Implementing social media data in algorithms for estimating crowdedness in tourist’s attractions

Author: Natali Polishuk
Supervisors: Nuno Otero, Didac Gil de la Iglesia
Examiner: Jenny Lundberg
Date: 31th May 2017
Subject: Social Media and Web Technologies
Level: Master
Course code: 5ME10E
Abstract

This thesis aims to explore the implementation of social media crowd data collection (Twitter posted tweets) within algorithms. The tweets data collection reflects tweets related to a tourist’s sight-seeing coordinates location: the collected tweets were posted in a 2 KM radius from the center of each selected sight-seeing. The total number of tweets collected was 5800 in Barcelona, and 1880 in Rome.

The application which will provide such service, will offer the users a navigation feature. The users will be able to insert several preferences (sight-seeing sites selections, hours and days of the week), and the application will offer them the three best paths. These paths will be based on three parameters: time of walking distance between the locations, average visiting time and estimated waiting time based on crowdedness. To evaluate this navigation system, three aspects were reviewed: user’s needs, data reliability and user experience. The data reliability evaluation represented as the reliability of the presented route calculation results. The user experience aspect relates to the user’s emotions and effects, enjoyment and system aesthetics (Bargas-Avila & Hornbæk, 2011) and evaluates the system characteristics of the system usability, complexity and functionality from the user’s perspective and attitudes toward the system (Hassenzahl & Tractinsky, 2006).

Four research questions were presented and answered: (1) what are the current habits and preferences of tourists, for the use of mobile phones as an information source while traveling? (2) “Does providing new information to tourists’ concerning crowdedness and rush hours, have influence on their plans and does it change their plans?” (3) Which features should a mobile application have, in order to support tourists’ activities when trying to plan sightseeing routes in cities? (4) Can social media data be reliably used in algorithms for measuring and estimating time aspects and crowdedness?

The data reliability evaluation results presented connection between the system results and the waiting time in practice. This proves that crowdedness can be predicted by this type of information. However, the tweets time analysis could be improved by reviewing and analyzing each tourist attraction separately.

For evaluating the user experience, five participants used the application with two aims: first, planning a trip using the application; second, for using the application in real time outdoors (as a “live experience”). The user experience evaluations findings showed users had a positive feedback concerning the features and design of the prototype. The users enjoyed planning the trip with this application and enjoyed the suggested route for navigating between the attractions in real time.
Keywords

Tourism, social media, navigation, data reliability, user experience, crowdedness, travel planning, tourist attractions, smart tourism, tourism technologies

Acknowledgements

This thesis is the final stage of a challenging and satisfying master program at Linnaeus University. The program of “social media and web technologies” has exposed me to critical thinking, professional knowledge, and innovation ideas.

First I wish to thank my supervisor Nuno Otero. Even though the official time frame for the formal supervision has long passed, Nuno agreed to supervise my thesis in detail. His suggestions and comments contributed to my learning skills in the research field, and to this thesis in specific.

Next, I wish to thank my second supervisor, Didac Gil de la Iglesia, who guided me from the initial thoughts about this thesis subject, and help me form this idea into a working prototype.

A big gratitude is to my mother and my mother-in-law, who motivated me to keep up the hard work, and looked after my two children, starting from the first day of this master program. Without their help and support, this could not have been done.

Another special thanks, belong to those who perform a full English and academic proofreading to this thesis, my friends: Lisa Ruehmann and Lital Letschert.

Last but not least, is to my close family. I wish to thank my beloved husband, Avner, who motivated and supported me in both good times and challenging times. Every time he said the right words and taught me what true love is all about. The best relationship anyone can ask for, is a partner that motivates you to reach the highest as possible. I also wish to thank both of my children: Yali and Lori. Although they might be still young, they taught me priorities, and what truly matters in life.

Thank you all!
# Contents

## 1 Introduction
1.1 Tourism Data Statistics ............................................. 1
1.2 Aims ................................................................. 2
1.3 Motivation .......................................................... 2
1.4 Research Questions & Research Method ......................... 4
   1.4.1 Research Questions ........................................ 4
   1.4.2 Research Method ........................................... 6
      1.4.2.1 User Needs Survey .................................. 7
      1.4.2.2 Data Reliability ...................................... 7
      1.4.2.3 User Experience (UX) ................................. 8
1.5 Thesis Structure ................................................... 9

## 2 Related Work
2.1 Tourism and Technology .......................................... 11
2.2 Market Search and Survey ...................................... 14
   2.2.1 Crowdedness as an Issue ................................ 14
   2.2.2 Trip Planning .............................................. 16

## 3 Initial Requirements and System Architecture ............ 18
3.1 Initial Requirements: User Needs Analysis ................. 18
   3.1.1 Prototype Implementation ................................ 22
3.2 System Architecture & Algorithm ............................. 23
   3.2.1 Tweet Data Analysis ...................................... 23
   3.2.2 System Architecture ....................................... 25
   3.2.3 Algorithm .................................................. 26
   3.2.4 System Flow ............................................... 28
3.3 Technology ....................................................... 30
   3.3.1 Collecting Data - Twitter ................................ 30
   3.3.2 Programing Language & Database ........................ 31
      3.3.2.1 Google Maps API: ..................................... 32

## 4 Data Reliability: Survey and Self-Reported Data ...... 33
4.1 Prototype - Version 1 ............................................ 33
   4.1.1 Home Page ................................................ 33
   4.1.2 ‘Selecting Tourist’s Attractions and Preferences’ Page 34
   4.1.3 ‘Selecting the Optimal Path’ Page ....................... 35
   4.1.4 ‘Path Output and Information’ Page ..................... 36
4.2 Methodology ...................................................... 37
   4.2.1 Participants ............................................... 37
   4.2.2 Procedure ................................................ 37
List of Figures

Figure 1, Work flow of the thesis ______________________________ 6
Figure 2, Screens from the example trip planning websites/applications ______________________________ 17
Figure 3, Answers regarding the frequency for collecting information/data about the destination____________________ 19
Figure 4, Information users collects by their mobile phone while traveling ________________________________________ 20
Figure 5, Sharing personal information on social media platforms ____________________________________________ 21
Figure 6, Visual of the tweets collection results by the hours of the day, Barcelona, Museum National Example, Tuesday __ 23
Figure 7, code snippet calculation of the estimated waiting time _______________________________________________ 24
Figure 8, System Architecture Layers __________________________ 25
Figure 9, Paths Calculation __________________________________ 27
Figure 10, User actions & algorithm calculation ___________ 29
Figure 11, Database Structure ____________________________________________ 32
Figure 12, Landing page screen ______________________________________ 33
Figure 13, Preferences expanded ____________________________________________ 34
Figure 14, Sites selections and preferences collapse __________ 34
Figure 15, informative message ____________________________________________ 35
Figure 16, Path choosing page ____________________________________________ 35
Figure 17, Output page ____________________________________________ 36
Figure 18, Results summary by waiting times ___________ 39
Figure 19, Home page screen ____________________________________________ 43
Figure 20, Selecting tourist’s attractions screen ___________ 44
Figure 21, Preferences expanded ____________________________________________ 45
Figure 22, Path output screen ____________________________________________ 46
Figure 23, Rating the experience using the Likert scale ___________ 49
Figure 24, Users testing the application in the streets of Tel Aviv ____________________________________________ 51
Figure 25, Application enjoyment ____________________________________________ 51
Figure 26, Page rating by Likert scale ____________________________________________ 52

List of Equations

Equation 1, Calculating the optimal path ______________________________ 26

List of Tables

Table 1, survey participants’ data __________________________________ 18
Table 2, Reliability evaluation participates data __________________________ 37
Table 3, All the recorded results concerning time __________________________ 38
Table 4, UX participants data ____________________________________________ 47
Chapter 1

1 Introduction

1.1 Tourism Data Statistics

Based on an annual survey concerning the preferences of the Europeans’ toward tourism, more than seven out of ten people (72%) had spent at least one night away from home for a holiday, and more than half (56%) who went on a holiday spent between 4-13 nights away from home (Flash Eurobarometer, 2015). Six out of ten people had traveled outside of their own country and 66% have reported to use the internet for planning their holiday (Flash Eurobarometer, 2015).

As for trip planning, travel choices are not decided entirely before trips but rather happen across the different phases of the trips, some travelers intentionally avoid pre-plans in order to “enjoy spontaneous travel experiences that comply with the unanticipated circumstances faced during trips” (Kah & Lee, 2014; Wilson & Becken, 2010).

Stewart and Vogt (1999) argued that travelers often change pre-trip plans while being introduced to new information during the trip itself.

Approximately 44% of trips had at least one unplanned visit site (Kah & Lee, 2014) and most travelers who gathered information while travelling might change the places they originally planned to visit (Kah & Lee, 2014; Fesenmaier, Vogt & Stewart, 1993). One of Kah’s and Lee’s (2014) conclusions was that the longer the length of stay, the more likely the respondents to partake in unplanned activities during the trip. Additionally, travelers make plans before their trips but these plans are “frequently changed when they are provided with new information arising from the use of navigation during the trip” (Kah & Lee, 2014).
According to research, experienced users are using mobile devices to look for traveling information before, during and after the trip (Hsieh, Wu, Tsai, Shih & Li, 2012).

Tourist’s attractions can use technology in many ways and some of them have been using it for their benefaction (Egger, 2013). The existence of implementations for using technologies in tourist’s attractions is mostly common, but by reviewing the ratio of implementations and studies in this subject, the latter is in a reduced relative (Egger, 2013).

1.2 Aims

This thesis aims to explore the implementation of social media crowd data collection (users posts and insights), into algorithms for calculating the best path route by time and crowdedness. This implementation was tested in order to predict crowdedness in the most popular tourist attractions and can be associated in the tourism field. An implementation of a new tourism technology application, which can predict crowdedness in tourist’s attractions, might improve the tourism experience and increase the number of visitors at the attractions sites. This application will do so by suggesting the tourists’ route paths, which may lead for spreading the tourists’ visits to the preferred sites throughout the length of a day and prevent the crowdedness during peak hours.

By collecting data from social networks, this application will offer the users a navigation feature, in which they can insert several preferences (such as sites selections, hours and days of the week and a starting point) and will be offered with the three best paths based on three parameters: time estimation of walking distance between the sites, average visit time and estimated waiting time at each site. This algorithm will plan the best path between tourist attractions and will help navigate more efficiently based on the user preferences.

1.3 Motivation

When traveling around the world and wishing to visit the most famous sites a tourist often encounters some, to severe, crowding at sights during certain days and hours. It is therefore reasonable to assume that visitors would prefer to avoid those times, given the chance to do so. Moreover, it is reasonable to assume that this will improve
traveling experience and by doing so, tourist attractions will also attract more visitors.

Tourism is a well-known income for countries and many countries try to improve this aspect constantly keeping in mind that the tourist's attitudes towards the experience will impact tourism in their country: both by tourists sharing experiences and by returning tourists (Flash Eurobarometer, 2015).

Social media data analysis can be used in more means than the individual user data analysis. It can be used for learning human behaviors, as can be seen in the field of psychology for identifying human mental disorder (Chang, Saravia & Chen, 2016) and for identifying users' personality traits (Thilakaratne, Weerasinghe & Perera, 2016). Analyzing a group of people can also be done by the post text content - for example in users' posts after a crime incident and its content (Marivate & Moiloa, 2016). This was also used for organizations analysis, same for reviewing the text content of the posts, in terms of text analysis for learning the costumers’ needs and feedback regarding a railway service (Yang & Anwar, 2016).

The application developed within this thesis does not analyze the social media data by the posts context, instead it is based only on the numbers of posts and location. This data was implemented in algorithms, for the recommended routes calculations. This information was implemented in algorithms related to the tourism field for a route calculation, where the suggested paths was based on popular tourist's attractions.

In the application developed, the users will review an optimal navigation path calculated taking into consideration the time and distance, as well as estimations of crowdedness based on social media data analysis.

Developing and evaluating this system should not focus merely on the calculation results of the data reliability and validity. Data reliability and validity, in this thesis, is referring to the reliability of the presented data calculation results. Are the places presented in the navigation as crowded in real time as the application algorithms have calculated? Is this information reliable and can users use it and trust the estimations to be accurate and according to the real-time crowdedness status?

