Faculty of Environmental Sciences Institute for Cartography

Master Thesis

Pedestrian Navigation with consideration of affective responses towards the environment extracted from Location-Based Social Media Data

submitted by Madalina Gugulica
born on 23.08.1989 in Iasi, Romania

submitted for the academic degree of
Master of Science (M.Sc.)

Date of Submission 15.02.2019

Supervisors Dr. Eva Hauthal
Institut für Kartographie, Technische Universität Dresden

Univ.Prof. Mag.rer.nat. Dr.rer.nat. Georg Gartner
Institut für Geoinformation und Kartographie, Technische Universität Wien

Reviewer Prof. Dipl.-Phys. Dr.-Ing. habil. Dirk Burghardt
Institut für Kartographie, Technische Universität Dresden

ORIGINAL TASK DESCRIPTION

Objective

Pedestrians' daily behaviour and decision-making in space is influenced, apart from time and distance factors, by an individual's subjective interpretation of the environment. Most of the current navigation services, however, employ time-optimized or distance-optimized algorithms and fail to fully meet the needs of the users. Mobile pedestrian navigation systems could be enhanced if people's affective responses to space would be taken into consideration and would be integrated in route planning algorithms.

The aim of this master thesis is to expand the "AffectRoute" study conducted by the Research Group Cartography within the Technical University of Vienna by exploring the use of location based social media data and natural language processing techniques such as sentiment analysis.

The basic idea is to develop a methodology for extracting, modelling and integrating people's affective responses to environments into a route planning algorithm and find out if location based social media data could provide enough and relevant information for such a methodology to be feasible and viable. The extractions and analysis of the people's perceptions on the environment will be based on methods developed at the Institute of Cartography from Technical University of Dresden. Their approach is to extract the information from written language contained within the metadata of georeferenced photos posted on social media networks such as Flickr, Panoramio and Instagram.

Description

Pedestrian navigation has become one important research topic in the community of cartographers, geographical information systems experts and location-based services experts. A great progress has been achieved over the last couple of decades by enhancing routing algorithms, expanding the understanding of spatial behaviour or developing technologies that aid people navigate indoors. Nevertheless, most of the current pedestrian navigation services employ either time-optimized or distance-optimized algorithms and fail to meet the needs of the users in terms of safety and satisfaction, which are mostly related to the environmental factors and personal interpretation of the space. Thus, researchers suggest that pedestrian navigation theories and technologies should apply a people-centric approach that incorporates the user's subjective relation to space.

A study conducted by the Research Group Cartography within the Technical University of Vienna proposed a crowdsourcing approach to collect and model people's affective responses to space and developed a routing algorithm (named AffectRoute) to aggregate and integrate the collected data for route planning. Even with the growing popularity of crowdsourcing in the era of Web 2.0, observing the low number of downloads the Emo-Map mobile application (the application was developed to facilitate this crowdsourcing approach) registered, is slightly over 100, it could be argued that additional sources of data are required for a successful implementation of the routing algorithm.

The emergence of Volunteered Geographic Information (VGI) as an alternative and powerful spatial data source on the Web offers us the possibility to extract information concerning people's affective responses to environments. Thus, the aim of this thesis is to expand the study mentioned above by exploring location based social media data, developing a methodology for extracting, modelling and integrating people's affective responses into a route planning algorithm and decide the feasibility and viability of such an approach.

The main idea is to retrieve georeferenced data for a specific area of interest from Location Based Social Networks (such as Instagram, Flickr, Panoramio) and extract the affect responses to the environment by applying natural language processing techniques. The affective responses will be furthered analysed and modelled in order to create a navigation network characterized by the emotional component of the human relationship with space and a routing algorithm will be developed. In the end the feasibility and viability of the approach will be tested and determined.

STATEMENT OF AUTHORSHIP

Herewith I declare that I am the sole author of the thesis named

„Pedestrian Navigation with consideration of affective responses towards the environment extracted from Location-Based Social Media Data“

which has been submitted to the study commission of geosciences today.

I have fully referenced the ideas and work of others, whether published or unpublished.

Literal or analogous citations are clearly marked as such.

Dresden, 15.02.2019 Signature

ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the support and encouragement of several special people. Hence, I would like to take this opportunity to express my gratitude to those who have been by my side and assisted me in a myriad of ways.

I would first like to express my heartfelt thanks to my supervisor Dr. Eva Hauthal and to Prof. Dr. -Ing Dirk Burghardt for the useful comments, remarks and engagement throughout the conceptual phase of this master thesis. Furthermore, I would like to thank Dr. - Ing Alexander Dunkel for all his help and for providing me the data necessary for the analysis performed within my work.

To my mother and my father, words cannot express my gratitude and appreciation for your support, encouragement and the untold number of sacrifices you made for me. I will always carry you in my heart!

Finally, I want to thank my dearest friends Alina, Katy and Vlad for being my pillars of strength in times of grief and sorrow and for encouraging me to carry on and not give up on my professional endeavours.

ABSTRACT

Research on environmental psychology has shown that humans perceive environments not only according to their physical features, but also in relation to their affective qualities. This affective interpretation of their surroundings often influences people's spatial behaviour and decision-making process during navigation tasks. Thus, studying people's environmental affective perceptions would lead to a better comprehension of people's spatial experiences and behaviour, as well as to the enhancement of certain category of location-based services, namely pedestrian navigation services.

This research regards Location-Based Social Media data, namely the metadata of geotagged images shared on social media photo platforms, as a potent and readily available resource for studying and extracting people's affective responses to the environment and proposes a methodology to harness this data for pedestrian navigation purposes. More specifically, within this thesis it was investigated how the affective responses to environments can be described in a structured way in relation to Location-Based Social Media data, the main characteristics and issues of this kind of data were identified and a methodology to extract them was developed. Furthermore, an innovative approach for the aggregation and modelling of the extracted data for the enhancement of pedestrian route planning was conceived and a subjective data layer was delivered.

The findings of the study show that, on one hand, that textual-based social media data can be regarded as significant data source to extract affective information for further integration in route planning algorithms, in oder to provide human-centred pedestrian navigation systems. On the other hand, due to its complex and complicated nature, extracting the affective response from this kind of data is a difficult task that raises an important number of issues that need further revision and new avenues for research should be explored.

Keywords: pedestrian navigation, location-based social media data, affective responses to environments

TABLE OF CONTENTS

ORIGINAL TASK DESCRIPTION II

STATEMENT OF AUTHORSHIP IV

ACKNOWLEDGEMENTS V

ABSTRACT VI

TABLE OF CONTENTS 1

LIST OF FIGURES 3

TABLES 4

FORMULAS 5

ABBREVIATIONS 6

1 INTRODUCTION 7
- 1.1 Motivation and problem statement 7
- 1.2 Research Objectives and Research Questions 9
- 1.3 Thesis Outline 9

2 THEORETICAL BACKGROUND 10
- 2.1 Affective Responses to Environments 10
  - 2.1.1 Construction and Quantification of Emotions 10
  - 2.1.2 Approaches to Measure Emotion 12
    - 2.1.2.1 Traditional Instruments of Measuring Emotions 12
    - 2.1.2.2 Emotional Expression through Language 16
    - 2.1.2.3 Sentiment Analysis 16
  - 2.1.3 Affective Qualities of the Environments 21
- 2.2 Location-Based Social Media Data 22
  - 2.2.1 Image-based VGI: Flickr and Instagram 23
  - 2.2.2 Benefits and Challenges of Using Location -Based Social Media Data 26
- 2.3 Pedestrian Navigation 27
  - 2.3.1 Human Navigation 27
  - 2.3.2 Navigation Systems 28
  - 2.3.3 Human-Centred Pedestrian Navigation Systems 29

