2.csv";
172         destSource = "arrivalsDest_2.csv";
173         origSource = "arrivalsOrig_2.csv";
174         triSource = "triton.json";
175         buttonsUpdate("#brect2");
176         updateAll();
177     });
178     d3.select("#brect3").on("click", function() {
179         incSource = "inc15_3.csv";
180         destSource = "arrivalsDest_3.csv";
181         origSource = "arrivalsOrig_3.csv";
182         triSource = "sophia.json";
183         buttonsUpdate("#brect3");
184         updateAll();
185     });
186     d3.select("#brect4").on("click", function() {
187         incSource = "inc15_4.csv";
188         destSource = "arrivalsDest_4.csv";

```

```

186     origSource = "arrivalsOrig_4.csv";
187     triSource = "sophia.json";
188     buttonsUpdate("#brect4");
189     updateAll();
190 });
191
192 // Create checkboxes that show and hide elements.
193 // For each checkbox, a binary variable is prepared to control whether it is ticked
194 // or not. Their values are initially set to true.
194 var showFlow = true;
195 var showInc = true;
196 var showTri = true;
197 // The boxes are created in a loop function, using the same building block
198 // variables as the buttons.
198 var checkBox = []
199 for (var i=0; i<3; i++) {
200     checkBox[i] = buttonGroup.append("g").attr("class", "checkBox"+(i+1))
201     // Define the boxes' dimensions and place them on the canvas
202     checkBox[i]
203         .append("rect")
204         .attr("id", "crect"+(i+1))
205         .attr("width", buttonH)
206         .attr("height", buttonH)
207         .attr("stroke", "black")
208         .attr("stroke-width", strokeW)
209         .attr("fill", buttonIdle)
210         .attr("x", buttonX)
211         .attr("y", buttonY + (button.length+i)*(buttonH+buttonDist))
212         .attr("rx", checkR)
213         .attr("ry", checkR);
214     // Place ticks inside all of the boxes
215     checkBox[i]
216         .append("path")
217         // The lines of the ticks are created using the SVG path mini-language
218         .attr("d",
219             "M"+buttonH/6+", "+buttonH/2+
220             "L"+buttonH/2+", "+buttonH*5/6+
221             "L"+buttonH*5/6+", "+buttonH/6
222         )
223         .attr("stroke-width", strokeW)
224         .attr("stroke", "black")
225         .attr("fill", "none")
226         .attr(
227             "transform", "translate(" +
228                 buttonX + ", " +
229                 (buttonY+(button.length+i)*(buttonH+buttonDist)) +
230                 ")"
231         )

```



```

232     .attr("alignment-baseline", "middle")
233     .attr("pointer-events", "none");
234 // Add labels to the checkboxes
235 checkBox[i].append("text")
236     .text(i==0 ?
237         "Show migration flows" : i==1 ?
238         "Show incidents" :
239         "Show search operation areas")
240     .attr("x", 2*buttonX + buttonH)
241     .attr("y", (buttonY+buttonH/2) + (button.length+i)*(buttonH+buttonDist))
242     .attr("text-anchor", "left")
243     .attr("alignment-baseline", "middle")
244     .attr("pointer-events", "none");
245 };
246
247 // Implement the checkboxes' functionality (on click)
248 // When a checkbox is clicked, change its binary variable
249 d3.select("#crect1")
250     .on("click", function() {;
251         if (showFlow) {showFlow = false} else {showFlow = true};
252         updateAll();
253         console.log(showFlow);
254     });
255 d3.select("#crect2")
256     .on("click", function() {;
257         if (showInc) {showInc = false} else {showInc = true};
258         updateAll();
259     });
260 d3.select("#crect3")
261     .on("click", function() {;
262         if (showTri) {showTri = false} else {showTri = true};
263         updateAll();
264     });
265
266 // This function sets the opacity of the ticks in all checkboxes to 1
267 function checkMarkReset() {
268     d3.select(".buttonGroup").selectAll("path").attr("opacity", 1)
269 }
270 // Functions to display or remove map elements
271 function displayPaths() {
272     // If the binary variable for the flow path checkbox is false, remove all flow
273     // paths and set the opacity of the tick in the checkbox to 0
274     if (!showFlow) {
275         d3.select(".flowGroup").selectAll("path")
276             .remove(),
277         d3.select(".checkBox1").selectAll("path").attr("opacity", 0);
278     }
279 };

```