Other than the evaluation of the data reliability, another important evaluation to be performed is the user experience (UX). The term user experience refers to the dimensions of the UX as user’s emotions and effects, enjoyment and system aesthetics (Bargas-Avila & Hornbæk, 2011). It evaluates the systems characteristics of the system usability,
complexity and functionality from the user’s perspective and attitudes towards the system (Hassenzahl & Tractinsky, 2006). We, as researchers, in the field of information technology systems, should look beyond data and numbers and ensure that the user interface is easy and efficient, so that our solution will be best used and improve the overall experiences for people.

The relevant stakeholders of the application are the parties that will have interest in the use of the system. It involves understanding multiple potential interests including both financial and end-users interests (MacCaulay, 1994; Sharp, Finkelstein & Galal, 1999). In the case of the application developed these are:

- Tourists will be the primary users, whom will enjoy the solution itself and will be frequently hands-on users of the solution (Easton, 1987; Sharp, Finkelstein & Galal, 1999).
- Countries DMO’s (destination marketing organizations) will be the secondary users. Users who are occasional and need intermediation between the system and the developers/owners (Easton, 1987; Sharp, Finkelstein & Galal, 1999). The DMO’s wishing to increase the tourism incomes by providing such service to tourists.
- Developers and researchers will be the tertiary users, users who are affected from introducing such a system and will be influence from its purchase (Easton, 1987; Sharp, Finkelstein & Galal, 1999), whom could use ideas for implementing social media data in system algorithms, so that it can be used in many aspects which were not used before.

### 1.4 Research Questions & Research Method

#### 1.4.1 Research Questions

*RQ1: “What are the current habits and preferences of tourists, for the use of mobile phones as an information source while traveling?”*

With RQ1, this thesis aims to inquire about tourist’s habits regarding the use of a mobile phone while traveling. For example, how often do tourists use navigation in their mobile phone? Do they collect data and change their sightseeing plans according to the information provided in the mobile phone applications? Information about these issues will be collected with a preliminary user’s needs survey.
RQ2: “Does providing new information to tourists’ concerning crowdedness and rush hours, have influence on their plans and does it change their plans?”

RQ2 investigates the assumption that providing new information (through the mobile application) while traveling, can change tourists’ path and pre-plans. For example, if new information is presented (such as crowdedness), the users will make new decisions.

More specifically, RQ2 inquires if tourists will try to avoid crowdedness and rush hours for a better traveling experience. This thesis will test whether tourists will prefer to stay at the nearest, but crowded location, or will they change their plan, in order to visit a site which is not as crowded but is further away.

RQ3: “Which features should a mobile application have, in order to support tourists’ activities when trying to plan sightseeing routes in cities?”

With RQ3, this thesis aims to inquire concerning the tourists’ preferences concerning trip plans and define requirements based on the findings and reflect upon those requirements within the design and development of the mobile application. Based on the pre-development user survey for understanding the users’ needs and requirements, relevant data will be collected and included in the development for a better user experience. In addition, research about the uses of technology in the tourism field will be done and will help predict the needed features for this mobile application.

RQ4: “Can social media data be reliably used in algorithms for measuring and estimating time aspects and crowdedness?”

With RQ4 this thesis aims to inquire the use of social media in algorithms (data collection and analysis), and establish their reliability as a source for crowdedness in selected locations. This thesis will aspire to draw information from a non-related topic to tourism, such as social media, and use it to create time lines for locations and tourists.

More specifically, RQ4 will investigate if information derived from social media live records, can be a good indicator of crowdedness at a certain location, based on GPS information and information people have agreed to share in their social media account.

This research is testing whether this information is useful to generate paths between touristic sites and support touristic activities, by using Twitter as the social media platform as to be described in chapter 3.3.1.
1.4.2 Research Method

For addressing the research questions mentioned above, the following work flow has been formulated (cf. Figure 1).

1. Literature Review and Related Work: this section covered similar studies and research in the field of mobile applications and tourism, trying to uncover reviewing tourists’ habits and preferences in order to start generating requirements for the application to be developed. Additionally, a market oriented search and survey was conducted in order to explore what already exists as commercial applications targeting similar issues.

2. User Needs Survey: based on the previous step, a survey was created to inquire about: (a) the users’ preferences and habits regarding traveling, (b) the use of mobile while traveling and (c) what popular features should such an application provide.

3. Collecting Social Media Data: collecting real-time information from Twitter (tweets published in a specific location) over a defined period (one week to 14 days).

4. Prototype Development (First Version): based on the previous steps, the application was developed following the information presented and researched.

5. Data Analysis and Prototype Implementation: the collected tweets data was analyzed for defining and evaluating a waiting
time value for each tweet (for example, each tweet valued as X minutes). This data was implemented in the application database, and in the application algorithms.

6. Data Reliability Evaluation: after the application was fully developed, an evaluation exercise compared the real waiting time vs. the application results.

7. Prototype Development (Second Version): based on some application issues uncovered in the data reliability evaluation another prototype was developed that solved them.

8. User Experience Evaluation: the second version of the prototype was tested following the users experience evaluation approach. With this evaluation the goals was to inquire if (a) the users enjoyed using the application; (b) did they find it useful and (c) did it improve their trip experience?


As can be seen in the workflow (cf. Figure 1) to evaluate this navigation and guidance system, three aspects were reviewed: the user’s needs, data reliability and user experience.

1.4.2.1 User Needs Survey

Prior to developing the prototype, a survey was distributed among future potential users. The users input was used to review and evaluate which features an application of this type should contain. In addition, it aided the definition of the scope of the application to be developed within this thesis. This survey reviewed the user’s behavior toward traveling and the usage of mobile phones. Within this survey, the personal social media usage of the participants, and the habits of sharing content while traveling were evaluated. Based on this survey, a list of requirements was created and the findings were implemented in the prototype development.

1.4.2.2 Data Reliability

For evaluating the usefulness of the collected data, the analysis and its reliability in the algorithm, eight users have tested the application in real-time and compared the application results to the real-time results. The question that needed to be answered was: Is the sightseeing site as crowded as the application estimated? After the users compared the results they documented it in a user survey, in which they noted down their results and were asked about other aspects from the application as can be seen in the evaluation section (Chapter no.4).
In addition, the author of this thesis also tested the reliability while being in Barcelona. The times were measured and different aspects of the application and its features were reviewed. This review was performed after the users had participated in the evaluation and noted down their results. The sites which had not been visited by the users, were reviewed by the researcher to ensure completeness of the data. The gathered data was documented by entering the values into a survey as well.

In this evaluation, the algorithm was tested based on users’ personal reports, and the results indicated if implementing social media data in algorithms to judge whether touristic places are crowded, is a reliable method.

1.4.2.3 User Experience (UX)

For evaluating the UX, the method and process of the user centered design (UCD) approach was applied. This method has three phases: design research, design and design evaluation (Williams, 2009). The design research was conducted applying two methods: market search and survey for similar systems, and a user survey distributed among users who have traveled abroad in the past year. In the design phase, based on the design research aspects, the prototype was developed and designed.

The final phase included the design evaluation - in which the prototype was tested among real users indoors and outdoors. The application was tested by two types of user experience: trip planning indoors and a live experience outdoors. The findings from this evaluation could help design an application which could serve two proposes (a) trip planning crowdedness estimation and (b) real time traveling crowdedness navigation:

1. **Trip Planning**

With this evaluation, the application was tested as a source for trip planning for traveling time. The application can serve users as a source for popular tourist attractions list, a recommended route path – and all in the stage of planning a trip abroad.

The evaluation included a survey about the different features of the application. Aspects such as design, screen flow, available features and user scenarios have been covered. Following the feedback from this phase, the same users continued to the next step (using the application in real-time / as a live experience). This part included a survey and access to the application.
2. Live Experience

The users from the first UX section (trip planning) tested the application in Israel, Tel Aviv city center. At this point the application data was fabricated. As the common users in Israel are not using Twitter often (based on information from social media users by region (Bullas, 2013)) and the thesis does not test the data reliability within this section it is not further discussed here.

A comparison from the user experience aspect was done between the trip planning and the live experience trip. Some features which may have seemed important before, might change priority while testing the application in real-time, and other requirements might come up when the application was tested outdoors. This part included an open user observation, group discussion and a questionnaire.

1.5 Thesis Structure

The thesis structure is as follows:

The first chapter is the introduction – this chapter describes tourism statistics, and the motivation for this thesis. It also presents the research questions, the research method approach and the thesis workflow.

The second chapter reviews the related work which was done in relation to this thesis. The related work ranges from the academic papers and conferences to the review of the market (which similar systems and applications already exists for the public users).

The third chapter displays the system architecture and based on the user’s needs survey the system design, the algorithm and the system charts and diagrams, and a technology description of how the implementation was performed.

The fourth chapter includes the data reliability evaluation description – the methodology, the evaluation results and the discussion concerning the results.

The fifth chapter, describes the UX evaluation - the methodology, the evaluation results (by evaluations phases, trip planning and live experience) and the discussion concerning the results.

The sixth chapter concludes with the results and discusses the research questions and hypotheses that were presented in the introduction chapter. This chapter reflects and reviews all the steps and the work that has been done within the scope of this thesis.
The seventh chapter presents summarized conclusions, and the limitations of the work done and suggests new lines of research to be explored in the future.

The eighth chapter contains the references used in this thesis.

The ninth chapter contains the appendices and the relevant materials, which were used for the evaluations and analysis.
Chapter 2

2 Related Work

2.1 Tourism and Technology

First, a better perspective of the potential needs and an overview of what has been done in this field of research before was needed in order to better inform the initial steps related to the research to be conducted. Reviewing the connection between social media and its use in tourism planning, revealed that only 8% of the travelers claimed that social media has an important role in their planning of trips (Flash Eurobarometer, 2015).

The term for using technologies in tourism and travel is “smart tourism”. It describes the use of the interconnection and synchronization between travel technologies (Gretzel, Sigala, Xiang & Koo, 2015; Huang, Goo, Nam & Yoo, 2017).

Several studies explore the potential of using digital technologies and social media as a marketing tool (Zeng & Gerritsen, 2014). In the review of Zeng & Gerritsen (2014), they have identified that sharing information in social media concerning traveling, may help tourists to plan their trip and affect their decision making.

Munar & Jacobsen (2014) reviewed the motivation of users’ sharing tourism experiences in social media. They categorized the users’ motivations into three types: individual action and personal cognition (e.g. as a personal documentation or a diary), self-centered motivation (others who acknowledge your knowledge and experiences), and community-related motivations (helping others) (Munar & Jacobsen, 2014).
Kah and Lee (2014) examined the “extend traditional theoretical and practical approaches of information behavior to spontaneous travel behavior, by incorporating the use of travel technology during trips” (Kah & Lee, 2014, p.669). Some of their analysis included the term “information behavior”. This describes information seeking, usage (active and passive) of the information resources and channels (Kah & Lee, 2014; Wilson, 2000). They have concluded that exposing the travelers to new information by technology, will lead to more unexpected visits, and will change the trip plans (Kah & Lee, 2014).

The field of GIS (Geographical Information System) uses social media data for innovation solutions presented in a map for end users. In the tourism industry the main uses for GIS are to provide digital files for internet and mobile mapping, attractions mapping and web sites interactive maps (Jovanović & Njeguš, 2008).

GIS is a powerful instrument for social aspects, which became a “media for constructive dialogue and interaction about social issues” (Sui & Goodchild, 2011, p. 1738). For example, organizing a mapping party in which users work together to map roads (Sui & Goodchild, 2011). These social aspects also addressed the social media location-based features, such as seeing a friend’s physically current location on interactive maps (Sui & Goodchild, 2011). This can help real time monitoring and decision making for the GIS users (Sui & Goodchild, 2011).

In the past, the collection of information related to geospatial data required expensive efforts, and was the majority of the cost for a GIS project (Longley, Goodchild, Maguire & Rhind, 2010). Nowadays, with Google Maps and similar open source data, this factor is no longer an obstacle. The geospatial data can be related to locations from social media: GIS solutions which use social media data can support a better understanding and mapping of human behaviors (Stefanidis, Crooks & Radzikowski, 2013). In the research of Stefanidis, Crooks & Radzikowski (2013), location-based tweet collection was used to predict a location as a “hot spot” based on user reports (e.g. news event). They analyzed the collection based on a wide geographical area and gained information related to the content of those tweets (Stefanidis, Crooks & Radzikowski, 2013).

Another important aspect to take into consideration concerns the navigation between touristic attractions in the same city/area and the current support supplied by different applications or websites.