3 PROPOSED METHODOLOGY 32
3.1 Workflow 32
3.2 Data Preparation 33
3.3 Approach to Extract the Affective Responses to the Environment 34
3.4 Approach to Aggregate and Model the Extracted Affective Responses 37
3.4.1 Spatial Extent of the Affective Responses and Aggregation Method 37
3.4.2 Quantification and Classification of the Affective Ratings 38
4 IMPLEMENTATION OF THE METHODOLOGY 42
4.1 Data Basis 42
4.2 Data Preparation and Extraction of the Affective Responses 43
4.2.1 Developed Algorithm 43
4.2.2 Software Resources 46
4.3 Aggregation of the Affective Responses 47
5 RESULTS 49
5.1 Results of the Implementation 49
5.2 Evaluation of the Developed Methodology 51
5.2.1 Assessment of the Data Preparation 51
5.2.2 Assessment of the Sentiment Analysis and of the Affective Classes Proposed 53
6 DISCUSSION AND FUTURE WORK 56
6.1 Weaknesses and Limitations of the Proposed Methodology 56
6.2 Future Work 58
7 CONCLUSIONS 60
References 62
APPENDICES VII

LIST OF FIGURES

Figure 1 *Core Affect* taken from Russell (2003) 11
Figure 2 Excerpt of Stockport Emotion Map (Nold, 2007) 14
Figure 3 “The First Published Map of World Happiness” 14
Figure 4 “Level of comfort according to three urban scenes” (Klettner et al., 2013b) 15
Figure 5 Distribution of ANEW and Warriner et al. words in valence-arousal-space 19
Figure 6 Distribution of BAWL-R and Warriner et al. words in valence-arousal-space 20
Figure 7 “3D visualisation of valence and arousal in Dresden” (Hauthal, 2015) 21
Figure 8 Methodology Overview 32
Figure 9 POS tagging explained 36
Figure 10 *The 12 – Point Affect Circumplex Model* (Yik et al., 2011) 39
Figure 11 Affective Classes (the horizontal axis represents the valence dimension, whereas the vertical axis represents the arousal dimension) 40
Figure 12 Spatial Distribution of the Geotagged Photographs 42
Figure 13 Implemented Algorithm for the Extraction of the Affective Responses 44
Figure 14 Excerpt from the Resulting Affective Responses Feature Class 46
Figure 15 Implemented Algorithm for the Aggregation of the Extracted Affective Responses 47
Figure 16 Spatial Distribution and Classification of the Extracted Affective Responses 49
Figure 17 Classification of the Extracted Affective Ratings (according to the 5 defined classes) 50
Figure 18 Results of the Classification of the Routing Graph’s Edges after the Aggregation of the Affective Ratings 51
Figure 19 Hashtag Segmentation Evaluation 52

TABLES

Table 1 Classification of Tags and their Functions (Sen et al., 2006 and Golder & Huberman, 2006) 24
Table 2 Meaning of Part-of-Speech Codes 36
Table 3 Difference between the performances of English and German affective words extraction 53
Table 4 List of Incorrectly Processed Affective Terms 55

FORMULAS

Equation 1 Weighted Average Valence 38
Equation 1 Weighted Average Arousal... 38
Equation 3 Average Valence of User's Textual Entry 46
Equation 4 Average Arousal of User's Textual Entry... 46

ABBREVIATIONS

ANEW	Affective Norms for English Words
APA	American Psychological Association
API	Application Programming Interface
BAWL-R	Berlin Affective Word List Reloaded
EXIF	Exchangeable Image File Format
FEAT	Facial Expression Analysis Tool
GPS	Global Positioning System
GUIDs	Global Unique Identifiers
LBS	Location-Based Services
LBSM	Location-Based Social Media
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
OSM	OpenStreetMap
POS	Part of Speech
PNS	Pedestrian Navigation System
PNSs	Pedestrian Navigation Systems
VNS	Vehicle Navigation System
VNSs	Vehicle Navigation Systems
VGI	Volunteered Geographic Information
URL	Uniform Resource Locator
Warriner et al.	English Affective Lexicon create by Warriner et al. (2013)

1 INTRODUCTION

Research on environmental psychology shows that humans perceive environments not only according to their physical features, but also affectively (Russell, 2003). This affective interpretation of their surroundings often influences people's spatial behaviour and decision-making process (Borst et al. 2009). Thus, studying people's environmental affective perceptions would lead to a better comprehension of people's spatial experiences and behaviour, as well as to the enhancement of certain location-based services, such as navigation systems, travel application and urban planning.

1.1 Motivation and problem statement

Humans must frequently navigate in unfamiliar environments and in order to complete their task they often need assistance. In the past, people had to refer to maps or simply to asking other people for instructions, but with the increasing use of mobile devices and the emergence of Location-Based Services (LBS), mobile navigation systems have been developed. Mobile navigation systems are composed of three models: positioning, route-planning and route-communication (Huang & Gartner 2009). The role of the route-planning module is to compute the optimal pathway between the origin and the desired destination. These mobile navigation systems were initially developed for car drivers and were later expanded for pedestrian use, yet employing the same time- or distance-optimized routing algorithms used in automotive navigation systems.

Research on spatial cognition and behaviour of pedestrians has suggested the necessity of adjusting pedestrian navigation systems to the users' needs (Gartner et al. 2011). Thus, new approaches of computing routes such as: routes with a minimal number of turns and routes with a minimal angle (Winter 2002), routes with least instruction complexity (Duckham & Kulik 2003), reliable routes minimizing the number of complex intersections with turn ambiguities (Haque et al. 2007), routes adapted to a pedestrian typology based on motion behaviour (Millonig & Gartner 2009) and "affect routes" (Huang et al., 2014) have been proposed.

Studying people's environmental affective perceptions has become the subject of more and more studies, as scientists started to acknowledge the importance and potential of harnessing such knowledge. The "AffectRoute" is one initial study, conducted by the Research Group Cartography within the Technical

University of Vienna, that regards people's affective responses to environments for navigation purposes. The study proposed a crowdsourcing approach to collect people's affective responses to space and designed a routing algorithm which aggregates and integrates the collected data for automatic route planning. Even with the growing popularity of crowdsourcing in the era of Web 2.0, motivating people to participate can be very challenging and as reported by the researchers in their publication (Huang et al. 2014) additional sources of data could be regarded for a successful implementation of the routing algorithm at a larger scale.

With the increasing use of social media networks (e.g. Twitter[1] Instagram[2], Flickr[3] etc.) and dissemination of mobile devices equipped with positioning sensors, large volumes of geotagged social media data are continuously produced. This user-generated content is quite diverse and among others, often contains subjective information such as the user's emotional response to the surrounding stimuli at a certain time in a certain place. This data could be considered a readily available resource for studying and extracting people's affective responses to the environment.

The aim of this master thesis was to develop a methodology which builds upon the ideas and methods proposed by the "AffectRoute" study and makes use of location-based social media data as an additional source for the extraction of the affective responses towards the environment. Furthermore, another goal of this research was to find a suitable way of aggregating the extracted data and deliver an affective-layer that could be further incorporated into a routing algorithm.