```

279 function displayInc() {
280     // If the binary variable for the incident checkbox is false, remove all incident
        circles and set the opacity of the tick in the checkbox to 0
281     if (!showInc) {
282         d3.select(".circleGroup").selectAll("circle")
283             .transition()
284             .duration(transitionDur)
285             .ease(d3.easeCubic)
286             .attr("r", 0)
287             .remove(),
288         d3.select(".checkBox2").selectAll("path").attr("opacity", 0);
289     }
290 };
291 function displayTri() {
292     // If the binary variable for the search areas is false, remove all search area
        paths and set the opacity of the tick in the checkbox to 0
293     if (!showTri) {
294         d3.select(".triton").selectAll("path")
295             .remove(),
296         d3.select(".checkBox3").selectAll("path").attr("opacity", 0);
297     }
298 }
299
300 // This function updates all map elements
301 function updateAll() {
302     // First, all map elements are created and the ticks in all checkboxes are made
        visible.
303     createIncidents();
304     createIncidents();
305     createTriton();
306     createFlows();
307     checkMarkReset()
308     // Then, the display-functions are called, removing map elements and hiding ticks
        in the checkboxes according to the state of the binary checkbox variables.
309     displayPaths();
310     displayInc();
311     displayTri();
312 }
313
314 // Add info text about hover functionality
315 var infoGroup = d3.select(".legendContainer").append("g").attr("class", "info")
316 var info = [];
317 // The lines for the text are created in a loop function
318 for (var i=0; i<2; i++){
319     info[i] = d3.select(".info").append("g")
320         // Add text...
321         .append("text")
322         .text(i == 0 ?

```

```

323     "Hover over map elements to get" :
324     "more information.")
325     // ...and place it on the canvas
326     .attr("x", buttonX)
327     .attr("y",
328         (buttonY+buttonDist+(i*0.5)*buttonH) +
329         (button.length+checkBox.length)*(buttonH+buttonDist)
330     )
331     .attr("text-anchor", "left")
332     .attr("alignment-baseline", "middle")
333     .attr("pointer-events", "none");
334 };
335
336 // Add info text about sources
337 var sourceGroup = d3.select(".legendContainer").append("g").attr("class", "info")
338 var source = [];
339 // The lines for the text are created in a loop function
340 for (var i=0, font=4; i<6; i++){
341     source[i] = d3.select(".info").append("g")
342         // Add text...
343         .append("text")
344         .text(i == 0 ?
345             "UK parliament (search operation areas)": i==1 ?
346             "Natural Earth (country geometries)": i==2 ?
347             "IOM Missing Migrants Project (incidents)": i==3 ?
348             "UNHCR (migration flows)": i==4 ?
349             "":
350             "Sources")
351     // ...and place it on the canvas
352     .style("font-size", font)
353     .attr("x", buttonX)
354     .attr("y", 207-(i*(font+1)))
355     .attr("text-anchor", "left")
356     .attr("alignment-baseline", "middle")
357     .attr("pointer-events", "none");
358 };
359
360 // Helper function to add spaces as thousand separator
361 // Adapted from an answer by user Tim Goodman on stackoverflow:
362 // http://stackoverflow.com/questions/3753483/javascript-thousand-separator-string-format
363 function addSpaces(nStr) {
364     if (nStr >= 10000) {
365         nStr += ',';
366         x = nStr.split(',');
367         x1 = x[0];
368         x2 = x.length > 1 ? ',' + x[1] : '';
369         var rgx = /(\d+)(\d{3})/;

```

```

370     while (rgx.test(x1)) {
371         x1 = x1.replace(rgx, '$1' + ' ' + '$2');
372     }
373     return x1 + x2;
374 }
375 else {
376     return nStr
377 }
378 }
379 // Load Italy and Greece
380 d3.json("itgr.json", function(itgr) {
381     // Bind data and create one path per GeoJSON feature
382     svg.select(".background").append("g").selectAll("path")
383         .data(itgr.features)
384         .enter()
385         .append("path")
386         .attr("d", path)
387         .attr("fill", "gainsboro")
388     // On mouseover, change fill
389     .on("mouseover", function() {
390         d3.select(this)
391             .transition()
392             .duration(400)
393             .ease(d3.easeLinear)
394             .attr("fill", "grey")
395     })
396     // On mouseout, go back to original fill
397     .on("mouseout", function() {
398         d3.select(this)
399             .transition()
400             .duration(400)
401             .ease(d3.easeLinear)
402             .attr("fill", "gainsboro")
403     })
404     // Add the name of the country as a tooltip
405     .append("title")
406     .text(function (d) {
407         return(d.properties.sovereignt));
408 });
409
410 // Load Europe
411 d3.json("europe.json", function(eur) {
412
413     //Bind data and create one path per GeoJSON feature
414     svg.select(".background").append("g").selectAll("path")
415         .data(eur.features)
416         .enter()
417         .append("path")

```

```

418     .attr("d", path)
419     .attr("fill", "gainsboro");
420 });
421
422 // Load search operation areas
423 function createTriton() {
424     d3.json(triSource, function(trit) {
425         var tritonGroup = d3.select(".triton").selectAll("path")
426         var triton = tritonGroup.data(trit.features);
427         triton.enter()
428             .append("path")
429         triton
430             .attr("opacity", 0.05)
431             // dashed fill
432             .attr("fill", "url(#diagonal-stripe-3)")
433             .attr("d", path)
434             // On mouseover, change opacity to 1
435         triton.on("mouseover", function() {
436             d3.select(this)
437                 .transition()
438                 .duration(400)
439                 .ease(d3.easeLinear)
440                 .attr("opacity", 0.5)
441         })
442         // On mouseout, return opacity to 0.05
443         .on("mouseout", function() {
444             d3.select(this)
445                 .transition()
446                 .duration(400)
447                 .ease(d3.easeLinear)
448                 .attr("opacity", 0.05)
449         })
450         // Add an info text about the search operation as a tooltip, using the correct
451         // text based on the current search operation displayed
452         .append("title");
453         triton.select("title").text(
454             triSource == "triton.json" ? tritonTooltip : sophiaTooltip)
455     });
456     displayTri();
457 }
458
459 // These are the info texts about the search operations that are used as a tooltip
460 var tritonTooltip =
461     "Operation Triton (October 2014 - May 2015) \n\n"+
462     "After two shipwrecks in October 2013, which left a total of over 500 migrants
463     dead, \n"+
464     "the Italian authorities started the Operation Mare Nostrum, which was conducted
465     by \n"+

```