As an example, the multinational technology company, Google, have implemented in its applications and websites the option in which users
can look for sites around their current location, and can initiate a navigation to one of the suggested sites, by using Google Maps (Egger, 2013).

Another implementation of a technology within the scope of tourism navigation was done in Italy by Borrego-Jaraba, Ruiz & Gómez-Nieto (2011) as part of a research project on using “smart posters” with NFC technology. Within this implementation, the user reaches a “smart poster” posted in the streets of the city, and by using NFC on their mobile phone, they can get data information about their current site, and could initiate a navigation to a close by site. This implementation only relies on the NFC data and does not use GPS (Borrego-Jaraba, Ruiz & Gómez-Nieto, 2011).

Self-navigation (traveling without the assistance of a formal tour guide) by tourists became more popular in the last years (Wu, Liang & Liu, 2013). Based on a survey conducted in the US - 81% of the users are now using their mobile phone for maps and directions while traveling (Keynote/Adobe, 2015).

Navigation between points by planning the shortest path is a well-known problem with applications beyond the tourism scope. The famous example is of the traveling salesman problem (TSP), which uses algorithms to solve the calculation for the shortest path between several locations (multi-spots) (Xu & Du, 2009). However, the classic solutions to the TSP problem do not address waiting times at a location/spot or crowdedness issues. These two issues can affect the actual time estimated for the route between touristic spots and influence the satisfaction of the solutions provided when consideration the potential application in tourism.

Navigation features can be used in the tourism filed, not just for reaching a certain location, but also within a game setting. This could be done by using a method of exploring a new place with a game technique such as hide and seek. The navigation part in this game can be done through the use of mobile phones (Eguma, Izumi & Nakatani, 2013). Another possible game theme for smart traveling, was done by allowing the user to resolve tasks in an open space. In addition to the problem solving, navigation features were included (Hung, Hsu & Weng, 2013).

Using data from other web sources and combining it with navigation output has been done before. One example can be found in the research of Pippig, Burghardt & Prechtel, in which they have combined data from Wikipedia and photos' platforms API’s, to show users' selections and interests. The output of this implementation, was a map that supplied
users with optimal route planning visuals (Pippig, Burghardt & Prechtel, 2013).

Another example of using web data sources for navigation aids is the possibility to suggest optimal paths and include sightseeing stops on the way to a destination (Sakaguchi, Izumi & Nakatani, 2013). This implementation proposed an opportunistic sightseeing navigation platform, based on retrieved data from photos platforms (such as “Panoramio”), in which a route will be generated based on the users mood (Sakaguchi, Izumi & Nakatani, 2013).

Regarding the use of traveling websites, Hsieh, et al., (2012), suggest that users seem to prefer to use blogs or social sites instead of traditional traveling websites, while they would also like to share their own individual travel experiences.

Kaplanidou and Vogt (2006) evaluated the users’ perceptions of tourism websites and its usefulness for trip planning. They found that accessibility, ease of navigation and content in the websites have an important impact on the formation of users’ perceptions (Kaplanidou & Vogt, 2006; Huang, Goo, Nam & Yoo, 2017).

2.2 Market Search and Survey

2.2.1 Crowdedness as an Issue

Approaches and business solutions concerning crowdedness in public spaces, were reviewed during the research phase. For example, smartphones provide the user with the possibility of informing themselves concerning the current crowd conditions of a certain sight or location, and could suggest directions in order to avoid this (Sassi, et al., 2015).

A main goal for solving those issues, would be to influence the behaviors of users in the long term in aspects of minimizing the visit time per site, or to balance the crowd at a sightseeing site (Sassi, et al., 2015).

The use of big data from social media for solving crowding issues, was researched. In the research of Wood, Guerry, Silver & Lacayo (2013), the data usage was not for planning a navigation route, but for getting a better perspective on visitors’ preferences at tourist’s attractions (Wood, Guerry, Silver & Lacayo, 2013). This research was based on Flickr photos geolocations/geotagging, and the researcher collected data related to 836 tourist’s sites, and found a correlation between the results.

1 https://www.panoramio.com/
in national parks (Wood, Guerry, Silver & Lacayo, 2013). In this research they discussed the limitations for using Flickr, and offered a better data collection to be provided by Twitter (Wood, Guerry, Silver & Lacayo, 2013).

Another method for estimating crowdedness can be based on people personal reports and estimations. This method was performed in the research of Goncalves, Kukka, Sánchez & Kostakos (2016), where the visitors of a restaurant were asked to report their personal waiting time in a kiosk stand before and after waiting in line. This approach had limitations such as technical problems, visitors’ motivations for reporting, and wrong time estimations (Goncalves, Kukka, Sánchez & Kostakos, 2016).

Avoiding crowdedness is not only beneficial for the tourist’s personal experience, but also in the context of preventing casualty accidents (Zhao, Qiu & Deng, 2016). In the research of Zhao, Qiu & Deng (2016), the crowd density was calculated, and the areas were classified with the values of safe, dangerous and alarming. This was calculated by accessing the users’ location from their mobile phones within an implementation of a GIS solution for the visualization of the results (Zhao, Qiu & Deng, 2016).

There are some sightseeing locations owners and random applications developers, that have identified crowdedness as a problem, and they provide the users with information about days and times when usually there is a crowded audience to be expected at the tourism site. An example for one of these applications is: “is it packed?” - This website offers the users information about the crowdedness of theme parks. It is relies on Twitter data, users’ data and other information which is not mentioned on the public website visitors (made for the end-users). The users can access a calendar which relays the estimated crowdedness in the theme parks. In addition, latest pictures taken at that location as well as latest tweets tweeted, and ask questions in a forum, are made available on this webpage.

Some theme parks have a mobile application for navigating the park. Disneyland, for example, has developed an application called “My Disney Experience - wdw”, this application allows the users to access the real-time waiting times, navigating (by using GPS) to attractions near the user and getting information about the theme park. Another mobile application for amusement park waiting lines if for the

“Liseberg”\(^4\) park in Sweden – this application helps the visitors to avoid queues and provide information. Both of these applications have inside information for the waiting lines times, and they do not rely on outsource information such as social media.

An additional way of estimating the crowdedness of a place is to use cameras to analyze the presence of people in an open space (Sassi et al., 2015). This kind of approach, which tries to give users information about the crowdedness of an area, is used in the application called the “Placemeter”\(^5\). This application does not rely on individuals’ information, but on cameras and reading images technologies. This technology uses images from individuals and street security cameras in private and public places, in order to estimate the crowdedness in the streets, shops and attractions. Based on their technology for video analysis, they offer a business solution for this issue.

Google also provides information about the crowdedness in restaurants and business stores\(^6\). The search engine can give information about the sites recommended visiting times, however Google has not provided any information about how the data collection is being done, and what factors are being used for providing this information (Reed, 2015).

2.2.2 Trip Planning

There are many applications that provide services for planning trips as well as navigational services. The number of these kinds of applications is increasing rapidly, and more features are implemented on a daily base. Following a few chosen and leading applications will be presented:

“Tripomatic”\(^7\) (cf. Figure 2, top) is a website, which provides the user with the option of planning a trip. It takes into consideration the estimated visiting time of each sight-seeing location and the shortest navigation between those. However, it does not take into consideration the trips starting time or the opening hours of the sites. “Tripomatic” is also accessible on a mobile device.

“RoutePerfect”\(^8\) (cf. Figure 2, middle) offers the user with the best route for traveling between cities, by defining locations, vacation type, number of nights and trip budget. This website provide booking services

---

\(^5\) [https://www.placemeter.com/](https://www.placemeter.com/)  
\(^6\) [https://plus.google.com/+google/posts/QY1c97V25Tz](https://plus.google.com/+google/posts/QY1c97V25Tz)  
\(^7\) [http://www.tripomatic.com/](http://www.tripomatic.com/)  
\(^8\) [https://www.routeperfect.com/](https://www.routeperfect.com/)
for the trip accommodation and suggest booking the accommodation as a package based on the planned route the user created in this website.

“Triposo”⁹ (cf. Figure 2, bottom) is a mobile application which offers the user information about the popular places. This information includes: sightseeing, restaurants, hotels and navigation. This application is commercial and offers the user to buy entry tickets (“skip the line” tickets), book an accommodation, and order attractions such as day trip for the selected city. Users can write a review on each attraction and share this information in social media.

![Screens from the example trip planning websites/applications](https://www.triposo.com/)

---

⁹ [https://www.triposo.com/](https://www.triposo.com/)
Chapter 3

3 Initial Requirements and System Architecture

3.1 Initial Requirements: User Needs Analysis

Before developing the prototype for estimating the crowdedness at a site, a user survey was conducted for collecting the user preferences and behaviors regarding tourism and mobile technologies. For gathering information about the use of mobile phones in tourism, a Google form survey (see Appendix A, 9.1) was distributed through sharing it on Facebook, through E-mails and the application WhatsApp. Questions concerning tourism habits, mobile phone usage, tourism information seeking and popular mobile features usage were posed. In addition, some demographic questions were posed to the participants.

In total 81 responses were collected. The participants were between the ages 19 and 52. The average age was 28 and the median age was 27 (cf. Table 1).

<table>
<thead>
<tr>
<th>Demographic Variables</th>
<th>No.</th>
<th>~%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>Female</td>
<td>51</td>
<td>63</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 26</td>
<td>37</td>
<td>46</td>
</tr>
<tr>
<td>27-35</td>
<td>34</td>
<td>42</td>
</tr>
<tr>
<td>&gt;36</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Nationality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Israeli</td>
<td>43</td>
<td>54</td>
</tr>
<tr>
<td>Swedish</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>German</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Italian</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>French</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Japanese</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other European Countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>America</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Africa</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>N/A</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1, survey participants’ data
Most of the responders (96%) have traveled abroad at least once in the past 3 years, therefore the collected data can reflect the general use of smartphones in tourism in today’s technology and the users’ can represent the traveler’s community.

The majority of the responders travel at least once a year or more (78%), and 77% claimed that they have collected data about the destination in “most of the cases” and “always” during the trip itself (cf. Figure 3).

As to obtaining the information for the trip, 86% of the participants are using their mobile phones as an information resource.

The most used and searched information, which a tourist is looking for on their mobile phone, is navigation (52 votes, 64%), up next are the recommended attractions (49 votes, 60%), sites opening hours (48 votes, 59%) and sites descriptions / history (48 votes, 59%) (cf. Figure 4).
The responders were also asked “What do you find as the most important features while navigating/traveling between the tourism sites?” and 58% of the responders mentioned that they find the fastest path by time as the most important feature in a navigation application. Also 46% state that the shortest walking path by KM/miles is the most important feature.

As part of this thesis's motivation for using social media and crowd data, the participants were asked about their behaviors on social media platforms regarding sharing information. Almost all the participants have shared information during their trip on social media platforms: only 6% stated that they did not share anything during the trip itself (not using the check-in location provided by Facebook, sharing photos or statuses). In Figure 5 the results of the questionnaire for the questions “What are your habits regarding posting and sharing in social media while traveling?” are presented. It is divided by the type of information that is being shared and how often. The users’ privacy settings for their private social media account are not displayed within Figure 5. Therefore, the willingness to share (private) information can only be used as an indicator but it does suggest that the number of users sharing their location in real time is growing.

![Figure 4, Information users collects by their mobile phone while traveling](image)
This information extracted from the user needs survey, helped designing the application and provided an added value to the motivation for developing such an application for the tourism crowdedness navigation.

Based on the survey and the according evaluation the following requirements for the prototype have been formed:

- All of the participants have stated that they collect data before and during the trip – therefore the application should be easy to use during the trip as well as ensuring that the user’s current situation such as current location is factored in.

- Navigation and recommended attractions are the most useful features for gathering information with the mobile phone while traveling.

- Most participants are interested in the shortest path to a sight-seeing location by time.

- The majority of the responders (94%) share traveling information on social media, which can be an indicator concerning their current location and can be factored in for the crowdedness in the tourist’s attractions.
3.1.1 Prototype Implementation

The application targets users who travel by foot between sight-seeing attractions as long as they are within walking distance (based on the user personal definition to a “walking distance”, mostly referring to a one hour walk).

In the process of developing the prototype, two versions of the application have been developed and implemented. The first version is based on the user analysis survey, literature review and the author’s personal background. The second version had minor improvements for drawbacks which became obvious during the data reliability evaluation.