On the whole, developing a methodology to harness already existing and free data for navigation purposes would lead to the facilitation of data collection and generally, to the enhancement of mobile pedestrian navigation services.

To extract the affective responses, the written language in the metadata of georeferenced Flickr and Instagram photographs is used. For testing purposes, the author collected data for a familiar and rather decreased pilot area, namely the city of Dresden (Germany).

[1] http://www.twitter.com
[2] http://www.instagram.com
[3] http://www.flickr.com

1.2 Research Objectives and Research Questions

The main goal of this thesis was to develop a methodology that, from the written language in the metadata of georeferenced Flickr and Instagram photos:

(1) Extracts people's affective responses to the environment applying natural language processing (NLP) and sentiment analysis techniques.

(2) Aggregates and models the extracted data and delivers an affective data layer suitable for a further integration into a routing algorithm.

The research questions that need to be addressed in order to meet the main objective of this research are as follows:

(1) How can the affective responses towards the environment be described in a structured way and how can they be extracted from location-based social media data?

(2) What are the main characteristics of text-based social media data and how should these be considered when extracting affective responses?

(3) The extracted affective responses will be initially stored as discrete data. How should be this data aggregated into an affective layer suitable for a further integration with the navigation graph used by the routing algorithm?

1.3 Thesis Outline

The thesis is structured into five chapters. Following the introductory chapter, the second chapter covers a comprehensive review of the theoretical background, treating the subjects of affect and affective responses, location-based social media data and pedestrian navigation. It summarises the theories and ideas on which this research is based, identifies the key issues and challenges and presents some of the related projects and studies. The third chapter exposes the methodology developed and presents each phase of it in detail explaining the approaches, algorithms and technology that has been used. Chapter four reveals the results of the implemented workflows and algorithms, visualises them and discusses the main difficulties encountered. Finally, in the fifth chapter conclusions are drawn and recommendations for future work and improvements are proposed.

2 THEORETICAL BACKGROUND

Based on the assumptions that people use social media platforms to share information, opinions and the way they feel about situations, people, objects and places, the possibility of extracting affective responses to environments from location-based social data is considered. In order to fulfil the objectives of this study, it is essential to understand what affective responses are, the nature and characteristics of this particular type of data and how the extracted information could benefit pedestrian navigation. Therefore, the aim of this chapter is to synthesize relevant research in the main fields underlying this thesis. The first part of this chapter introduces the concept of affective response and is followed by a brief description of location-based social media data as a form of Volunteered Geographic Information. Pedestrian navigation is then treated in the final part of the chapter.

2.1 Affective Responses to Environments

The affective response represents the general internal neuropsychological state of an individual within a given situation. In other words, it is an individual's emotional reaction of a single stimulus or a set of various stimuli. Thus, as asserted by studies in environmental psychology (e.g. Mehrabian & Russell,1974; Bondi et al., 2007), environments are perceived not only according to their physical features, but also in terms of a person's affective response towards them. Before exploring further the relationship between environments and the emotions they evoke in individuals or studying the methods available to measure and extract this affective information, it is important to discuss how emotions are constructed and structured.

2.1.1 Construction and Quantification of Emotions

The concept of emotion seems to be widely understood, nevertheless, it is surprisingly difficult to come up with a solid definition. In the last century psychologists have offered a variety of definitions for this complex construct, each focussing on various manifestations or components of emotion. Of particular interest for this research is the vantage point from which James Russell describes emotion and its components. According to Russell (2003), emotion is described as the mental categorization of a neurophysiological state. Thus, he proposes the term "core affect" to define the omnipresent assessment of one's current internal neurophysiological state as well as the one of the primitives responsible for the construction of any affective states such as emotions and moods. Furthermore, in reference to his approach of structuring emotions, which will be discussed in the following, the core affect could also be described as the component that gives the emotional state its hedonic tone [4] and felt energy.

In modern psychology there are two broadly accepted approaches used to classify the vast range of affective states. The first approach considers that emotions are discrete and fundamentally different constructs while the second approach asserts that emotional states can be classified according to a dimensional modelling of the core affect. Representative for the first approach is the work of Ekman and Friesen (1971), who after studying emotions in relation to facial expressions proposed the theory that there is a set of universal emotions which constitutes the primary components of the affect and from which all the other emotional states can be derived. They initially proposed a list of six basic and universal emotions that included anger, disgust, fear, happiness, sadness and surprise.

The dimensional approach simplifies the quantification of emotional states to the measurement of two or three dimensions of the core effect. Russell's (1980) approach of quantifying emotional states, which is the one significant for this research, is based on the use of a two-dimensional scale of valence (how negative-unpleasant or positive-pleasant is the emotional state) and arousal (how enervated – deactivated or energized –activated is the emotional state). According to these two perceived dimensions of the core affect, the emotions can be then located in a two-dimensional affective space by plotting the values of valence and arousal for the core affect (see Figure 1).

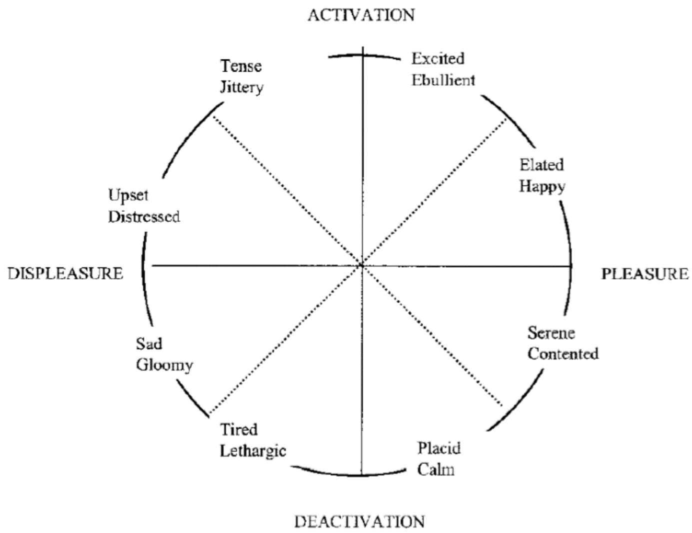


Figure 1 *Core Affect* taken from Russell (2003)

[4] property of a sensory or other experience relating to its pleasantness or unpleasantness (APA Dictionary, 2018)

To exemplify Russell's approach a basic emotion such as happiness can be regarded as a core affect with a high positive value of the hedonic tone and a high value of the arousal, whereas the core affect for the emotional state of anger would have negative valence and high arousal values.

2.1.2 Approaches to Measure Emotion

Due to their complexity, emotions cannot be measured directly, but solely by quantifying their manifestations. The multiple views on defining emotions lead to a debate on which manifestation is sufficient or necessary to quantify emotions. A current general favoured solution to this debate is to treat emotions as a multifaceted phenomenon manifested through the following components (Battachi et al., 1996 as cited in Hauthal, 2015):

- physiological reactions (e.g. changes in the heart rate, faster breathing, sweating etc.)
- motoric reactions (e.g. gestures, facial expressions)
- linguistic reactions (e.g. lexical and syntactical constructions, words choices)
- subjective affective reactions (e.g. feeling comfortable, happy etc.)