```

463 "the Italian Navy. The goal of Mare Nostrum was to save migrants from dying in
    inci-\n"+
464 "dents on the Mediterranean Sea. \n\n"+
465 "The Operation Triton replaced Mare Nostrum in October 2014. Triton is conducted
    \n"+
466 "by Frontex (the EU's border security agency). The main goal of Triton is to
    secure \n"+
467 "the European borders against illegal immigration. Compared to Mare Nostrum, the
    \n"+
468 "operational area of Triton was reduced and limited to areas near the Italian
    coast."
469
470 var sophiaTooltip =
471 "Operation Sophia (May 2015 - present) \n\n"+
472 "In April 2015, over 1000 migrants died in two shipwrecks on the Mediterranean
    Sea. \n"+
473 "These incidents lead to an expansion of the Operation Triton in May 2016 to
    include \n"+
474 "a wider area up to close to the Libyan coast. The operation was later renamed
    Sophia, \n"+
475 "after a baby who was born on a military vessel in the operation."
476
477 // Incident function, creating bubbles for each incident where people died or went
    missing
478 var incTotalGroup = svg.select(".circleGroup").append("g").attr("class", "
    incTotalGroup")
479 var incGroup = svg.select(".circleGroup").append("g").attr("class", "incGroup")
480 function createIncidents() {
481   d3.csv(incSource, function(data) {
482     var inc = incGroup.selectAll("circle")
483     // Sort the incidents, so that the smaller bubbles are drawn on top
484     .data(data.sort(function(a,b) {
485       return b.TOTAL-a.TOTAL}),
486       function(d) {return d.OBJECTID}
487     );
488     // Remove old incident bubbles (since this function is called each time the
    data is updated)
489     inc.exit()
490     .transition()
491     .duration(transitionDur)
492     .ease(d3.easeCubic)
493     .attr("r", 0)
494     .remove();
495     // Enter new incident bubbles
496     inc.enter()
497     .append("circle")
498     // Place the bubbles on the map, using the coordinates in the source data
499     .attr("cx", function(d) {

```

```

500     return projection([d.x, d.y])[0];
501   })
502   .attr("cy", function(d) {
503     return projection([d.x, d.y])[1];
504   })
505   .transition()
506     .duration(transitionDur)
507     .ease(d3.easeCubic)
508     .attr("opacity", 1)
509     .attr("fill-opacity", 0.2)
510   // Scale each bubble, according to the number of dead and missing in the
511   // source data
512   .attr("r", function(d) {
513     return Math.sqrt(parseInt(d.TOTAL) * incScale)
514   })
515   // Define the appearance of the bubbles
516   .attr("fill", "red")
517   .attr("fill-opacity", 0.2)
518   // One particular bubble has a black stroke (needed in the user study)
519   .attr("stroke", function (d) {return d.OBJECTID != 465 ? "darkred" : "black"
520   })
521   .attr("stroke-width", 0.5)
522   .attr("id", function (d) {return "inc"+d.OBJECTID})
523   // Set migrant route as class
524   .attr("class", function (d) {return d.MIGRANTROU.split(" ")[0]} );
525   // Change opacity and stroke width on mouseover/mouseout
526   d3.select(".incGroup").selectAll("circle")
527     .on("mouseover", function() {
528       d3.select(this)
529         .transition()
530         .duration(0)
531         .ease(d3.easeLinear)
532         .attr("fill-opacity", 1)
533         .attr("stroke-width", 1.8)
534     })
535     .on("mouseout", function() {
536       d3.select(this)
537         .transition()
538         .duration(200)
539         .ease(d3.easeLinear)
540         .attr("fill-opacity", 0.2)
541         .attr("stroke-width", 0.5)
542     })
543   // Add info on the incident as a tooltip: number of migrants, location
544   // description, and date
545   .append("title")
546   .text(function(d) {
547     // Use a different syntax if the number of migrants is 1 (singular form) or

```