The first version was used in the evaluation of the data reliability (see chapter 4), and the second version was used in the evaluation of the user experience (see chapter 5).
3.2 System Architecture & Algorithm

3.2.1 Tweet Data Analysis

For collecting social media data, social media platforms offer developers an open API REST calls, for gathering information from their database. In case for Twitter, several API REST calls were performed based on the sight-seeing coordinates radius (up to 2 KM radius from the center of each sight-seeing). The tweets were collected in real time – each time a tweet was posted during the selected time frame, it was saved in a DataBase. The total number of tweets collected were 5800 tweet in Barcelona (reflecting 14 sight-seeing, during active hours), and 1880 tweets in Rome (reflecting 6 sight-seeing, during active hours). Further details about this technology will be review in chapter 3.3.1.

After collecting the data, data analysis was performed. Sites opening hours and days had to be considered, to know which tweets had to be counted and calculated and which needed to be excluded. In Figure 6 a visual graph about the tweets collection results distributed by the hours of the day is shown.

![Tweets Distribution Example](image)

Since the collected tweet data had different values and counts between the sight-seeing locations (some provided a big collection of data sets and some only had a small data), each site had to be analyzed separately.

In order to make the analyzed data reliable, a normalization of the results for each location was performed. For each site the highest result per hour was used as the highest value (See Appendix B, 9.2). For example: At one of the selected sites (based on internet data and personal experience) the maximum waiting time for standing in a queue
was three hours. In this case, the highest number of tweets in an hour was set as a three-hour waiting time, and each tweet got a “tweet value”. For example, the highest number of tweets in one hour at another site is 30, and the maximum waiting time is three hours - in this case each tweet will be valued as a 6 minute waiting time:

\[6 \text{ min} \times 30 \text{ (number of tweets)} = 180 \text{ min} = 3 \text{ hours}\]

This number will later be used for the calculation of the estimated waiting time.

For calculating the estimated waiting time at a tourist attraction by a given time, the number of tweets had to be reviewed in a comprehension to the average daily tweet number at the tourist attraction. For example, the number of five tweets during an hour could indicate a long waiting time at one site, and at another site it could indicate there is no waiting time. The minimum number of tweets for a waiting time calculation was set to the following condition: if the tweet number of the selected hour is lower than the average daily tweets per site divided by two (code snippet in Figure 7).

```
function getWaitingTime($currentPoint, $weekDay, $hourTime)
{
    $query = "SELECT * FROM TweetHour WHERE Site_name = 'CurrentPoint' AND DayOfWeek = 'WeekDay';"
    $result = mysql_query($query);
    if ($result)
    {
        while($row = mysql_fetch_array($result))
        {
            if ($hourTime == 22)
            {
                $timeUnit = $row['TweetValue'];
                $HourTimeNew='h'. $hourTime;
                $timeCalc=(($row['h_08']]+$row['h_09']+$row['h_10']]+$row['h_11']+$row['h_12']]+$row['h_13']
                +$row['h_14']]+$row['h_15']]+$row['h_16']]+$row['h_17']]+$row['h_18']]+$row['h_19']
                +$row['h_20']]+$row['h_21']]+$row['h_22']]+$row['h_23'])/24;
                $TweetNum=$row['HourTimeNew'];
                if($TweetNum<$timeCalc)
                {
                    $waitingTime="00:00:00";
                }
                else
                {
                    $time=explode(":\",$timeUnit);
                    $sec=$time[2]*60;
                    $sec=$sec+$timeNum;
                    $diff = $sec;
                    $waitingTime = sprintf(’%02d:%02d:%02d’, ($diff / 3600), ($diff / 3600 % 60), $diff % 60);
                }
            }
            else
            {
                $waitingTime="00:00:00";
            }
            return $waitingTime;
        }
    }
```

*Figure 7, code snippet calculation of the estimated waiting time*

Previous data concerning waiting time was not available for some of the attraction sites. Furthermore, some sight-seeing locations are considered to be “open” sites, where there are no queues. With a
situation such as those, assumption has to be applied for the waiting time and algorithm.

Open sites were presented as indicators for crowdedness of a site - for example if the beach is crowded, then the waiting time indicates the estimated time for catching a spot to lay on the beach or sit in a restaurant. Collected data from a sight-seeing location during closing hours were removed, since this could not indicate the crowdedness of the location.

Collected data from a sight-seeing location during closing hours were removed, since this could not indicate the crowdedness of the location.

3.2.2 System Architecture

The system architecture (Figure 8) is composed by the following layers:

- The Services layer that get requests from Google Maps API
- The application layer that is used by the delivery mechanism of the browser and addresses the DB resources. After the user select its preferences for sight-seeing locations and travel time, a server request is sent to the Resources layer for the database request.
- The Resources layer that includes the database, containing information about the sight-seeing data (coordinates, visit hours), and the tweets analysis and collection (number of tweets per hour, tweet value).

After getting the proper information based on the user request, this information is implemented in the calculations for optimal path algorithm.

Next, after obtaining the optimal path from the algorithm, a request is sent to the Google Maps API service and a map is presented to the user in the web application.

![Figure 8, System Architecture Layers](image)
3.2.3 Algorithm

For calculating the best path, considering waiting and walking time, a calculation algorithm was implemented. The basic calculation is a sum of the interests’ points walking time (by using distance divided by walking speed), plus the estimated waiting time and the average visit time (cf. Equation 1).

This calculation was set as a simple solution, but it is mostly affected from the results of the tweet analysis for the estimated time aspects. This algorithm is only a “container” for values to be settled in, as it was specify in section 3.2.1 (the tweets values analysis). This algorithm only sums up and calculates time values based on the results from the tweets analysis.

\[ d_x = \text{distance}, s_x = \text{speed}, t(w)_x = \text{waiting time}, t(v)_x = \text{visit time} \]

\[
\sum \left( \frac{d_1}{s_1} + t(w)_1 + t(v)_1 \right) + \left( \frac{d_2}{s_2} + t(w)_2 + t(v)_2 \right) \ldots \rightarrow n
\]

Equation 1, Calculating the optimal path

The algorithm is based on an exact algorithm, which means that the users sight-seeing selection of locations is sorted into permutations (ordered combinations), next the algorithm checks which of the possible combination has the shortest by travelling time. This is similar to the traveling salesman problem (TSP) – however, in this case the waiting time and the visiting time, are parameters in the calculation, and not just the shortest walking time by distance. Another parameter that is included in the algorithm is the average visiting time - each location has a visiting time which indicates the time a visitor spends there.

The parameters of the waiting time and average visiting time affects the time of arrival for the next point - if the user reaches the next point during rush hour, the parameters values will change accordingly.

Each combination is summed up to the total amount of time and compared to the previous one in order to get the minimum total time (cf. Pseudo Code, Figure 9). For each point within each combination from the users selected points, some properties have been collected:

1. 'Place' - the place name for retrieving data from the database (DB).

2. 'Distance Time from Next Point' - retrieving the distance time data from the DB.
3. 'Average Visiting Time' - retrieving the average visiting time for the selected place based on different sources such as the sites website and user feedback.

4. 'Waiting Time' - the estimated waiting time based on the tweet analysis.

5. 'Arrival Hour' - calculating the arrival time based on summing up the distance, the visiting time and the waiting time, based on the local time in the city.

Each combination is summed up to the “Total Time”. This property is the one that is used for comparison to receive the combination that...
requires the least total amount of time. The system is offering the user three optimal paths – therefore three combinations are being saved and compared within the calculation.

After calculating the three optimal paths the data is presented to the user through Google Maps and the user can navigate use the Google Map application (installed on their mobile phone) to navigate to each site.

3.2.4 System Flow

The proposed system has several components. In the sequence UML diagram in Figure 10, the user actions and the system algorithm calculation is described. As it can be seen that the user in the first section, selects the relevant data from the application form (selectData()). The selected data refers to tourism site selection, trip preferences and properties. After the user selects their sight-seeing sites, the system calculates a permutations for each of the selected sites (calcCombinations()).

For the calculation, the system uses the information provided in the database server and it enters a loop until all calculations for the combinations are done (forEachCombination() -> getData() -> calc()). After the three optimal paths were calculated, the user can select one of the paths which they decide to be their optimal path (getOptimalPath()). In the next step the Google Map server generates the expected output by using the Google Maps API\(^\text{10}\).

\(^{10}\) API stands for application programming interface
Figure 10. User actions & algorithm calculation
3.3 Technology

3.3.1 Collecting Data - Twitter

For the purpose of developing the prototype, data was collected from one social media platform - Twitter\textsuperscript{11}. The collected data were “tweets” which is the textual information users post in this platform. Twitter is a public social media platform with 328 million active users\textsuperscript{12}, where users can post publicly and share their thoughts with the world.

The use of Twitter in tourism exists in several countries such as the USA, Spain, France, Germany and many more (Hays, Page & Buhalis, 2013). Twitter users can actively connect with the world in three ways (Emerald Group, 2014):

1. By including outside links in their tweets, usually shortened url’s,
2. By addressing other Twitter users either by mentioning them or “re-tweeting”\textsuperscript{13},
3. By using “hashtags”, words starting with a #, that users identifies a topic by category or a one wording description.

Some tweets might be ‘buried’ really fast which is due to the fact that 92% of retweeting and replying to tweets occurs within a one-hour period after the original tweet was posted (Hays, Page & Buhalis, 2013).

Another popular social media platform is Facebook as it currently has over a billion users\textsuperscript{14}. However, for this specific research, Facebook was not chosen as a source for the algorithm. Although it could have been used for the implementation, Facebook is used mostly for sharing information on a personal level (many of the users accounts settings are defined as “private”) in comparison to Twitter, where all the data is publicly accessible (Hays, Page & Buhalis, 2013). Additionally, for this research and prototype, Twitter data is more accessible in comparison to Facebook data from a developer perspective: Twitter API REST calls require less resources for collecting information based on the user location. Twitter platform is meant to reflect upon events that take place in the present and in real-time events as they occur (Hays, Page & Buhalis, 2013).

\textsuperscript{11} https://twitter.com/
\textsuperscript{12} https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/
\textsuperscript{13} “re-tweeting” mean’s that one user is sharing / tweeting another person’s tweet
\textsuperscript{14} according to Facebook’s own statistics they had 1.23 billion daily active users on average for December 2016” (https://newsroom.fb.com/company-info/)
As mentioned, the selected social media platform for the data gathering is Twitter. The application prototype database is based on data gathered during the time period between June 19th, 2015 and July 3rd, 2015 which means it covers a time frame of 14 days.

For the prototype data was collected for two of the most touristic countries for Europeans - Spain (20%) and Italy (20%) (Flash Eurobarometer, 2015). For the evaluation process, the use of history data collection was chosen. By doing so, the evaluation would not be affected by other parameters (such as the possibility of reflecting upon a day for the evaluation without any tweets or data). For the future work beyond the scope for this thesis, the data could be in relation to real-time Twitter feed through the use of the Twitter API combined with the history data collection and analysis. Through this the calculations would become more reliable and might reflect upon the current situation better.

3.3.2 Programming Language & Database

For developing the prototype, a web application was created based on HTML5\textsuperscript{15}, CSS3\textsuperscript{16}, jQuery\textsuperscript{17} and PHP\textsuperscript{18}. For setting the web application to be responsive between the different mobile, tablet and desktop devices, a CSS framework “Bootstrap”\textsuperscript{19} was used, which can be easily implemented to design responsive interfaces using the Cascading Style Sheets language.

In order to save system time and resources, and since the suggested sites for the user are from a limited list, some of the data was previously saved in a MySQL\textsuperscript{20} database (cf. Figure 11):

1. The distance between the suggested sites (by time and KM/miles) - in order to avoid Google API calls for calculating distances. This saves system resources.
2. Site data information such as longitude and latitude, opening hours, average visit time, off work days.
3. Tweet analysis values and numbers saved by days of the week and hours from 8AM to 11PM, due to the relevant of the crowdedness hours.