2.1.2.1 Traditional Instruments of Measuring Emotions

Traditionally, instruments, that measure the emotional manifestations mentioned above, range from simple self-reports to sophisticated high technology – based physiological methods that track brain waves or eye movements. Often, the instruments are classified as non-verbal (objective) instruments and verbal (subjective) instruments.

The most often used verbal instrument is the self-report (e.g. questionnaire and interviews) that require respondents to report their emotions through the use of a set of rating scales or verbal protocols. Barrett et al. (2007) pointed out that self-reports are the most direct way to gather information about an individual's subjective feelings. Nevertheless, there is a series of disadvantages associated with these methods, namely the limited access to verbal statements on emotions and the lack of details.

The non-verbal instruments could be further divided into physiological measurements and behavioural observations. Physiological recordings use various sensors (e.g. electrocardiogram, electromyogram, galvanic skin response) to provide an objective measurement of the affective responses experienced by the study participant. The advantage of this method is that the person does not even have to recognize the physiological changes and this way cannot manipulate them. Behavioural observations consist of the interpretation of expressive motoric reactions (the facial, vocal, and postural expression) that accompany the emotion with the aid of interpretation tools such as Facial Expression Analysis Tool (FEAT; Kaiser &Wehrle, 1992). This method has similar advantages and disadvantages as the physiological measurements one. The observed reactions can hardly be manipulated; however, their occurrence is not strictly determined by a person's affective reactions caused by the environment. Consequently, the data collected from physiological recordings and behavioural observations without a subjective interpretation may be insufficient to precisely identify the individual's emotional state.

These traditional meanings of measuring emotions have been exploited by the pioneers in the field of emotional cartography and in the following some of their studies are presented.

The first study to collect emotions related to space was the project of Chris Nold, *Bio Mapping*. Nold produced a series of emotion maps of communities using his 'Bio Mapping' device, a Global Positioning System (GPS) enabled, wearable device that measured galvanic skin response (Nold, 2009). For instance, one noteworthy map he produced is the Stockport Emotion Map. The map is based on the results of the biometric measurements in terms of emotional arousal experienced by the participants of his experiment walking around the town of Stockport (United Kingdom) and includes also sketches and notes of the participants describing the town (see Figure 2)

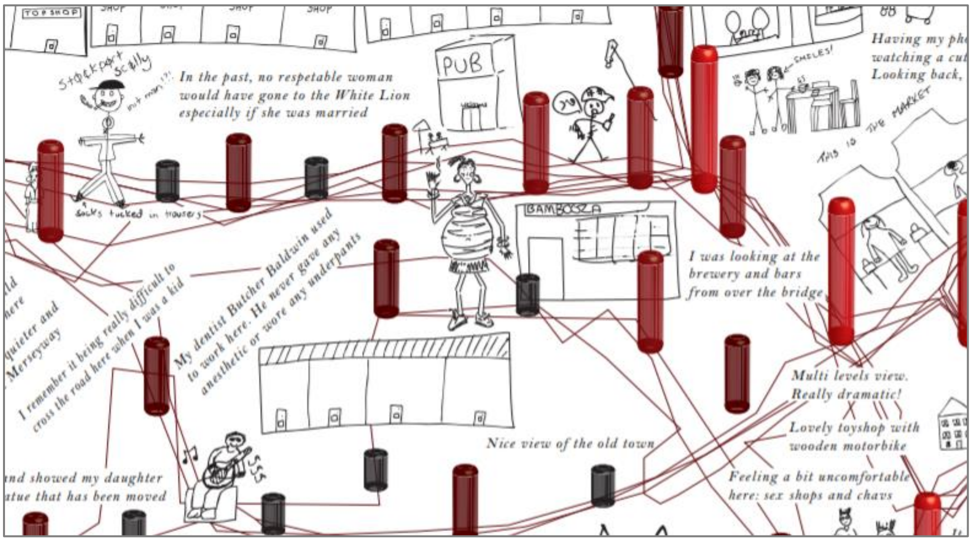


Figure 2 Excerpt of Stockport Emotion Map (Nold, 2007)

Another project that collects affective responses through traditional verbal instruments is the project of Adrian White (2007), who assumed that the estimation of the subjective well-being of a country can represent a way to measure the underlying level of happiness experienced by a population. After collecting data interviewing more than 80,000 people worldwide, he created the “First Published Map of World Happiness”, a choropleth map based on the measurements of the subjective well-being (Figure 3)

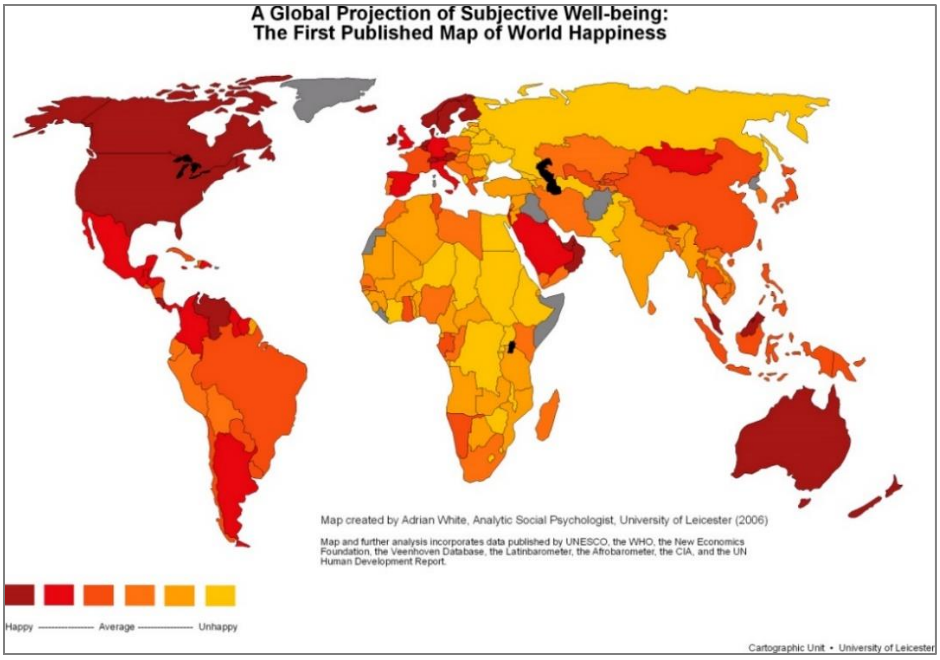


Figure 3 “The First Published Map of World Happiness”[5]

[5] Available at: http://www.zonu.com/images/0X0/2010-09-07-12097/World-happiness-2006.jpg (accessed on 28.12.2018)

Some other projects have adapted the traditional self-reports to the technological trend and developed mobile applications to crowdsource data regarding the emotions people experience in space.

Mody et al. (2009) asserted that emotions are a critical aspect of how people experience places and proposed that collecting and sharing such information would offer social context to location-based services. They introduced WiMo, a location-based social networking tool that enables users to share and geotag their emotions evoked by the encountered environments. A mobile application was developed to enable the users to input information about the level of comfort they experienced, their general feeling about a certain location (if they liked it or rather disliked it) and add personalised short comments to share them with other users.