```

    if there is no location description (leave out this line)
545     if (d.TOTAL == 1) {
546         if (d.LOCATIONDE == "") {return(
547             parseInt(d.TOTAL) + " dead or missing migrant \n" +
548             d.DATEREPORT
549         )} else {return(
550             parseInt(d.TOTAL) + " dead or missing migrant \n" +
551             d.LOCATIONDE + "\n" +
552             d.DATEREPORT
553         )}
554     }
555     else {
556         if (d.LOCATIONDE == "") {return(
557             parseInt(d.TOTAL) + " dead or missing migrants \n" +
558             d.DATEREPORT
559         )} else {return(
560             parseInt(d.TOTAL) + " dead or missing migrants \n" +
561             d.LOCATIONDE + "\n" +
562             d.DATEREPORT
563         )}
564     }
565 });
566 // Group the incident data by migrant route (central or western Mediterranean)
567 var incByCountry = d3.nest()
568     .key(function (d) {return d.MIGRANTROU.split(" ")[0]})
569     .rollup(function(v) { return d3.sum(v, function(d) { return d.TOTAL }) })
570     .entries(data);
571 // Create bubbles showing the total number of dead or missing per migrant route
572 var incTotal = incTotalGroup.selectAll("circle")
573     .data(incByCountry);
574 // Remove old bubbles
575 incTotal.exit()
576     .transition()
577     .duration(transitionDur)
578     .ease(d3.easeCubic)
579     .attr("r", 0)
580     .remove();
581 // Enter new bubbles
582 incTotal.enter()
583     .append("circle").append("title")
584 // Set the coordinates for the bubbles, based on the migrant route
585 incTotal.attr("cx", function(d) {return d.key == "Central" ? 145 : 340})
586     .attr("cy", function(d) {return d.key == "Central" ? 170 : 100})
587     .transition()
588     .duration(transitionDur)
589     .ease(d3.easeCubic)
590     .attr("opacity", 1)
591     .attr("fill-opacity", 0.2)

```



```

592 // Scale each bubble, according to the total number of dead and missing
593 .attr("r", function(d) {
594     return Math.sqrt(parseInt(d.value) * incScale)
595 })
596 // Define the appearance of the bubbles
597 .attr("fill", "darkred")
598 .attr("fill-opacity", 0)
599 .attr("stroke", "darkred")
600 .attr("stroke-width", 1)
601 .attr("stroke-dasharray", "3,3")
602 .attr("id", function (d) {return "inc"+d.OBJECTID})
603 // Set info as a tooltip: name of the migrant route and total number of
migrants
604 incTotal.select("title")
605 .text(function(d) {return(
606     d.key + " Mediterranean total" + "\n" +
607     d.value + " dead or missing migrants"
608     )});
609 displayInc();
610 });
611 };
612
613 // The following helper functions will aid in the creation of flow paths. They are
based on trigonometric rules summarized in an answer by user John Douma on
stackexchange:
614 // http://math.stackexchange.com/questions/409689/how-do-i-find-a-point-a-given-distance-from-another-point-along-a-line?rq=1
615 function slope (x1,y1,x2,y2) {
616     return (y2-y1)/(x2-x1);
617 };
618 function distance (x1,y1,x2,y2) {
619     return Math.sqrt(Math.pow((x2-x1),2) + Math.pow((y2-y1),2));
620 };
621 function shiftX (startX, dist, slp, flowDir) {
622     return startX + Math.sqrt((Math.pow(dist,2))/(1+Math.pow(slp,2)))*flowDir
623 };
624 function shiftY (startX, startY, newX, slp) {
625     return slp*(newX-startX)+startY;
626 };
627
628 // Load flow data. After all the data are loaded, the function analyze() is called.
629 function createFlows() {
630     d3.queue()
631     .defer(d3.csv, destSource)
632     .defer(d3.csv, origSource)
633     .defer(d3.csv, "regionCoord.csv")
634     .defer(d3.csv, "originCoord.csv")
635     .await(analyze);

```

```

636 };
637 // Helper function to join two csv files. This join is based on code by Timo
        Grossenbacher provided in the online guide "Learn JS Data"
638 // http://learnjsdata.com/combine_data.html
639 function join(lookupTable, lookupKey, mainTable, mainKey, select) {
640     var l = lookupTable.length,
641         m = mainTable.length,
642         lookupIndex = [],
643         output = [];
644     for (var i = 0; i < l; i++) { // loop through l1 items
645         var row = lookupTable[i];
646         lookupIndex[row[lookupKey]] = row; // create an index for lookup table
647     }
648     for (var j = 0; j < m; j++) { // loop through m items
649         var y = mainTable[j];
650         var x = lookupIndex[y[mainKey]]; // get corresponding row from lookupTable
651         output.push(select(y, x)); // select only the columns you need
652     }
653     return output;
654 };
655 // Prepare the data for the flows: join data of the number of migrants with
        coordinates
656 function analyze(error, arrDest, arrOrig, regCoord, origCoord) {
657     // This join is based on a code by Timo Grossenbacher provided in the online
        guide "Learn JS Data"
658     // http://learnjsdata.com/combine_data.html
659     // Do the join for the destination regions
660     var joinDest = join(regCoord, "Location", arrDest, "Location", function(numbers,
        coord) {
661         return {
662             country: numbers.Country,
663             region: numbers.Location,
664             totalmig: +numbers.totalmig,
665             // coordinates for the countries (middle part of the flow line)
666             cx1: projection([+coord.cx1, +coord.cy1])[0],
667             cy1: projection([+coord.cx1, +coord.cy1])[1],
668             cx2: projection([+coord.cx2, +coord.cy2])[0],
669             cy2: projection([+coord.cx2, +coord.cy2])[1],
670             // coordinates for the regions (branches of the flow line)
671             rx1: projection([+coord.cx2, +coord.cy2])[0],
672             ry1: projection([+coord.cx2, +coord.cy2])[1],
673             rx2: projection([+coord.rx, +coord.ry])[0],
674             ry2: projection([+coord.rx, +coord.ry])[1]
675         };
676     });
677     // Do the join for the origin regions
678     var joinOrig = join(origCoord, "Origin", arrOrig, "Origin", function(numbers,
        coord) {

```