\textsuperscript{15}Hyper Text Markup Language 5
\textsuperscript{16}Cascading Style Sheets 3
\textsuperscript{17}Cross-platform JavaScript library \url{https://jquery.com/}
\textsuperscript{18}Server-side scripting language \url{http://www.php.net/}
\textsuperscript{19}\url{http://getbootstrap.com/}
\textsuperscript{20}Structured Query Language Database \url{https://www.mysql.com/}
3.3.2.1 Google Maps API:

Many of the services for calculating the path between the sight-seeing locations rely on Google Maps API. Google Maps provides information and calculations between destinations, based on distance and time. However, relying on Google Maps API might have limitations such as the source of calculation: Google not providing information about “how” the result was calculated, but only setting up the final calculation output. There is no option to edit those calculations, and the easiness of the usedness in this API overcome the limitations it might has. Google Maps API also has the option to visualize the output data on a map. The data can be visualized in the web application itself with an implemented map, and can also open the data in the (native) Google Maps application. Other services this API has to offer were used, such as the navigation features Google Maps provides.
Chapter 4

4 Data Reliability:
Survey and Self-Reported Data

4.1 Prototype - Version 1

Following the requirements from the user needs survey, a prototype was implemented and tested in the data reliability evaluation.

4.1.1 Home Page

The landing page is the first page (cf. Figure 12), where the users can select the city they wish to visit or are visiting at the moment. The prototype made available the two following cities: Rome and Barcelona. The design is simple. Each city is represented through a picture of the town with a famous sight-seeing location displayed in it. This way the user can easily identify the correct city. Additionally, the user can be inspired by the photograph and the location. In addition to the picture representation the cities name is layered over it to ensure that each town can be distinguished from one another. The header and footer have the same design pattern in order for the application to keep its consistency.

Figure 12, Landing page screen
4.1.2 ‘Selecting Tourist’s Attractions and Preferences’ Page

The page that is now presented is for the user preferences selection and sites selection. On this page, the user can select an attraction from a list of popular tourist’s attractions (indoor and outdoor) (cf. Figure 14). The option to add additional properties such as: adding the user’s current location or an address as the starting point, plan the path based on other days of the week and hours is also given.

The option for adding the new preferences is collapsed in order to reduce the cognitive load while filling this screen (cf. Figure 13). This screen fits by size to a mobile touch screen.

![Figure 14, Sites selections and preferences collapse](image)

![Figure 13, Preferences expanded](image)
4.1.3 ‘Selecting the Optimal Path’ Page

After submitting the relevant data, the user will get three optimal path options for the shortest path, based on the given time frame and the selected tourist’s attractions (cf. Figure 16). The estimated time is based on waiting times at the sites and the walking time. The indicators for the crowdedness at the sites is depicted through the use of circle in the colors of green (not crowded, less than 10 minutes), yellow (a bit crowded, between 10 to 30 minutes) and red (very crowded, more than 30 minutes). The user can choose his preferred path based on the presented data and personal preferences.

In case it was not possible to calculate a path since some sites were closed during the selected days or hours, an informative message (cf. Figure 15) will be displayed to the user that one of the selected sites is closed and that the plans must be changed accordingly either through altering the initially selected sight-seeing locations or the personal preferences.

![Figure 15, informative message](image)

![Figure 16, Path choosing page](image)
4.1.4 ‘Path Output and Information’ Page

After choosing the optimal path, the next screen presents detailed information about the selected path. Each site has relevant data such: arrival time, time from the next point, estimated waiting time, and average visiting time (cf. Figure 17).

The user can use the Google Maps application by clicking on the button (“Navigate by Google Maps”) for navigating this path, and for the visual, the walking path route is displayed in a Google Map thumbnail.
4.2 Methodology

4.2.1 Participants

For verifying the reliability of the data, some users agreed to participate in the evaluation and test the data reliability in real time in Barcelona (Spain), during the month of August 2015.

Eight users have tested the application in real time, and compared the results of the recommended path calculated by the application, with the estimated waiting time to the real waiting time.

The participants’ ages were 27 to 60 and the average age was 38 (cf. Table 2).

After getting the results from the users, the author of this thesis has decided to test the application in real time as well and compare the results to the reported ones. All results were then noted down.

4.2.2 Procedure

Each tester was asked to provide their feedback through a provided link on the web application. The feedback form was created by using the Google Survey tool. The survey included questions about the user’s demographic data such as age, grid questions about the time results from the application vs. real time (actual time), reasons for choosing one path or another and general feedback (See appendix C, 9.3).

However, during this feedback not all the tourist’s attractions were reviewed, since different users chose different paths and interests). In order to complete the feedback on all attraction sites, they were all reviewed by the researcher of this thesis.

Finally, the users recorded their answers and results in the provided survey and wrote a general feedback from their personal user experience. Since the users did not get a previous introduction to the system, receiving their feedback for user experience was most advantageous.

<table>
<thead>
<tr>
<th>Demographic Variables</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>&lt; 28</td>
<td>2</td>
</tr>
<tr>
<td>29-33</td>
<td>4</td>
</tr>
<tr>
<td>&gt;34</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Reliability evaluation participates data
4.3 Evaluation Results

The results were collected based on two aspects: end users and the researcher test. The results from the potential end users were collected between the dates of August 12th to August 17th, 2015. The results from the researcher were collected from September 13th to September 17th, 2015.

<table>
<thead>
<tr>
<th>Site</th>
<th>Number of Tests</th>
<th>percentage of correlation</th>
<th>percentage of over-estimation</th>
<th>percentage of under-estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Boqueria</td>
<td>5</td>
<td>100% (5)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Estatua de Colom</td>
<td>4</td>
<td>75% (3)</td>
<td>0</td>
<td>25% (1)</td>
</tr>
<tr>
<td>Plaça Espanya</td>
<td>4</td>
<td>75% (3)</td>
<td>25% (1)</td>
<td>0</td>
</tr>
<tr>
<td>Parc de la Ciutadella</td>
<td>4</td>
<td>75% (3)</td>
<td>25% (1)</td>
<td>0</td>
</tr>
<tr>
<td>Platja de la Barceloneta</td>
<td>3</td>
<td>100% (3)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Museu Nacional</td>
<td>3</td>
<td>0</td>
<td>100% (3)</td>
<td>0</td>
</tr>
<tr>
<td>Plaça Catalunya</td>
<td>2</td>
<td>0</td>
<td>100% (2)</td>
<td>0</td>
</tr>
<tr>
<td>La Font Màgica</td>
<td>2</td>
<td>100% (2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Catedral de Barcelona</td>
<td>2</td>
<td>0</td>
<td>100% (2)</td>
<td>0</td>
</tr>
<tr>
<td>Teatre del Liceu</td>
<td>2</td>
<td>50% (1)</td>
<td>50% (1)</td>
<td>0</td>
</tr>
<tr>
<td>Torre Agbar</td>
<td>1</td>
<td>100% (1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>32</strong></td>
<td><strong>66% (21)</strong></td>
<td><strong>31% (10)</strong></td>
<td><strong>3% (1)</strong></td>
</tr>
</tbody>
</table>

100% connection between the application and the waiting time in real-time.

100% non-connection between the application and the waiting time in real-time.

Table 3. All the recorded results concerning time
Some of the results correlated – meaning that the application managed to analyze and calculate correctly the estimated waiting time for each site. However, some of the results did not correlate, due to multiple parameters, to be elaborated in the following.

In this evaluation a total of 14 sites were tested. However, three sites in the application data could not be analyzed and calculated like the rest. This happened even though tweets were collected in the analysis part and the navigation application did offer the users the option to navigate there. The sites were: the “must see” attractions of Gaudi: Sagrada Familia, Casa Mila and Parc Guell. These mentioned sites require a pre-booking reservation: the visitors need to use the internet or an automatic machine for purchasing tickets and there are no waiting lines.

It became apparent that the Barcelona municipality became more efficient on those tourist’s attractions and the visitors need to book a place in advance for a specific visiting hour since the site is limited by the number of visitors per hour/day. Therefore, these sites could not be estimated by crowding or to be considered as part of the navigation application.

As can be seen in Figure 18, 21 of the reports (66%) reported that the timing results had a connection between the applications estimated time and the actual crowdedness at the sight-seeing location. As for the remaining 11 reports (34%) - most of the results (10 reports, 31%) reflected an over-estimation of the application. An over-estimation means that the application estimated a waiting time, when in practice, the place was not crowded. Only one test resulted in an under-

---


estimation, in which the application estimated the place was not crowded, when in practice it was.

As can be seen in Table 3, each site had a different number of testers due to the users’ preferences. In addition, the researcher of this thesis has visited all the suggested sites in the application (in Barcelona) and compared the data from the application to the crowdedness in practice. A total of four sites did not show connection between the application and the waiting time in practice. Three sites were under-estimated by the application - Museu Nacional, Plaça Catalunya and Catedral de Barcelona. On those sites the waiting time in practice was shorter than the application calculated. One site (Estatua de Colom) was over-estimated - the waiting time in practice was longer than what was calculated. This site has an elevator for viewing the landscape from the top of the statue - the elevator can only carry two people at a time (based on the formal instructions at the site itself). Therefore, regardless of the number of people in line, the waiting time might take up to 30 minutes due to the elevator's capacity.

The participants of the survey were asked questions regarding their waiting time. When asked to elaborate on time differences between the estimated waiting times from the application to the waiting times in practice, the users have mentioned that some of the sites were open-air sites, and therefore there are no queues. However, even at an open-air sight-seeing location waiting times can occur when the location has a special place, which might be currently occupied, or they want to photograph a nice scenery.

When users were asked regarding their choice of one out of three compatible paths (which one they have chosen and why), all users have mentioned that the application did not consider their current location in the navigation path so neither of the paths was good enough. However, this option does exist in the application (the user can select the starting point as his current location or insert an address as a starting point), but in this case, all users have overlooked this option in the application. This was a good example for a quick fix for the user experience. Thus, it was concluded that this feature needs to be more accessible and visible to the end users.

4.4 Discussion

The majority of the results between the application estimation waiting time and the waiting time in practice were compatible. In order to evaluate whether there were any factors that have an impact on the
design or the participants, and were not taken into consideration, the data reliability was additionally tested in Barcelona by the author of this thesis as well. As mentioned in the results section, some sites cannot be affected by waiting times and queues from an outside parameter/factor as the social media data collected suggested. Those sites require a reservation and in advance purchased ticket in order to avoid crowdedness and to provide the tourists a better experience. By using this method, the crowdedness can be controlled, and the tourists will be able to enjoy the attraction. It is the authors believe that this method will be used more in the future for mega toured sites. For example, in Barcelona - the tourist cannot enter the tourism site without a ticket which is for a specific entry time - for example between 13:00 to 13:30. The tourist can order this ticket through the sight-seeing’s website, an automatic machine outside the site or at the cashier counter if there is one. If the researcher had not tested the data reliability personally in Barcelona, they might not have become aware of this factor for the application’s navigation.

Based on the results it can be assumed that the social media data (i.e. Twitter) could help predict whether a place is crowded. However, other factors should be taken into consideration for the data reliability as well. Factors could be the “transport” to the actual sight-seeing place as was seen in the Estatua de Colom: the location itself could contain a big crowd, but the elevator to get to the viewing level can only carry 2 people at once. Thus, one would have to wait in line for 20 minutes even if there are only six people ahead in line.

There were three sites in which the application predicted that they were more crowded than the real-time situation. Those sites should be evaluated more carefully as they indicate that other factors should be considered in addition to the tweets history. One optional solution could be an algorithm that calculates using both the ‘live’ Twitter API calls as well as the analysis of the tweets history data. As can be seen above, the application predicted successfully the waiting time for the majority of the sites.

Since all users in this evaluation had not noticed the application feature for using their current location as the starting point in the route planning, this issue affected the users’ personal feedback and experience.

They felt that an important feature was missing, and some of them expressed disappointment from the system. Their choice for getting their optimal path out of three options, could not include their current
location. Even though this feature was in fact part of the prototype, but in the presented user interface (UI) design it was overlooked.

Since the evaluation of the applications data reliability did not include the researcher as part of the evaluation group in real time (as an observer or an instructor), some information about the users’ process might be missing:

- How was the testing done in the field?
- Have the users had any issues with operating the application?
- Is the reported real waiting time accurate?

For the discussion and the conclusions of this thesis, an assumption was made that the evaluation results were accurate.
Chapter 5

5 UX - User Centered Design

5.1 Prototype - Version 2

After the first data reliability evaluation, insights about the user interface led to several changes that were implemented in the screen flow and the UI design for the second prototype. Four screens have changed: the home page, the user sites selection and preferences screen (which was split into two screens instead of one), and the path output and information page.