Similarly, the EmoMap Project derived from the assumption that each person perceives the urban environments in a subjective way. In oder to collect the affective data, they developed a three-level emotion-space model that combines discrete emotions with the dimension of valence and the affective qualities of the environment. Vienna was chosen as a trial area and the participants were asked to submit their responses regarding how pleasant/unpleasant they perceived the surroundings, their current mood (stressed/relaxed/bored/excited) and environmental qualities in terms of traffic, noise, smell, attractiveness or level of safety. The purpose of the study was two-fold: extraction and visualisation of data regarding the affective relations people establish with the environment and enhancement of pedestrian navigation systems. The study concluded that people perceive green spaces as the most pleasant and the areas with heavy traffic as unpleasant and stressful, while attractiveness, level of comfort and safety are factors that determine route choices (Klettner et al., 2013a).

Figure 4 “Level of comfort according to three urban scenes” (Klettner et al., 2013b)

2.1.2.2 Emotional Expression through Language

In psychology, emotional expressions are both verbal and non-verbal behaviours that describe internal affective states. They are way of communicating affective states and occur with or without self-awareness. As mentioned before, non-verbal behaviour such as facial expressions, laughing, crying etc. indicates the general emotion the individual experiences; however, it is sometimes not sufficient to precisely identify the nuances of more complex emotional states. Language, on the other hand, is the tool that gives humans the opportunity to express the subtleties and complexity of their emotions.

In the field that studies the relationship between language and emotion, three directions of research can be distinguished (Lindquist et al., 2016). Firstly, there is the psychological constructionist approach, which assumes that language influences the construction of emotions in the first place when affect is made meaningful as an instance of an emotion category. Secondly, in opposition to the first approach, the emotion regulation model assumes that words can feedback to modulate emotions, helping to regulate emotions after their construction. Finally, there is the approach that considers that language is just a tool to exteriorise the affective states after they occur. It is based on the assertion that most words have an affective connotation that can be derived through empirical methods and stored as emotional vocabularies of different languages.

In connection with the content that people share on social media platforms, emotional expressions are typically not explicit but rather coded with the affective connotations of the words the users choose to describe the content of the data they are sharing. Thus, to analyse the emotional expressions on social media, the last approach is typically employed.

2.1.2.3 Sentiment Analysis

In recent years, the rapid growth of social media - product reviews, blogs, microblogs, and social networks, led to an increasing interest in extracting affective information from this generated vast data sources. Hence, a new application of NLP, sentiment analysis, emerged.

Sentiment analysis (also known as opinion mining) refers to multi-disciplinary field that uses NLP, text analysis and computational linguistics to systematically identify, extract, quantify and study people's affective states, opinions or attitudes towards various entities and their characteristics expressed in written language. Generally speaking, it is the task that aims to determine the attitude of a writer with respect to some topic or the overall tonality of a document. Predominantly, sentiment analysis is a semantic analysis task that focuses on determining the overall positive, negative or neutral affective connotation of a document, phrase or sentence. Due to the syntactical and lexical complexity of human language, sentiment analysis is a very challenging task and over the recent years different approaches that aim to identify sentiments as precisely as possible have been developed.

Concerning the manner in which the sentiment analysis is performed, the methods can be grouped into three main approaches: machine learning approaches (supervised sentiment analysis), lexical approaches (unsupervised sentiment analysis) and hybrid approaches (that combine both supervised and unsupervised methods).

Supervised sentiment analysis methods are the ones whereby applications, called classifiers, are developed and trained to make classification decisions based on manually annotated domain-specific corpus data. Most of supervised approaches, employ machine learning algorithms such as Support Vector Machines (Mullen and Collier, 2004; Prabowo et al., 2009; Wilson et al., 2005) or Naïve Bayes (Pang and Lee, 2004; Wiebe et al., 2005; Tan et al., 2009). Within these algorithms, texts are represented as vectors of features and depending on the words-frequency models (bag-of-words and lexeme-based features are the more commonly used (Pang and Lee, 2008)), the classifiers can reach higher level of performance. The main limitation of supervised approaches is the fact that these applications require vast amounts of domain-specific training data in order to learn to effectively classify texts and the creation of the annotated corpus is tedious and costly.

The unsupervised sentiment analysis approach is based on the creation of dictionaries and lexicons of affectively annotated words, by assigning to each item of the words collection a score that indicates how strong is the relation between that word and each polarity is assigned. The classification of a text using these scores is then performed by calculating the overall polarity score. The most common scoring methods are generally adding the positive values and subtracting the negative values of the single words in a text. (Liu, 2012; SentiWordnet2).

The limitation of the keywords spotting approach lies in the method's incapability to recognise negation, adjectives' comparative and superlative forms or sense figurative speech. To exemplify one of the cases, the following sentence 'This place is not attractive at al' with the keywords spotting technique will be misclassified as overall positive due to the positive affective connotation of the adjective 'beautiful', when in fact the sentence has a highly negative polarity caused by the negation 'not' and the adverb 'at all', which has the role to amplify the negative sentiment.

Possible affective lexicons are ANEW (Affective Norms for English Words; Bradley & Lang, 2010), Warriner et al. (2013), BAWL-R (Berlin Affective Word List Reloaded; Võ et al., 2009) or SentiWordNet 3.0 (Baccianella et al., 2010). ANEW is a collection of 2,476 English words described with values ranging from 1 to 9 for the dimensions of valence, arousal and dominance. The lexicon proposed by Warriner et al. (2013) extended the ANEW lexicon to 13,915 English words including gender, age and educational differences in emotion norms and they also "included stimuli from nearly all of the category norms (e.g., types of diseases, occupations, and taboo words) collected by Van Overschelde, Rawson, and Dunlosky (2004)" (Warriner et al., 2013, pp.1).

The BAWL-R (The Berlin Affective Word List Reloaded) lexicon contains 2,901 German with associated values for valence [-3.0, 3.0], arousal [1.0, 5.0] and imageability [1.0,7.0]. SentiWordNet is based on the WordNet lexical database for English language created in the Cognitive Science Laboratory of Princeton University under the direction of psychology professor George Armitage Miller starting in 1985 and directed in recent years by Christiane Fellbaum (Wikipedia, 2018). To each synonym set of WordNet three sentiment scores: positivity, negativity and objectivity. Each of the three scores ranges in the interval [0.0, 1.0], and their sum is 1.0 for each synonym set (Baccianella et al., 2010).

The hybrid approaches, also called concept-based approaches (Cambria et al., 2013), are methods that employ deep learning and machine learning algorithms and are based on Web ontologies or semantic networks to first perform semantic text analysis to identify implicit meaning/features associated with natural language concepts and subsequently classifies texts affectively making use of affective lexicons and counting cooccurrences. These approaches are able to analyse multi-words constructions and sense figurative speech; however, as Cambria et al. (2013, pp. 19) emphasize, they heavily rely on the depth of the knowledge they are based on and "without comprehensive resource that encompasses human knowledge", the sentiment analysis system will perform poorly in identifying the semantics of natural language text.

Despite all the challenges and performance limitations, the lexicon-based methods' accessibility and economy make it to be extensively used for sentiment analysis tasks. Recently, the weaknesses of the method have been explored through studies and projects that enhanced this approach by developing methods of reliably classifying grammatical special cases such as negations and comparative and superlative forms of adjectives (Hauthal, 2015) or even sensing and classifying cases of sarcasm (Agrawal, 2018) which led to improved lexicon-based sentiment analysis methods. Due to accessibility reasons this study employs the basic approach of lexicon-based sentiment analysis; however, the enhanced approaches are considered for the future work and the improvement of the research.