```

679     return {
680         country: numbers.Country,
681         region: numbers.Location,
682         totalmig: +numbers.totalmig,
683         // coordinates for the countries (middle part of the flow line)
684         cx1: projection([+coord.cx1, +coord.cy1])[0],
685         cy1: projection([+coord.cx1, +coord.cy1])[1],
686         cx2: projection([+coord.cx2, +coord.cy2])[0],
687         cy2: projection([+coord.cx2, +coord.cy2])[1],
688         // coordinates for the regions (branches of the flow line)
689         rx1: projection([+coord.cx2, +coord.cy2])[0],
690         ry1: projection([+coord.cx2, +coord.cy2])[1],
691         rx2: projection([+coord.rx, +coord.ry])[0],
692         ry2: projection([+coord.rx, +coord.ry])[1]
693     };
694 });
695 // Group the data by country
696 var destByCountry = d3.nest()
697     .key(function(d) {return d.country})
698     .entries(joinDest);
699 var origByCountry = d3.nest()
700     .key(function(d) {return d.country})
701     .entries(joinOrig);
702 // Sort the data so that the flows are lined up in the right way
703 destByCountry[0].values.sort(function(a,b) {
704     return b.ry2-a.ry2});
705 destByCountry[1].values.sort(function(a,b) {
706     return a.ry2-b.ry2});
707 origByCountry[0].values.sort(function(a,b) {
708     return a.ry2-b.ry2});
709 origByCountry[1].values.sort(function(a,b) {
710     return b.rx2-a.rx2});
711 // In order for the flows to line up, the starting points of the outer branches
712 // of the flow need to be shifted along the edge of the middle part of the flow
713 // line f
714 function shiftFlows(nested) {
715     // Outer loop: Loop through the countries
716     for (var i=0; i<nested.length; i++) {
717         var sumLeft = 0;
718         // Inner loop: loop through the regions. For each flow path element, cumulate
719         // the values of all the previous elements in the group
720         for (var j=0; j<nested[i].values.length; j++) {
721             nested[i].values[j].sumLeft = sumLeft;
722             sumLeft = nested[i].values[j].sumLeft + nested[i].values[j].totalmig;
723         };
724         // Calculate the total value of all region flows
725         for (var j=0; j<nested[i].values.length; j++) {
726             nested[i].values[j].sumTotal = sumLeft;

```

```

724     };
725     // Adapt the coordinates for the starting points of the outer branches
726     for (var j=0; j<nested[i].values.length; j++) {
727         var
728             cx1 = nested[i].values[j].cx1,
729             cy1 = nested[i].values[j].cy1,
730             cx2 = nested[i].values[j].cx2,
731             cy2 = nested[i].values[j].cy2,
732             rx1 = nested[i].values[j].rx1,
733             ry1 = nested[i].values[j].ry1,
734             rx2 = nested[i].values[j].rx2,
735             ry2 = nested[i].values[j].ry2,
736             totalmig = nested[i].values[j].totalmig,
737             // slp is the slope perpendicular to the middle part of the flow line
738             slp = -1/slope(cx1,cy1,cx2,cy2),
739             // sumLeft and sumTotal determine the distance of the shift
740             sumLeft = nested[i].values[j].sumLeft,
741             sumTotal = nested[i].values[j].sumTotal,
742             dist1 = (sumLeft+totalmig/2)/flowScale,
743             dist2 = (sumTotal/2)/flowScale;
744         // Shift the starting point of the line
745         // First, the point is shifted based on the sum of the width of all
previous paths. This aligns all branches next to each other
746         nested[i].values[j].rx1 = shiftX (cx2, dist1, slp, -1);
747         nested[i].values[j].ry1 = shiftY (cx2, cy2, nested[i].values[j].rx1, slp);
748         // Then, the point is shifted in the other direction, based on the total
width of all paths. This lines the branches up with the main flow line
749         tx = nested[i].values[j].rx1;
750         ty = nested[i].values[j].ry1;
751         nested[i].values[j].rx1 = shiftX (tx, dist2, slp, 1);
752         nested[i].values[j].ry1 = shiftY (tx, ty, nested[i].values[j].rx1, slp);
753     };
754 };
755 }
756 // Do the shift for destination and origin lines
757 shiftFlows(destByCountry);
758 shiftFlows(origByCountry);
759
760 // Sum the values of all region values by country
761 var countrySummed = d3.nest()
762     .key(function(d) {return d.country})
763     .rollup(function(v) {
764         var cx1 = d3.mean(v, function(d) {return d.cx1});
765         var cy1 = d3.mean(v, function(d) {return d.cy1});
766         var cx2 = d3.mean(v, function(d) {return d.cx2});
767         var cy2 = d3.mean(v, function(d) {return d.cy2});
768         return{
769             sum: d3.sum(v, function(d) {return d.totalmig}),

```