5.1.1 Home Page

The design of the application’s starting page was not altered (cf. Figure 19). An additional city was added. Afterwards there are three cities available within the application. The additional city was present within the UX evaluation.

Figure 19, Home page screen
5.1.2 ‘Selecting Tourist’s Attractions’ Page

Based on the users’ feedback from the data reliability evaluation, where many users have missed the personal “preferences” section (as described in the data reliability evaluation chapter), the decision was done to split the screen of “Selecting tourist’s attractions and preferences” into two: “Selecting tourist’s attractions” (cf. Figure 20) and “Preferences” (cf. Figure 21).

![Image of Selecting tourist’s attractions screen](Image)

*Figure 20, Selecting tourist’s attractions screen*
5.1.3 Preferences Page

This screen provides the user with a starting point to select either from an address or from the current location. The user can also choose a different time or day of the week for the creation of the navigation and suggestions from the application for a future trip (cf. Figure 21).

![Preferences screen](image)

*Figure 21, Preferences screen*

5.1.4 ‘Selecting the Optimal Path’ Page

Remained the same as version 1.
5.1.5 ‘Path Output and Information’ Page

In this screen (cf. Figure 22), a new feature was added. A new button below the information data for each sight-seeing site, allows the user to navigate to every sight-seeing location separately from their current location using Google Maps. The researcher of this thesis valued this feature as important while reviewing the application in Barcelona. In addition, the user can navigate all the routes by Google Maps.

![Figure 22, Path output screen](image-url)
5.2 Methodology

5.2.1 Participants

The user experience of the application was assessed by a group of five participants in January 2017. The participants (cf. Table 4) used the application for planning a trip and for using the application in practice, in an outdoors live experience.

According to the research of Nielsen (2012), it is reasonable to assume that five users can identify 80% of the insights for an evaluation, even on websites. Based on this research the evaluation was conducted with a group of five people.

The participants’ age range was between: 28 to 38, with the average age of 31. Their technological background knowledge varied from low to an expert level, based on their self-declaration toward this subject.

All participants used the web application from the “Firefox” browser.

Four of them had an Android running smartphone (operating system), and one participant used an iPhone.

5.2.2 Procedure

The users have performed two activities using the application; trip planning and navigating with the application.

5.2.2.1 Trip Planning

The participants were asked to plan a 3-day trip to Rome over the weekend (Friday to Sunday). The scenario was presented to them through the Google form they were asked to fill out within the first section (see Appendix D, 9.4). When the process was completed, the users were asked to independently evaluate their process and give their feedback via an online Google form survey.

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Age</th>
<th>OS</th>
<th>Technological background</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>28</td>
<td>iOS</td>
<td>Low</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>29</td>
<td>Android</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>29</td>
<td>Android</td>
<td>Expert</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>38</td>
<td>Android</td>
<td>Intermediate</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>29</td>
<td>Android</td>
<td>Expert</td>
</tr>
</tbody>
</table>

Table 4, UX participant’s data
The following information was sent to all participants in advance: information on the applications purpose, actions that needed to be taken during the process, and the links to the user survey as well as the web application.

The survey included questions about the demographics concerning each participant such as name, age and mobile model. It also included a scenario section with questions concerning the complete process, and a Likert grid type five-point scale with phrases about the users’ feelings and experiences towards the application (see Appendix D, 9.4). The use of Likert scale was used to measure the participant feelings, opinions and attitudes toward the experience (Likert, 1931). Since the participants were asked to rate different aspects of the experience, this method can help measuring from several response options, which best reflects the participant opinion to the statement (Gliem & Gliem, 2003).

The phrases for the Likert scale were based on the research by Ceipidor, et al. (2015) and Jang, et al. (2016) concerning UX evaluations.

5.2.2.2 Live Experience

The outdoor evaluation of the application was conducted by the users navigating the streets of Tel-Aviv. As was mentioned previously, tweets data was not collected for this city, and the application showed results based on fabricated data. Since the users did not test the data reliability, and had to focus only on the user experience, they were instructed to conduct as though the fabricated crowdedness in the application was true in practice. This evaluation included open observation, group discussion, and a user survey.

The users’ interaction with the application were observed, and the outdoor navigation with the application was performed as one group. The users’ insights and discussions while navigating were noted down by the researcher of this thesis. At the end of the navigation evaluation, a group discussion was conducted, and conclusions were drawn from the users’ experience and insights of the completed process. Finally, a Google form survey was distributed among the users (See Appendix E, 9.5).

The form contained an “after trip” section and a general feedback section (“application screens”). In the “after trip” section the users were asked about their experience using a Likert grid type five-point scale, and questions about their actions while navigating. In the “application screens” section, the users were asked to rate each screen in the application and write a short paragraph about possible improvements.
The phrases and the rating questions were based on the research of Ceipidor, et al. (2015) and Jang, et al. (2016) UX evaluations.

5.3 Evaluation Results

The results from the participants were gathered from January 11th to January 14th, 2017.

The parameters that were evaluated were both the trip planning and the live experience including the outdoor navigation.

5.3.1 Trip Planning

The participants received a scenario in which they had to plan a trip in Rome, based on a given starting point (a specific address) and a number of days.

Firstly, the users were asked to describe their process of planning a trip. All users have tried and compared different paths in order to find the best path for their personal preferences. Most users (4 out of 5) preferred to choose the path which avoided crowdedness, over the time-shorter path.

Next, the users were asked about their planning results – between all users there was no identical path. Each user chose a different path in comparison to the other testers.

Lastly, the users were asked to rate their experience by a Likert grid type five-point scale. In Figure 23, it can be seen that overall the experience was positive - average of over 4 out of 5 for each phrase. Only one user did not find the application to be as easy to plan the trip with.

![Figure 23, Rating the experience using the Likert scale](image)
All users agreed that the new information (concerning the crowdedness) which was offered to them by the application, has caused them to change their previously planned path plans. However, since the trip planning forced the users to navigate between the screens constantly (in order to try new paths and new destinations) - there were some bugs and improvements suggestions. Overall the applications design received positive feedback from the users.

5.3.2 Live Experience

In the outdoor activity three methods were used: participant observation, group discussion and a personal user survey.

5.3.2.1 Participant Observation

For this evaluation, the participant observation method was used, in which the researcher was part of the group activities (Singh, 2010). The users were all aware of the researchers’ presence. However, they were unaware and had no knowledge that they were being observed closely concerning their experience and impressions during the process.

Throughout the length of the navigation experience, the participants and their discussions were observed and evaluated. It was noted, that their discussions frequently were concerning the different options the application presented. Eventually, the shortest path that was presented by the application was the one that was chosen. They used the Google Maps application for navigating between the tourist’s attractions. Since all users had previously participated in the trip planning evaluation, they were all familiar with the application screens and the expected output. Problems to operate and use the application were not noted, and no “how to” questions were asked.

5.3.2.2 Group Discussion

Following the outdoor activity, the participants were asked to express their insights concerning the UX evaluation process, and provide a review for the application from their perspective. There was an active discussion concerning their impression of the application. All of the participants thought the concept of the application was really good, and that it could help them during planning of future trips and navigating tourist attractions in popular cities in real time.

Participants also shared their suggestions for the application and what, in their opinion, was missing from the experience. Some participants suggested to switch the order of the screens for selecting
tourist attractions and the preference page (i.e., make the preference page first and make the tourist attractions selection page second).

Some participants have suggested that a visualization of the location of the different sites in relations to one another should be added to the tourist attractions selection page. This can be done by adding a map presenting all of the selected locations so that the user will be able to thoughtfully choose attractions that are close to one another for a one day trip. Finally, the participants shared their traveling experience from other places around the world, and how this sort of application could help change their experience for the best.

![Figure 24, Users testing the application in the streets of Tel Aviv](image)

5.3.2.3 Survey

Following the outdoor activity and the group discussion, a survey was distributed among the participants regarding their experience with the application and the outdoor navigation.

All users agreed that they found it easy to navigate using the application, and that they think that avoiding crowdedness can improve their trip experience. All users have changed their sites selection and adapted their route due to the results from the application.

As for the question "what did you enjoy most?" (cf. Figure 25), only one of the user mentioned that they enjoyed planning the trip with the application more than using the application in practice. It was stated that it helps to estimate the trip duration for each sight-seeing location.

![What did you enjoyed most?](image)
The other four users who enjoyed using the application most, wrote that they found it useful and easy, and that they like the ease of use with navigating through the use of Google Maps - since it is very reliable, familiar and easy to use. They also found the possibility to change their route at any time or any point, very useful.

Next, the users were asked to give their opinion concerning the applications screens providing a rating and suggestions for improvements as can be seen in Figure 26.

The home page was rated high, with no suggestions for improvements. In the selecting tourist attractions page - the rating had an average of 3.6 (out of 5), and the users suggested some improvements. Five out of five users wanted to get information about other tourist attractions in relations to one another to be able to easily inform themselves which sites are close by. They wish to have a visual of all of the suggested sites on a map, or divided by area. Also, they wish to get information about the tourist attraction (what it is, opening hours, and historical data).

The preferences page was rated relatively low (average of 3 out of 5) and some improvements were suggested: the preferences page should precede the sites selections page. Thus, a user who wants to change their sites selection, would not be compelled to insert the same data about their preferences repeatedly. Some have mentioned they wish they could insert an ending time and that the screen is “not colorful enough as the other pages”.

The screens of selecting the optimal path, and the path result and map were rated by all users with a 5 (“love it!”), they liked the visuals and that their inserted info was bright and clear.

![Application Page Rating](image)

*Figure 26. Page rating by Likert scale*
5.4 Discussion

The analysis of the evaluation indicated that the live experience navigation had a more satisfying result concerning the user experience, when compared to the trip planning. However, they have expressed that the trip planning option is an important and useful component of the application, although some improvements should be done.

**Evaluation of the trip planning:** Some assistance was given to the participants while they were in the process of the trip planning, as it was the first time that they used the application. Therefore, the assistance of the researcher was given and the application goals and process were explained and technical help was provided. Through taking this step it also gave opportunity to gather additional feedback from the new end users.

The users were asked to give their honest opinion and true feedback and keep their opinion strictly professional, making the feedback as reliable as possible.

As can be seen within the results no chosen paths were identical between any of the tester’s. All users have planned a different trip from one another - which can indicate that there is no “perfect navigation” or “best navigation”. Every user had their own criteria for the “best path” - some avoided crowdedness, some wished to consider the shortest time or shortest distance and some wished to have a “chill” trip in which they see one or two tourist’s attractions per day instead of seeing all six sites in Rome which could be done within one day as well.

**Evaluation of the outdoors navigation:** At this point of the evaluation, no technical difficulties had arisen, since all users previously used the application to plan the trip. In this evaluation, the researcher was a part of the group and had the opportunity to observe the group’s experience and guide them while they experienced some minor problems.

The group discussion, which followed the navigation activity, was a helpful source of insight, since all users had a relevant input to provide about the experience. The users listened to one another and agreed or disagreed with each other on different subjects. The group discussion helped create an initial opinion before the Google survey and it might have provided some users with a positive attitude about the application as, during the group discussion, some of the users talked about it in a positive manner.
Within the survey section, all users mentioned the application could improve their trip experience. These thoughts were mentioned in all of the UX evaluations and it seems that this sort of application is missing from the market. The users in the UX evaluation were told that all of the presented data was true and were asked to ignore the real-time results. By doing so, the users could actually test the use of such an application and not be negatively influenced or disappointed by a “waste of time” or a “wrong trip recommendation”.

Navigation through the use of the Google Maps application has proved to be very successful. Users have mentioned their dislike of traveling applications that compel the use of navigation developed by the different providers of smartphones and operating systems as they tend to have bugs and produce mistakes for the user. The Google Maps application is popular and reliable, therefor easier and more comfortable to use.

Most users had the Google Maps application installed except for the iPhone user - this user had to install the application especially for the evaluation.

The analysis of the ratings on the application screens revealed that, the lowest rated screens are the ones where the user had to insert their initial trip data. All users wanted more information on those screens and wished that they showed more visual elements. A visual element could be, for example, a map for showing all the tourist attractions locations at once.
Chapter 6

6 General Discussion

This thesis investigated the potential of using social media data and data analysis algorithms in a software implementation within the field of tourist’s attractions and crowdedness. The results presented in this thesis, are very extensive. It starts with a literature review and explores similar products in the technological tourism market. Next, a description of the system architecture is prepared and a working prototype implemented. Lastly, two evaluations were performed - one regarding the data reliability aspect and one concerning the user experience aspect.