Visualising the distribution of the ANEW and Warriner et al. words in the valence-arousal space (Figure 5), reveals the improvement brought by Warriner et al. to the ANEW lexicon and that is the reason for which the current research uses the Warriner et al. affective lexicon for the sentiment analysis of the English text. Moreover, for the affective classification of the German text, the BAWL- R affective lexicon is used.

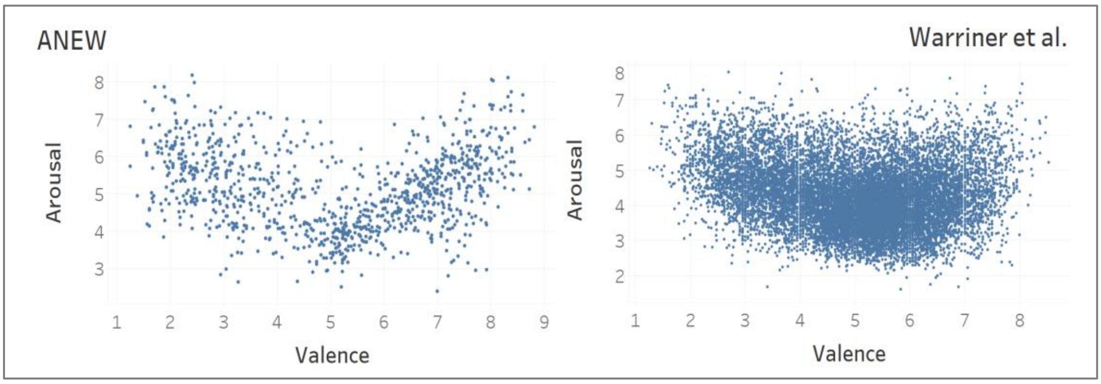


Figure 5 Distribution of ANEW and Warriner et al. words in valence-arousal-space

On another note, it is essential to mention the differences between the two different lexicons used for the sentiment analysis of German and English words, namely BAWL-R and Warriner et al. In the first place, as it can be observed in the distribution of the words of the two lexicons in the valence-arousal space, the boomerang-shaped distribution (see Figure 3) of the BAWL-R data is caused by the fact that the lexicon does not include positive words with high arousal such as taboo words (Võ et al., 2009). Secondly, the remarkable difference between the number of words each lexicon contains (13,915 English words in comparison to only 2901 German words as depicted in Figure 6) might lead to a better performance of the sentiment analysis carried out on English text.

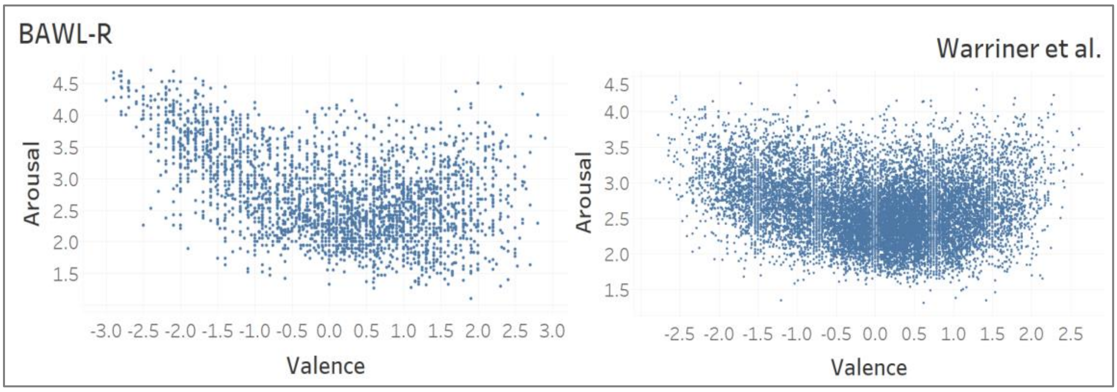


Figure 6 Distribution of BAWL-R and Warriner et al. words in valence-arousal-space

In the last decade, understanding the potential of this tool, sentiment analysis has been adopted by researchers in geosciences to study the relations between human emotions and space. Some noteworthy projects and studies are presented in the following.

Mislove et al. (2010) performed lexicon-based sentiment analysis (based on the ANEW affective lexicon) on Twitter georeferenced data and generated a series of hourly density-preserving cartograms representing the current mood of people in the United States.

Another project, 'Beautiful picture of an ugly place' applies sentiment analysis to the comments of Flickr photos to extract the affective responses of the users regarding the quality of the images and the places photographed. The study was based on data collected for five locations: Krakow, Warsaw, Wisla, Auschwitz and Dachau. The affective responses were quantified on a two-dimensional positive-negative-scale based on a corpus - generated lexicon of adjectives with opinion and sentiment values they have generated with the textual data collected (Kisilevich et al., 2010).

Hauthal and Burghardt (2014) developed a methodology for extracting location-based emotions from the written language in the metadata of georeferenced Flickr and Panoramio photos using lexicon-based sentiment analysis. Their approach addressed the complexity of language and grammar by considering various grammatical issue such as negations or words that could have an effect of amplification or attenuation on the affective state expressed. Thus, procedures to deal with these grammatical issues were developed and various visualisations of the extracted emotions were proposed. (Figure 7)

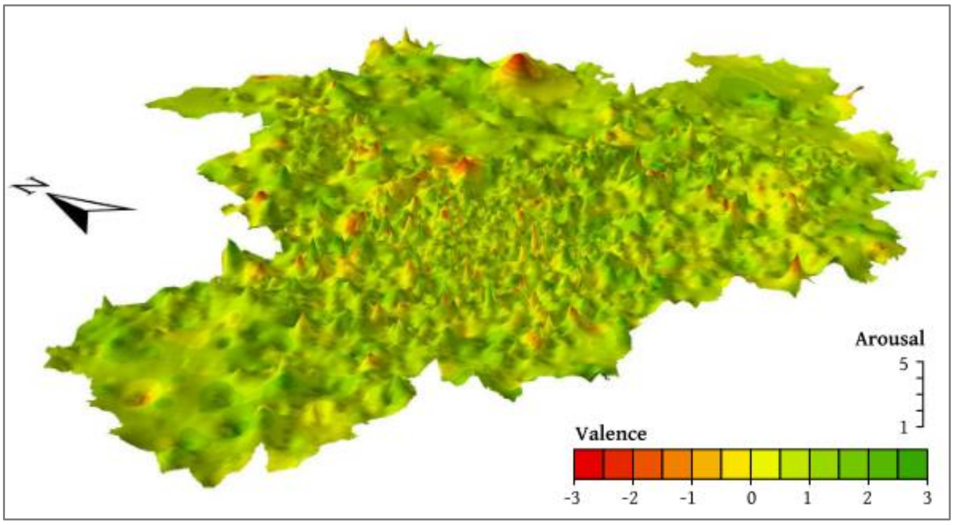


Figure 7 "3D visualisation of valence and arousal in Dresden" (Hauthal, 2015)

2.1.3 Affective Qualities of the Environments

The following quote "Affect is central to conscious experience and behaviour in any environment, whether natural or built, crowded or unpopulated. Because virtually no meaningful thoughts, actions, or environmental encounters occur without affect" (Ittelson, 1973, p. 16; Izard, 1977; Zajonc, 1980 as cited in Ulrich, 2003) illustrates a conjecture that has been by researchers in the field of environmental psychology multiple times proven through a multitude of empirical means. To mention a few examples, Kaplan & Kaplan (1989) analysed the relation between the 'environmental encounters' and the affective state of humans by monitoring the responses of people in an 'outdoor challenging program' and found that natural environments have a positive effect on the emotional state of individuals. Another study conducted by MacKerron & Mourato (2013), that crowdsourced information about people's level of happiness, their location and the context they were in, provided a new line of evidence on the links between nature and subjective wellbeing. In another project, EmoMap, data was collected for the city of Vienna with the help of a mobile application asking people for their feelings in terms of level of comfort, safety, attractiveness or relaxation related to the location they found themselves at. The researchers divided the study area into green spaces with light traffic and with heavy traffic spaces to compare the affective responses of the contributors. As expected, they observed a positive influence of the green spaces on people's feelings (Klettner et al., 2013a).