```

770     x1: cx1,
771     y1: cy1,
772     x2: cx2,
773     y2: cy2
774   }
775 })
776   .entries(joinDest);
777
778   // After all the data has been prepared the right way, call the functions that
       create the lines. The second instantiation of each line is a ghost with a
       greater width that facilitates hovering over thin lines
779   lineFlow(joinDest, ".destGroup", -1, true);
780   lineFlow(joinDest, ".destGroupGhost", -1, false);
781   lineFlow(joinOrig, ".origGroup", 1, true);
782   lineFlow(joinOrig, ".origGroupGhost", 1, false);
783   lineCountry(countrySummed, true);
784   lineCountry(countrySummed, false);
785 };
786
787 // Function to generate the middle part of flow line (country total)
788 function lineCountry (data, show) {
789   var flowSelect = d3.select(show ? ".countryGroup" : ".countryGroupGhost").
       selectAll("path");
790   var flow = flowSelect.data(data, function (d) {return d.key});
791   flow.enter().append("path").append("title");
792   flow
793     // Create a straight line using the SVG path mini-language
794     .attr("d", function(d) {return(
795       "M" + d.value.x2 + "," + d.value.y2 +
796       "L" + d.value.x1 + "," + d.value.y1
797     )})
798     // Define appearance
799     .attr("stroke", "#4d4d4d")
800     .attr("stroke-linecap", "butt")
801     .attr("fill", "none")
802     // Set opacity of the ghost line to 0
803     .attr("opacity", show ? 1 : 0)
804     .attr("id", function (d) {return d.key})
805     // Add tooltip
806     .select("title")
807     .text(function(d) {
808       return (
809         "to " + d.key + "\n"
810         + addSpaces(d.value.sum) + " migrants"
811       )})
812     // Set stroke width (higher for the ghost line)
813     flow.transition()
814     .duration(transitionDur)

```

```

815     .ease(d3.easeCubic)
816     .attr("stroke-width", show ?
817         function(d) {return (d.value.sum)/flowScale} :
818         function(d) {return (d.value.sum)/flowScale + 5}
819     );
820 };
821
822 // Function to generate the outer branches of the flow line (origin and destination
823 // regions)
824 function lineFlow(data, flowSubGroup, flowDir, show) {
825     var flowSelect = d3.select(flowSubGroup).selectAll("path");
826     var flow = flowSelect.data(data);
827     var slp, dist, ctrlx, ctrly;
828     flow.enter()
829     .append("path").append("title");
830     flow
831     // Define appearance
832     .attr("stroke", "#4d4d4d")
833     .attr("stroke-linecap", "butt")
834     .attr("fill", "none")
835     // Set opacity of the ghost line to 0
836     .attr("opacity", show ? 0.75 : 0)
837     .transition()
838     .duration(transitionDur)
839     .ease(d3.easeCubic)
840     // Create a quadratic Bezier curve using the SVG path mini-language
841     .attr("d", function(d) {
842         // For the Bezier curve, a control point is created. It lies in the direction
843         // of the extension of the middle part of the flow line. The distance from the
844         // starting point of the branch of the flow line is 0.5 times the distance between
845         // the branches' starting and end point
846         slp = slope(d.cx1,d.cy1,d.cx2,d.cy2);
847         dist = distance(d.rx1,d.ry1,d.rx2,d.ry2)*0.5;
848         ctrlx = shiftX(d.rx1,dist,slp,flowDir);
849         ctrly = shiftY(d.rx1,d.ry1,ctrlx,slp);
850         return (
851             "M" + d.rx1 + "," + d.ry1 +
852             "Q" + ctrlx + "," + ctrly + "," +
853             d.rx2 + "," + d.ry2
854         );
855     });
856     // Set stroke width (higher for the ghost line)
857     .attr("stroke-width", show ?
858         function(d) {return d.totalmig/flowScale} :
859         function(d) {return d.totalmig/flowScale+10});
860     // Add tooltip
861     flow
862     .select("title")
863     .text(function(d) {

```