In the related work chapter, some of the technologies used in the tourism field were presented. Kah & Lee (2014) claimed that new information presented to the user while traveling will change their plans regarding the planning of a trip - the results reported in this thesis suggest the same. During the reliability and usability evaluations, the users changed their planning of the trip and chose a path less crowded and with a shorter waiting time.

The users’ behavior in both evaluations, supported Wilson (2000) suggestions for the “information behavior” which exists in tourism. The “information behavior” within those evaluations is referring to exposing the users to new information (the list of the tourist’s attractions) through the application. This creates awareness for more sightseeing locations and the users choose to visit more sight-seeing locations based on the new information presented to them. This increased their traveling satisfaction both during the trip planning phase as well as during the outdoor live experience evaluation. The new information supplied lead them to more unexpected visits and new places to explore just as Kah & Lee (2014) stated.
In the related work chapter, a summary of the users’ motivations for sharing tourism experiences on social media was reviewed, and three motivation categories mentioned: individual action and personal cognition, self-centered and community-related (Munar & Jacobsen, 2014). The third category, which is helping others (a community-related motivation), can be integrated within this applications purposes: this application reliability is relying and effected by the number of tweets – the more users share their location in real-time on social media, the more accurate the algorithm calculation will be. If the users who have shared their current location on Twitter in real-time, would know they are ‘helping’ other tourists for the estimation of the crowdedness, this motivation could increase the tweet numbers per location and more Twitter users will have the motivation for sharing this information.

All of the users’ that participated in the evaluations, were well aware of the concept of “Self-navigation” (Wu, Liang & Liu, 2013) and had previous experience with planning and navigating their own preferred path using mobile applications.

The users preferred the use of the navigation feature they were familiar with (Google Maps) in the application, due to the easiness of use and familiar content, similarly to the results of Kaplanidou & Vogt’s research (2006). In addition, the application was designed with applying a user-friendly structure and navigation between screens. This affected the users’ perceptions of the application in the evaluation of the planning of the trip, leading to positive feedback, just as suggested by the research of Kaplanidou & Vogt’s (2006).

In the introduction chapter, four research questions were presented to frame the explored subject of implementing social media data in algorithms for crowdedness estimations.

The first research question, “What are the current habits and preferences of tourists, for the use of mobile phones as an information source while traveling?” was answered in the different phases of the thesis. This was initiated with a self-reported survey in which the users reported about their habits and preferences towards using a mobile phone while traveling. Most of the responders (52, 64%), in this phase mentioned they use their mobile phone while traveling mostly for navigation between the tourist’s attractions and gathering information about those tourist’s attractions. During the evaluations those habits and preferences were described in detail, and the users were observed by the researcher of this thesis. The users also answered the surveys about their real-time behavior, their preferences concerning the present and the future use of the application. The users enjoyed using the application
and appreciated that the application uses a navigation feature through the Google Maps application – the users appreciated the familiar use with an application they already had experience with.

The mobile phone is a resource while traveling nowadays and the use of it was natural to all of the users. Additionally, the users expected the application to provide more information and features regarding traveling such as: information about the attraction, opening times and the options of suggesting a trip plan based on several days and selected tourist’s attractions.

The results concerning the second research question, *Does providing new information to tourists’ concerning crowdedness and rush hours, have influence on their plans and does it change their plans?* seem positive. In the live experience during the UX evaluation study, the users changed their planned paths when they were presented with information and estimations concerning a certain attraction to be crowded during the selected time frame. In addition, the application exposed the users to new tourist’s attractions which they were not aware of before. As for the influence of crowdedness estimations and rush hours on the changes of a trips plan, this was just partially observed. Some participants really loved the concept of avoiding crowdedness in an attraction, but some of the testers found the easiness and distance between the locations more important. For example – even though an attraction was crowded, they wished to visit this attraction during the morning hours of the day, where they needed more patience, in comparison to the evening time. Some wished to visit the attractions ordered by their distances, so they would need to walk less and some just preferred to see an attraction before the other one based on personal preferences. The avoiding of crowdedness, was also a factor in the planning of the navigational path and decisions taken, but it was not the only one.

The third research question, *Which features should a mobile application have, in order to support tourists’ activities when trying to plan sightseeing routes in cities?* was also addressed in this thesis. A user centered approach was applied for the implementation of the application. For designing this mobile application, the results from the user’s needs survey were taken into consideration while developing the prototype. Eventually, the evaluations of the implementation revealed insights that were not taken into consideration in the user needs survey stage, and the results from the evaluations helped the research to be more user oriented. The features, according to the surveys and the evaluations, which an application of this sort should contain are:
• Flexibility for the users planning of the route - the users should be able choose different sites and change their selections easily and insert initial data for the route planning as a starting point and date,

• A list of the best and popular tourist’s attractions,

• Visual design of the output for increased user experience,

• The importance of presenting data reliable (the suggested route should reflect the actual waiting times).

It was helpful to have the users involved in the process of planning and implementing and not just evaluating the system. This was perceived as a big contribution to this study.

The fourth research question, "Can social media data be reliably used in algorithms for measuring and estimating time aspects and crowdedness?" was answered with the help of the data reliability evaluation. The results suggested that this indeed was the case. More specifically, collecting social media data concerning a certain location, with the right and accurate analysis, can predict if a place is meant to be crowded during certain hours. However, it cannot be the only factor as will be detailed in the “future work & limitations” chapter. The implementation of such data in a new application, is easy to follow-through with as the API’s offered by the social media platforms are well documented and have an extended feature list. Based on the data reliability evaluation, the majority of the results had a connection with comparing the applications output to the real-time crowdedness. This data might become more accurate in future times, if the younger generation will share even more information with others in social media platforms, and the social media data numbers will increase, potentially reflecting crowdedness.
Chapter 7

7 Conclusions, Limitations & Future work

7.1 Conclusions

This thesis addressed two important topics: (a) testing the potential of using social media data for the estimation of crowdedness and (b) evaluating the users’ experience towards such a solution. Both of the evaluations presented positive findings.

The data reliability evaluation showed a connection between the users’ reported results and the estimated waiting time in practice – which suggests that this kind of information can predict crowdedness. However, the tweets time analysis could be improved by reviewing and analyzing each tourist attraction separately.

The UX evaluation findings reveal that the users had a positive impression of the features and design of the prototype. The users enjoyed planning their trip with the application and found the suggested navigation route feature useful.

7.2 Limitations

This study has several limitations. During the prototype development, it may have not contained all of the possible application features suggested by the users in the user’s needs survey. The possible features to improve the application were large by number, and would probably have been developed in vain since they did not contribute to answer the research questions. However, the possible end users, wished
to have more possible features, such as a suggested route, information about the touristic site, the missing of these features had a negative effect on their user experience.

Some limitations are concerning the data collection numbers from Twitter. The tweets were collected based on a certain sight-seeing longitude and latitude coordinates location. However, not all user’s accounts allow twitter to expose their current location (“geotagging”). It is not a default setting by Twitter\textsuperscript{23}, and users’ needs to activate it on their own. This can also affect the numbers and the waiting times estimations for the calculations.

Regarding the data reliability evaluation, one of the limitations was that the information was evaluated only using history data. By using real live data, gathered through a longer data collection period (e.g. data to be collected over an entire year and not just in a week or two), the results might have been more accurate. There is also the chance that the results were correct due to a coincidence – the results were not tested during an entire week and/or a complete day. It was only tested by different individual users’ each time, during a specific time which was based on the users’ choice and convenience. If the results would have been observed and watched for a whole day, maybe then the results had not been as positive or could have been more mismatched.

In addition, the number of participants in the evaluations could have been larger as only 13 users participated in both evaluations (user experience - 5; data reliability - 8). This is especially critical for the reliability evaluation, for which a larger number of users testing the application during different times of the day could have improved the reliability of the results.

7.3 Future Work

As for future work, every aspect of thesis can be further extended and developed. The data collected data was only from Twitter. Collecting data from other social media platforms can provide higher accuracy and richness. As for the time duration of the collected data – the collection period could be extended for a longer period of time.

The prototype is currently developed to collect data from the systems database and perform the calculations by using this data. This is helpful for future use and development, so that only the data analysis part of the implementation will have to be updated while inserting other factors

\textsuperscript{23} \url{https://support.twitter.com/articles/78525}
to the analysis process. Other factors based on the specific attraction, beside the social media data, need to be considered for calculating the estimated time results (for example ordering tickets in advanced or capacity limitations at the sites transport such as the elevator capacity at the Estatua de colom attraction). Each attraction should be examined in detail in order to make the proper adjustments for its own limitations. By reviewing each attraction on its own, the data created and available will become more reliable.

Based on the mentioned suggestions for the future work, this require more resources which can be achieved by addressing governmental innovation funds such as “VINNOVA” in Sweden, or “The Israeli innovation authority” in Israel.

The results emerging from the studies conducted in this thesis suggest that social media data can be used in algorithms and to be implemented in other algorithms for different research purposes. In this case, the social media data has estimated crowdedness in tourist’s attractions, but the data can be used for other new purposes as well, for example, to inspire researchers and for developers to gather new ideas.

24 http://www2.vinnova.se/en/
8 References


Wilson, J., & Becken, S. (2010, November 24–26). Who has the most sunshine? The battle for the most attractive destination. Presented at the Auckland NZ Tourism Hospitality and Research Conference, Auckland.


Appendices

9.1 Appendix A
Tourism Habits and the Use of Mobile Phone

Tourism habits and the use of mobile phone
*Required

Gender *

Age *

Nationality

Have you traveled abroad in the last 3 years? *
- Yes
- No

How often do you travel abroad? *
- Once a year
- More than once a year
- Less than once a year
Tourism habits - How often do you collect data about the destination... *

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>In some cases</th>
<th>In most cases</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before the trip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the trip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After the trip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you use your mobile phone while traveling as a resource for information about the destination? *

- Yes
- No

Your source for your mobile network connection while traveling *

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>In some cases</th>
<th>In most cases</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data (3G/4G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wi-Fi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How do you prefer to get information through your mobile phone about tourism sites? *

- Download applications about the destination
- Search Information in the search engines
- Other: _______________________

What type of information do you mostly look for through your mobile phone while traveling? *

Up to 4 marks

- Sites description / history
- Navigation
- Users reviews
- Opening times
- Prices and discounts
- Special events
- Recommended attractions
- Other: _______________________

What do you find as the most important features while navigating/traveling between the tourism sites? *

Up to 2 marks

- Sites capacity at the moment
- Shortest walking path (by km/miles)
- Fastest path (by time, include public transport)
- Sites based on personal preferences
- Beautiful/interesting walking path/route
- Other: _______________________

Please name the mobile applications you know regarding tourism

______________________________
Please specify your use in social media for getting information about tourism sites
Facebook, Instagram, Twitter, Youtube etc.