Mehrabian and Russell (1974) introduced the term *affective quality* to describe the properties of a stimulus prone to produce a change in the affective state of the individual.

They also offered evidence that all stimuli, including large-scale environments, are perceived in terms of their affective qualities. In this context, the affective quality of an environment is "the emotion-inducing quality that people verbally attribute to that place" (Russell & Pratt, 1980, pp. 312) and is often expressed with describing words such as beautiful, ugly, peaceful, hectic, exciting, boring and so on.

Since the user generated content from of social media platforms often contains affective information related to places in the form of words representing emotional expressions of the user or affective qualities of the place, location-based social media data could be considered as a source of extracting information regarding the affective qualities of the environments.

2.2 Location-Based Social Media Data

Entering the era of Web 2.0, the ubiquity of location-sensing mobile devices and the inception of new web services that enable people to share various digital content led to the production of abundant volumes of geo-referenced information. Goodchild (2007) defined this type of information as Volunteered Geographic Information (VGI), which could be described as collections of digital spatial data produced by citizens using appropriate tools to capture and broadcast their views and spatial knowledge on dedicated Web platforms, e.g. OpenStreetMap[6] (OSM), Wikimapia[7], Google MyMaps[8] or Flickr.

In their works Craglia et al. (2012) and Antoniou et al. (2010) classified VGI based on the type of explicit/implicit geographic footprint being captured and the type of explicit/implicit volunteering. Explicit-VGI takes place when the volunteers' main purpose is to focus on mapping activities and explicitly annotate the disseminated data with geographic content (e.g. geometries in OSM). On the other hand, in the implicit volunteering the geographic footprint is just a by-product of the information dissemination, the contributor's main purpose being other than sharing his/her geographic knowledge. Data that is implicitly associated with a geographic location can be: a body of text, an image or a video referring to or associated with a specific location (e.g. geotagged microblogs - Tweets, Flickr geotagged images, Instagram geotagged photos and videos or Wikipedia articles that refer to geographic locations). The volunteered geotagged data coming from social

[6] https://www.openstreetmap.org
[7] http://wikimapia.org
[8] https://www.google.com/mymaps

media services is also referred to as Location-Based Social Media (LBSM) data and it is classified into image-based (e.g. Instagram, Flickr) and text-based data (e.g. Twitter).

2.2.1 Image-based VGI: Flickr and Instagram

Image-based VGI is mostly derived from implicit volunteering on Web platforms such as Flickr, Panoramio, Instagram etc., on which contributors upload pictures of a particular geographic object or surrounding and annotate it with a geospatial reference. The shared photos can be annotated with a variety of metadata that includes textual tags, title, geographic position, and capture time. The process of assigning annotations to the uploaded photos is known as tagging and it has the purpose of making data detectable if searched or browsed again. Golder and Huberman (2006, pp. 203) proposed the following detailed taxonomy of the tags, which comprises of seven categories of tags related to the functions one tag performs:

(1) Identifying what (or who) it is about.

(2) Identifying what it is.

(3) Identifying who owns it

(4) Refining categories.

(5) Identifying qualities or characteristics.

(6) Self-reference.

(7) Task organising.

Furthermore, Sen et al. (2006) following the work done by Golder and Huberman collapse the taxonomy of tags in just three general classes which can be seen in the Table1. While most of the tags provided on photo platforms are factual or personal tags that are subject related, some other tags are subjective tags, that could be used to extract information about the contributor's affective reaction to the object or the surroundings captured in the image uploaded.

Tags Classification by Sen et al. (2006)	**Tags' Functions by Golder & Huberman (2006)**
Factual Tags	• *Identifying what or who is about.* • *Identifying what it is.* • *Identifying who owns it.* • *Refining categories.*
Subjective Tags	• *Identifying qualities or characteristics.*
Personal Tags	• *Self-reference* • *Task organizing*

Table 1 Classification of Tags and their Functions (Sen et al., 2006 and Golder & Huberman, 2006)

Regarding the format of the annotations provided by the contributors of social media data, it is important to mention the term *hashtag*. Normal tags are typically composed of single keywords, whereas a *hashtag* is a unique tagging format represented by a word or unspaced phrase with a prefix symbol, #, that associates a user-defined tag with the content shared. This particular way of tagging has the purpose of facilitating the process of searching and browsing content, proving the users with extra functionalities for organising, sharing, save or publish the search results of social media content.

One last type of tags encountered in the metadata of photos, is the geotag. Geotagging is the process of adding geographical identification metadata to the uploaded media. This metadata is usually represented by either a pair of geographic coordinates (Valli & Hannay 2010) or contributor-assigned geospatial descriptions of the objects/surroundings captured in the form of textual labels, geo-names.

For the practical part of this research, VGI disseminated through the social media services of Flickr and Instagram, two photo sharing services that contain millions of geotagged images contributed by Web users from all over the world, was used.

Flickr, launched by Ludicorp (a Vancouver-based company) in 2004, emerged from tools originally created for a multiplayer online game (Neverending) initially as a chat room (where users could also exchange photographs) and since 2005 as its current form, a commercial image and video hosting service. Users can label the uploaded content with titles, brief descriptions and the content may be also tagged (by the uploader or by other users, if the uploader of the content allows it).

Since its dawn, Flickr has changed ownership several time, in 2005 being acquired by Yahoo! and since April 2018 belonging to SmugMug. Accompanying these changes in ownership, new features have been developed and provided by the service. In 2006, Flickr introduced the geotagging feature. Geotags may be provided by the photo uploader either by means of synchronization with track logs from the built-in GPS of their devices, or by manually locating photos using a map interface. Flickr stores in the metadata of each geotagged image a location accuracy level ranging from 1 (world level) to 16 (street level) and automatically assigns it depending on the precision of the GPS coordinates or the zoom level of the map used to locate the image. When geotagging their content, users have the possibility to either georeferenced the location from which the photo was taken or the location of the immortalized scene. However, in case the case GPS coordinates are collected from their devices, their standpoint is the one georeferenced (Hollenstein & Purves, 2010).

Instagram, launched in 2010 by Kevin Systrom, is a photo and video-sharing social networking service currently owned by Facebook Inc. The platform allows users to upload photos and videos and share them with their followers directly on Instagram or on other social media platforms such as Facebook and Twitter. The uploaded content can be edited with various filters and organized with tags.