```

859     return d.region + "\n"
860     + addSpaces(d.totalmig) + " migrants"});
861 };
862
863 updateAll();
864
865 </script>
866 </body>
867 <svg height="10" width="10" xmlns="http://www.w3.org/2000/svg" version="1.1">
868   <defs>
869     <pattern id="diagonal-stripe-3"
870       patternUnits="userSpaceOnUse"
871       width="10"
872       height="10">
873       patternUnits="userSpaceOnUse"
874       width="10"
875       height="10"
876     >
877     <image
878       xlink:href =
879       "data:image/svg+xml;base64,PHN2ZyB4bWxucz0naHR0cDovL3d3dy53My5vcm
880       cvMjAwMC9zdmcnIHdpZHRoPScxMCCgaGVpZ2h0PScxMCC+CiAgPHJlY3Qgd2lkdGg
881       9JzEwJyBoZWlnaHQ9JzEwJyBmaWxsPSd3aG10ZScvPgogIDxwYXRoIGQ9J00tMSwx
882       IGwyLCOyCiAgICAgICAgICAgTAsMTAgbDEwLCOxMAogICAgICAgICAgIE05LDEExI
883       GwyLCOyJyBzdHJva2U9J2JsYWNrJyBzdHJva2Utd2lkdGg9JzZmLz4KPC9zdmc+"
884       x="0"
885       y="0"
886       width="10"
887       height="10"
888     >
889     </image>
890   </pattern>
891 </defs>
892 </svg>
893 </html>

```

B Questionnaire of the online survey

Star symbols indicate questions where an answer was required. Annotations in monospace font in curly brackets were not seen by the participants taking the questionnaire.

Welcome!

Thank you very much for taking part in this online user study on geographic visualization of migration data. The results of this study will be part of Gordon Bühler's master thesis at the University of Zurich. Your contribution as a participant is vital for a positive outcome to the thesis.

The online user study will take about 20 minutes to complete.

This study will take part in two groups. In order to be assigned to one of the groups, please state the day of the month of your birthday below.

★ On what day of the month were you born?

{This page only shows up if the answer to the question "On what day of the month were you born?" was <=15 (group: interactive)}

Getting set up

Please open the following link in a separate browser tab or on a second screen:

<http://tiny.cc/migstudy-a>

This link will bring up the visualization that is the focus of this study. Make sure that you view the visualization on a big enough screen (**mobile devices**, such as smartphones or tablets, will **not work**).

Hint: you could open the visualization on your computer, and the questionnaire on your smartphone. If you choose this option, just start the questionnaire on your smartphone:

<https://lamapoll.de/migstudy>

Please click next when you have brought up the visualization in a separate browser tab or on a second screen.

{This page only shows up if the answer to the question "On what day of the month were you born?" was >15 (group: static)}

Getting set up

Please open the following link in a separate browser tab or on a second screen:

<http://tiny.cc/migstudy-b>

This link will bring up the visualization that is the focus of this study. Make sure that you view the visualization on a big enough screen (**mobile devices**, such as smartphones or tablets, will **not work**).

Hint: you could open the visualization on your computer, and the questionnaire on your smartphone. If you choose this option, just start the questionnaire on your smartphone:

<https://lamapoll.de/migstudy>

Please click next when you have brought up the visualization in a separate browser tab or on a second screen.

A first look

Please take a moment to take a first look at the visualization. It shows migration flows and migrant deaths on the Mediterranean Sea in the year 2015. Make sure you understand the visualization. Do not explore the content of the map yet, but focus on the technical understanding.

Click next when you feel comfortable with the visualization.

★ How are you feeling?

Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

At this moment, I am tired and low on energy.

I am excited to see what waits for me in this study.

{This page only shows up if the answer to the question "On what day of the month were you born?" was <=15 (group: interactive)}

Explore the map now. Do you spot something interesting or remarkable in the visualization? Please take short notes.

If you wish, you can now take some time and explore the visualization on your own. When you find something that you find interesting or remarkable, please take a short note in the field below. Feel free to take as many notes as you like. Move on as soon as you want to.

★ How do you like this visualization so far?

Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

Working with this visualization is tiring and cumbersome.

The interactive elements make this visualization more interesting.

The interactive elements make working with the visualization quicker and more effective.

I would prefer a static visualization over this interactive one.

Generally, when I see an online news story with a map, I will take some time to explore the map.

Generally, when I see an online news story with an interactive visualization, I am excited to explore it.

{This page only shows up if the answer to the question "On what day of the month were you born?" was >15 (group: static)}



Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

★ **How do you like this visualization so far?**
Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

Working with this visualization is tiring and cumbersome.

With some interactive elements, the visualization would be more interesting.

With some interactive elements, working with the visualization would be quicker and more effective.

I would prefer an interactive visualization over this static one.

Generally, when I see an online news story with a map, I will take some time to explore the map.

Generally, when I see an online news story with an interactive visualization, I am excited to explore it.

It is our duty in Europe to help refugees by giving them a safe space to live.

If we take in too many foreign people, it will hurt our identity.

The media are not focussing enough on the problems that the refugees bring for Europe.

The media should do more to show the suffering that refugees are seeing from.

In the following section, you will be given a number of tasks to solve with the help of the visualization.

Please click next to continue.

★ **Please state your personal *interest* in the following areas**

Please rate your interest from 1 (very little interest in the topic) to 5 (very high interest in the topic).

	1 – very little interest	2 – little interest	3 – neutral interest	4 – high interest	5 – very high interest
Maps	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data visualization	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Migration	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online news stories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

★ **Please state your personal *experience* in the following areas**

Please rate your expertise from 1 (very little knowledge of the topic) to 5 (very high knowledge of the topic).

	1 – very little knowledge	2 – little knowledge	3 – neutral knowledge	4 – high knowledge	5 – very high knowledge
Data visualization	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Migration	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online news stories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Maps	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do the dark grey flow lines represent?

- Migrants that travelled to Europe
- Migrants that travelled away from Europe
- Migrants that travelled to Europe or away from Europe

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

What do the transparent red bubbles represent?

- Number of migrants that reached Europe
- Incidents where migrants died or went missing