What are your habits regarding posting and sharing in social media while traveling? *

<table>
<thead>
<tr>
<th></th>
<th>Couple of times a day</th>
<th>Once a day</th>
<th>Several times during the trip</th>
<th>Once in the whole trip</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check-in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photos/videos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How often do you use Twitter? (for posting or following) *

- Never
- Once a month
- Once a week
- More than once a week
- Daily
- Other: [ ]

Submit

Never submit passwords through Google Forms.
9.2 Appendix B
Barcelona Tweets Collection and Analysis

<table>
<thead>
<tr>
<th>La Boqueria</th>
<th>Estatua de Colom</th>
<th>Plaça Catalunya</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
</tr>
<tr>
<td>8 0 1 1 0 2 1</td>
<td>8 0 1 1 0 2 1</td>
<td>8 0 1 1 0 2 1</td>
</tr>
<tr>
<td>9 1 0 3 0 0 0</td>
<td>9 0 0 0 1 4 0</td>
<td>9 3 1 6 4 2 3 3</td>
</tr>
<tr>
<td>10 0 0 0 0 0 0</td>
<td>10 1 1 0 1 0 0</td>
<td>10 4 1 0 9 5 5 2</td>
</tr>
<tr>
<td>11 0 0 0 4 0 0</td>
<td>11 0 0 0 1 0</td>
<td>11 7 3 1 3 5 1</td>
</tr>
<tr>
<td>12 2 1 4 2 2 0</td>
<td>12 1 0 1 2 0 1</td>
<td>12 10 4 1 5 2 8 5</td>
</tr>
<tr>
<td>13 2 1 0 3 0</td>
<td>13 0 0 0 1 0 1</td>
<td>13 7 9 5 5 7 1 2 9</td>
</tr>
<tr>
<td>14 1 1 2 4 2</td>
<td>14 0 0 0 1 0</td>
<td>14 8 3 1 0 4 5 9</td>
</tr>
<tr>
<td>15 2 1 0 0 0</td>
<td>15 0 0 0 1 2 2</td>
<td>15 4 0 8 4 0 2 2 6</td>
</tr>
<tr>
<td>16 1 1 4 1 2</td>
<td>16 0 0 0 0 0</td>
<td>16 7 2 7 5 1 0 3 4</td>
</tr>
<tr>
<td>17 2 1 3 3 0 1</td>
<td>17 0 0 0 1 2</td>
<td>17 5 7 7 6 1 0 2 5</td>
</tr>
<tr>
<td>18 1 1 4 1 0</td>
<td>18 0 0 0 1 3</td>
<td>18 6 7 1 2 9 4 8</td>
</tr>
<tr>
<td>19 3 0 2 3 0 1</td>
<td>19 0 0 3 1 1</td>
<td>19 8 6 7 8 1 0 8</td>
</tr>
<tr>
<td>20 5 2 0 4 0 0</td>
<td>20 0 0 3 1 0</td>
<td>20 7 6 8 8 1 2 1 5 6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sagrada Familia</th>
<th>Plaça de la Barcelona</th>
<th>Parc de la Ciutadella</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
</tr>
<tr>
<td>8 0 1 0 0 1 0 1</td>
<td>8 0 1 0 0 0 0 0 1</td>
<td>8 0 1 0 0 0 0 0 1</td>
</tr>
<tr>
<td>9 0 0 1 3 0 1 1</td>
<td>9 1 0 0 2 0</td>
<td>9 3 1 3 0 0 0 1</td>
</tr>
<tr>
<td>10 1 2 1 0 0 0 1</td>
<td>10 0 0 2 0</td>
<td>10 1 2 1 0 0 1</td>
</tr>
<tr>
<td>11 0 0 5 3 2 2 0</td>
<td>11 0 0 1 3 1 0</td>
<td>11 0 0 2 0</td>
</tr>
<tr>
<td>12 1 1 0 2 4 1</td>
<td>12 0 1 1 1</td>
<td>12 3 1 4 0 1</td>
</tr>
<tr>
<td>13 1 1 4 1 0</td>
<td>13 0 2 5 1 2</td>
<td>13 0 3 3 1 3 1 0</td>
</tr>
<tr>
<td>14 2 0 1 3 2</td>
<td>14 0 1 1 0 1</td>
<td>14 0 1 1 0 1</td>
</tr>
<tr>
<td>15 1 0 1 3 2</td>
<td>15 0 1 1 1</td>
<td>15 0 1 1 0 1</td>
</tr>
<tr>
<td>16 0 0 8 1 0 2 1</td>
<td>16 0 3 3 3 1 3 0</td>
<td>16 0 4 2 3 0</td>
</tr>
<tr>
<td>17 2 0 1 0 0</td>
<td>17 0 3 3 3 1</td>
<td>17 0 4 2 3 0</td>
</tr>
<tr>
<td>18 1 6 3 3</td>
<td>18 0 2 4</td>
<td>18 0 4 2 3 0</td>
</tr>
<tr>
<td>19 3 8 4 1 3</td>
<td>19 0 2 4 2 2</td>
<td>19 0 5 1 0 0 3 2</td>
</tr>
<tr>
<td>20 2 2 3 2 2 0 0</td>
<td>20 0 2 4 2 1</td>
<td>20 0 1 2 1 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monjuïc Nacional</th>
<th>La Font Màgica</th>
<th>Torre Agbar</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON (Closed) TUE WEN THER FRI SAT SUN</td>
<td>MON (no show) TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
</tr>
<tr>
<td>8 2 0 1 2 0 0</td>
<td>8 0 0 1 1 0 0 0</td>
<td>8 0 0 1 1 0 0 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plaça Reial</th>
<th>La Pedrera (Casa Milà)</th>
<th>Parc Güell</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
<td>MON TUE WEN THER FRI SAT SUN</td>
</tr>
<tr>
<td>8 0 0 1 1 0 0</td>
<td>8 0 0 1 0 0</td>
<td>8 0 1 0 1 0 1</td>
</tr>
<tr>
<td>9 0 0 4 1 2 0 1</td>
<td>9 0 1 0 0</td>
<td>9 0 1 0 0 1</td>
</tr>
<tr>
<td>10 1 1 7 6 2 1 0</td>
<td>10 1 0 1 0 0</td>
<td>10 1 0 1 0 0</td>
</tr>
<tr>
<td>11 2 0 5 0 1 2</td>
<td>11 0 0 2 0 1</td>
<td>11 0 0 2 0 1</td>
</tr>
<tr>
<td>12 1 1 6 2 1 1 2</td>
<td>12 0 0 1 1 0</td>
<td>12 0 0 1 1 0 1</td>
</tr>
<tr>
<td>13 0 0 8 2 4 2 3</td>
<td>13 0 1 1 0 1</td>
<td>13 0 1 1 0 1</td>
</tr>
<tr>
<td>14 5 0 5 4 2 6 1</td>
<td>14 0 1 0 2 1</td>
<td>14 0 1 0 2 1</td>
</tr>
<tr>
<td>15 0 2 7 4 8 3 4</td>
<td>15 0 1 0 2 1</td>
<td>15 0 1 0 2 1</td>
</tr>
<tr>
<td>16 2 4 0 5 0 1</td>
<td>16 0 1 0 2 1</td>
<td>16 0 1 0 2 1</td>
</tr>
<tr>
<td>17 2 4 0 5 0 1</td>
<td>17 0 1 0 2 1</td>
<td>17 0 1 0 2 1</td>
</tr>
<tr>
<td>18 2 2 3 1 9 1</td>
<td>18 0 1 0 2</td>
<td>18 0 1 0 2</td>
</tr>
<tr>
<td>19 4 2 3 1 9 1</td>
<td>19 0 1 0 2</td>
<td>19 0 1 0 2</td>
</tr>
<tr>
<td>20 0 0 3 2 3 1 0</td>
<td>20 0 1 0 2</td>
<td>20 0 1 0 2</td>
</tr>
</tbody>
</table>

max 30 min max 2 hours max 1 hour

max 2 hours max 30 min max 20 min

max 30 min max 1 hour max 30 min
<table>
<thead>
<tr>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THUR</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THUR</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THUR</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**max 30 min**

<table>
<thead>
<tr>
<th>MON</th>
<th>TUE</th>
<th>WED</th>
<th>THUR</th>
<th>FRI</th>
<th>SAT</th>
<th>SUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**max 1 hour**
9.3 Appendix C
Tourism Sites Timing in Barcelona

Tourism Sites Timing In Barcelona
This short feedback will help estimate the waiting time for tourism attractions in Barcelona by comparing the application result to your real time feedback.

Age

What was the time difference from the **actual waiting time** to the **application result estimated time**?
For ex. if the application result was "15 min estimated waiting time", but in real time you have waited 30 min - than you should select "15 min longer" since you have waited 15 min longer than expected. Please select the answer that is the closest to your result.

<table>
<thead>
<tr>
<th></th>
<th>30 min longer</th>
<th>15 min longer</th>
<th>Was a match</th>
<th>15 min shorter</th>
<th>30 min shorter</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Boqueria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estatu de Colom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaça Catalunya</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagrada Familia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaça de la Barceloneta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parc de la Ciutadella</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Museu Nacional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Fort Mágica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torre Agbar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaça Espanya</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Pedrera (Casa Milà)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parc Güell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teatre del Liceu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catedral de Barcelona</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please provide us with information about different timing for specific sites that you could elaborate your feedback on.

For example: "In site X, the waiting time was 2 hours more from the application result." "In the application for site X, the result was 1 hour waiting, but there were no waiting lines!" etc.

In the second screen you needed to choose 1 of 3 compatible paths. What was your choice and why?

Please provide feedback about your choice, was is time vs km? was it based on the crowdedness of the sites?

Additional general feedback

Submit

Never submit passwords through Google Forms.
9.4 Appendix D
Trip Planning – Rome

Trip planning - Rome

With the application "Is it Crowded"

*Required

Name *
Your answer

Age *
Your answer

Mobile phone model *
Your answer

Scenario

You need to plan a trip to Rome while you will stay there during the weekend (Friday to Sunday). Try using the application for planning the trip (select sites, hours, days). You will stay in a Hotel where it’s address is "Via della Vetrina, 500186 Roma".
Please describe your planning process *
How did you manage to plan? did you reviewed the application before/while planning? Have you changed your route often?
Your answer

Please describe your planning results *
Which sites will you visit by each day? Day 1, Day 2, Day 3
Your answer

Please rate the following phrases from disagree to agree *

<table>
<thead>
<tr>
<th>phrases</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I manage to navigate between the screens easily</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I could identify the buttons and links in the screens</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I liked the page layout (elements positioning)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I liked the application colors and design</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I found it easy to plan the trip using the app</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The application revealed me to new information</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The new information changed my plans for the trip</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The application met my expectations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
9.5 Appendix E
Live Experience – Tel Aviv

Live Experience - Tel Aviv

With the application "Is it Crowded"

*Required

Name *

Your answer

After Trip

After you have used the application in the streets of Tel Aviv, you must have some new insights about the application.

Please rate the following phrases from disagree to agree *

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I found it easy to navigate using the app</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I think avoiding crowdedness can improve my trip experience</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
I have changed my sites selections and re-navigate due to the application results *

- Yes
- No

I have tried different sites combinations for a better result *

- Yes
- No

If you have answered any of the phrases in “Yes” please describe the situation.
Your answer

What did you enjoyed most? *

- Planning the trip using the application
- Using the application in real time

Please specify
Your answer

NEXT
Application Screens

Please rate and review the screens in application.

Home Page

IS IT CROWDED?

Barcelona
Rome
Tel Aviv

Feedback

Please rate the screen *

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needs Improvements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Love it!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What would you add/change to this screen by features or look? What was missing? *

Your answer
Selecting Tourism Sites

Please rate the screen *

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Love it!</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needs Improvements</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

What would you add/change to this screen by features or look? What was missing? *

Your answer
Preferences

IS IT CROWDED?

YOU CAN ADD SOME...
Add a starting point address
Use my current location as a starting point
Future hour:
**Default time = current hour**
Different day:
**Default Week day = current day**

Calc Recommended Path

Please rate the screen *

<table>
<thead>
<tr>
<th>Needs Improvements</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Love it!</th>
</tr>
</thead>
</table>

What would you add/change to this screen by features or look? What was missing? *

Your answer
Selecting the Optimal Path

**IS IT CROWDED?**

**YOUR OPTIMAL PATHS TO CHOOSE FROM:**
- Not crowded - Less than 10 minutes
- A bit crowded - Between 10 to 30 minutes
- Very crowded - More than 30 minutes

**PATH 1**
- Total KM: 1.14 | Total Time: 05:34
- Place 1: Hatochara
- Place 2: Neve Tzedeck
- Place 3: Independence Hall

**PATH 2**
- Total KM: 1.95 | Total Time: 05:41
- Place 1: Hatochara
- Place 2: Independence Hall
- Place 3: Neve Tzedeck

**PATH 3**
- Total KM: 1.95 | Total Time: 05:56
- Place 1: Neve Tzedeck
- Place 2: Independence Hall
- Place 3: Hatochara

Please rate the screen *

Needs Improvements

1  2  3  4  5  Love It!

What would you add/change to this screen by features or look? What was missing? *

Your answer
Path Result & Map

**PATH 1**
TOTAL KM: 3.6 | TOTAL TIME: 03:08

- **Place 1:** Colosseum
  - Arrival: 11:31
  - Time: 00:00
  - Estimated Waiting: None
  - Average Time: 09:30:00

- **Place 2:** Piazza Navona

- **Place 3:** Spanish steps

Navigate Here From Current Location

Navigate All Route by Google Maps

Please rate the screen *

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Needs Improvements

Love it!

What would you add/change to this screen by features or look?
What was missing? *

Your answer

[BACK] [SUBMIT]

Never submit passwords through Google Forms.