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

How does the size of each bubble correspond to the number of migrants?

If you don't know, please do not answer.

- With its area
- With its diameter

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

Find the incident represented by a bubble with a black border. In what period did it happen?

- Jan-Mar
- Apr-Jun
- Jul-Sep
- Oct-Dec

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

-
-
-
-
-

This question was easy to answer.

-
-
-
-
-

In the area of what destination country did it happen?

{This is a follow-up question. It only shows up after the previous question has been answered}

If you don't know, please do not answer.

- Italy
- Greece

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

-
-
-
-
-

This question was easy to answer.

-
-
-
-
-

★ **How many people died or went missing during this incident?**

{This is a follow-up question. It only shows up after the previous question has been answered}

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

-
-
-
-
-

This question was easy to answer.

-
-
-
-
-

★ How many migrants went to Calabria between April and June 2015?

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

Find the single incident with the largest number of dead or missing migrants over all of 2015. In what period did it happen?

- Jan-Mar
- Apr-Jun
- Jul-Sep
- Oct-Dec

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

★ How many people died or went missing during this incident?

{This is a follow-up question. It only shows up after the previous question has been answered}



{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

Again referring to the large incident from the last page: What could be some of the reasons that this incident happened? Please give a maximum of three reasons, with reason 1 being the most important reason.

Reason 1

Reason 2

Reason 3

★ Rank the periods by the total number of migrants.

Drag and drop the period with the highest number to the top, and the period with the lowest number to the bottom.

↑↓ Jan-Mar	<input type="text"/>
↑↓ Apr-Jun	<input type="text"/>
↑↓ Jul-Sep	<input type="text"/>
↑↓ Oct-Dec	<input type="text"/>

★

I am confident that my answer is correct.

This question was easy to answer.

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

★ How do you agree with this statement?

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The civil war in Syria was the main driver of migration to Europe across the Mediterranean Sea in 2015.

This question was easy to answer.

What was the period with the highest total number of dead or missing migrants?

- Jan-Mar
- Apr-Jun
- Jul-Sep
- Oct-Dec

★ For this question, focus on the island of Lesbos.

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

Most migrants that went to Lesbos came from Syria.

This question was easy to answer.

★

{This is a follow-up question. It only shows up after the previous question has been answered}

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

I am confident that my answer is correct.

This question was easy to answer.

★ How does the number of dead or missing migrants relate to the total number of migrants?

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The number of dead or missing migrants is a direct consequence of the total number of migrants.

This question was easy to answer.

★ How does the number of dead or missing migrants relate to the search operation areas of the Operations Triton and Sophia?

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The extent of the search operation areas had a strong effect on the number of dead or missing migrants.

This question was easy to answer.

★ How do you agree with the following statement?

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The European community should do more to prevent the death of migrants on the Mediterranean Sea.

★ How can the number of migrant deaths on the Mediterranean Sea be reduced?

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

Identifying and neutralising the networks of smuggling and human trafficking.

Working to end the conflicts that drove migrants from their country.

Helping countries outside of Europe to take in more refugees.

Finding migrant boats on the Mediterranean Sea before they shipwreck.

Making Europe unattractive as a destination for migrants by swiftly sending them back.

Not intervening when incidents happen, so that migrants do not assume an easy passage.

**Did you spot anything else that is interesting or remarkable?
Please take short notes.**

This is the last question that involves using the visualization. If you wish, you can again note some interesting or remarkable findings in the field below. This is absolutely optional. You can spend as little or much time here as you enjoy.

★ How are you feeling?

Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The visualization helped to fully immerse myself in the topic.

At this moment, I am tired and low on energy.

The tasks felt tedious and boring for me.

Working with the visualization felt meaningful to me.

{This page only shows up if the answer to the question "On what day of the month were you born?" was <=15 (group: interactive)}

{This page only shows up if the answer to the question "On what day of the month were you born?" was >15 (group: static)}

★ **Now after having worked with this visualization, how do you like it?**

Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

The interactive elements made working with the visualization quicker and more effective.

I would prefer a static visualization without any interactive elements.

The interactive elements made the visualization more interesting.

The interactive elements were distracting for solving the tasks.

★ **Now after having worked with this visualization, how do you like it?**

Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

With some interactive elements, working with the visualization would have been quicker and more effective.

I would prefer a static visualization without any interactive elements.

With some interactive elements, the visualization would have been more interesting.

Interactive elements could have been distracting for solving the tasks.



Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

- I think that I would like to use this type of visualization frequently.
- I found the visualization unnecessarily complex.
- I thought the visualization was easy to use.
- I think that I would need the support of a technical person to be able to use this visualization .
- I found the various elements in this visualization were well integrated.
- I thought there was too much inconsistency in this visualization .
- I would imagine that most people would learn to use this type of visualization very quickly.
- I found the visualization very cumbersome to use.
- I felt very confident using the visualization .
- I would need to learn a lot of things before I could get going with this visualization .



Please indicate your agreement with the following statements from 1 (strongly disagree) to 5 (strongly agree).

1 - strongly disagree 2 - disagree 3 - neutral 4 - agree 5 - strongly agree

- It is our duty in Europe to help refugees by giving them a safe space to live.
- If we take in too many foreign people, it will hurt our identity.
- The media are not focussing enough on the problems that the refugees bring for Europe.
- The media should do more to show the suffering that refugees are fleeing from.

You have reached the end of the online user study. Thank you so much for taking part and completing the survey to the end.

Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Gordon Bühler