Users have also the possibility to geotag their content, however they are constrained to the selection of the location from a pre-defined list of places provided by Facebook. When uploading the image, a list of suggested locations is provided. If the photo to be uploaded to Instagram contains geographic coordinates in its EXIF (Exchangeable Image File Format) metadata, the Instagram application suggests places that are near the coordinates provided (EXIF metadata contains geographic coordinates if the built-in GPS of the device used to capture the image was activated at the moment the image was taken). In case this type of metadata is not available, the application suggests locations near the current upload location identified by their device (Cvetojevic et al. 2016). In case the location they want to share cannot be found in the pre-defined list, the user has the possibility to “create a new place” on Facebook. Consequently, a geotagged image does not provide the exact geographic coordinates of the location from which the image was taken or uploaded, fact that leads to a series of issues concerning the location accuracy of data.

2.2.2 Benefits and Challenges of Using Location -Based Social Media Data

LBSM data has become a valuable stream of information, as a supplement to more traditional datasets in environmental and geographical analyses. This data contains important volumes of information about people's travel itineraries, activities or behaviour in various environments. These large volumes of these particular types of information could be hardly gathered through more traditional meanings of collecting data such as surveys, or even through much modern approaches such as crowd-sourcing. As a result of the availability, economy and the potential to include more diverse, more detailed, more local, and more contextualized data, mining geotagged social media data has gained significant attention in the last couple of decades. Topics such as landmark and hotspot discovery from geotagged photos (Kisilevich et al., 2010; Yang et. Al, 2011), place semantics extraction (Hollenstein & Purves, 2010), behaviour modelling (Jankowski et al, 2010), context-aware location recommendations (Huang, 2016) or extraction of space related emotions (Hauthal & Burghardt, 2016) exemplify the research directions that this data has initiated.

The disadvantages of using LBSM data are mostly represented by the quality issues associated with it. The quality issues could be divided into positional inaccuracies and semantic inaccuracies. As mentioned before, there exist several approaches to geotag the content share on Flickr or Instagram: storing the geographic coordinates with the use of GPS, manually positioning the photo on a map interface or choosing the location from a pre-defined list of places. Thus, positioning inaccuracies can be the result of faulty GPS measurements, contestable level of spatial awareness of the user or location mismatch due to the constraint of choosing from a pre-defined list. According to Hollenstein and Purves (2010) the majority of Flickr photos have assigned an accuracy level between 12 and 16. Zielstra and Hochmair (2011) found that concerning Flickr data for Europe the mean error distance for positional accuracy in is 58.5 m. Cvetojevic et al. (2016) have studied the positional accuracy of objects in a set of Instagram photographs. They found that the offset between the identified objects and the Instagram location ranges from 2 m to 24 km (median: 85 m, mean: 635 m) and 52% of the locations investigate were less than 100m from the object. They correlate the positional inaccuracies of Instagram photos with several possible reasons. They consider that large offsets are caused by the user's inclination towards increased privacy (i.e., obscuring his or her exact location), lack of local spatial knowledge or the absence of an appropriate Instagram location nearby.

Since users can freely annotate and tag the content they are sharing, the text content of social media has a low degree of formal semantic and syntactic accuracy. This issues such as misspelling, typing errors, polysemy and synonymy, the use of different inflection forms and in particular of multiple-words- hashtags make the automatic analysis of the data very complicated. Therefore, text normalisation strategies need to be applied before performing any data analysis.

2.3 Pedestrian Navigation

2.3.1 Human Navigation

Daniel Montello (2005, p. 257) defines navigation as the "coordinated and goal-directed movement through the environment by organisms or intelligent machines" and characterized it as a process which consists of two components: locomotion and wayfinding. Montello defines locomotion as the coordinated physical movement of an entity in an environment, the movement being constrained by the proximal surroundings. Wayfinding, on the other hand, is described as the efficient, goal-directed and planned movement of an entity in an environment.

Downs and Stea (1973) defines wayfinding as a four-step process comprising of orientation, route selection, route monitoring and destination recognition. According to research in peoples' spatial cognition and behaviour, each of these stages is influenced by a series of factors that could be classified as follows. On one hand, there are a series of internal factors consisting of socio-demographic characteristics such as gender, age and health condition (Daamen and Hoogendoorn, 2003) or even culture, beliefs, level of education, lifestyle and attitudes (Holden,2000). On the other hand, there are also a series of external factors that influence pedestrians' behaviour and these are represented mainly by the characteristics of the navigation task (length of the trip, level of familiarity), the qualities of the infrastructure (type, attractiveness, shelter) and environmental properties (ambient, weather conditions) (Daamen and Hoogendoorn, 2003). Millong and Schnechtner (2008) provided a further classification of the external factors according to different dimensions of the route qualities into physical (distance, activity), emotional (attractiveness, safety) and cognitive qualities (complexity, landmarks).

In the past, people had to refer to maps or simply to asking other people for instructions to find their ways in unfamiliar environments. The increasing availability of mobile devices with embedded positioning sensors led to the emergence of mobile navigation systems (Millonig and Schechtner, 2008) that are designed to assist their users in the wayfinding process. These mobile navigation systems were initially designed for vehicle navigation and are currently the most mature and successful applications of Location-Based Services (LBS). Nevertheless, due to the expanding use of personal smart phones and their advanced technical features that support personal navigation tasks the focus shifted towards the development of pedestrian navigation systems (PNSs).

2.3.2 Navigation Systems

Navigation systems, both pedestrian and vehicle, usually consist of three modules: positioning, route planning and route communication (Huang & Gartner, 2009). Typically, the route planning module is based on a graph data structure. A graph is defined by a list of nodes (vertices) which usually represent the street junctions and a list of edges which connect sets of two nodes and represent the street segments. The graph might be directed or bi-directional, depending on the mode of transportation and the traffic directions. In a directed graph, edges represent one-way relationships, while in a bi-directional graph the connections go both ways. The graph's edges are associated with weights representing the street segments' lengths, speed limits or other information relevant to the navigation tasks such as traffic information or even street illumination, comfort level, safety level and so on (in the case of route planning for pedestrians). To automatically plan a route based on the navigation graph, the classic Dijkstra algorithm and A* algorithm are commonly employed (Huang et al., 2014) and conventionally, the shortest or fastest routes are provided.

PNSs formerly applied the same functionalities and design guidelines as the Vehicle Navigation Systems (VNS), yet the user requirements of PNSs are markedly different from those of VNSs. Research in the field of pedestrian navigation has identified the main aspects, in which the users of PNSs differ from those of VNSs and they are as follows:

- pedestrians have a higher degree of freedom in the locomotion; thus, limiting their locomotion only along road networks is not suitable for their needs (Corona & Winter, 2001)

- while drivers can rely only on GPS positioning technology, due to the mixture of environments (indoor and outdoor) pedestrian navigate through, additional positioning technologies that function indoors are required.

- drivers are less conscious about their surroundings because they have to concentrate on the traffic rules and flows; however, pedestrian do perceive their surroundings and the perceptions might often influence their decision-making process during navigation.

- pedestrian rely on the environmental cognition and perception when navigating; thus, the metrical verbal instructions of VNSs are not appropriate for pedestrians and could rather be replaced by instructions related to environmental information, such as salient features or objects (Rehrl et al. 2010; Michon & Denis 2001).

In order to ensure the usefulness (utility and usability) of PNSs, navigation systems should be human-centred and the previously mentioned needs and constraints of pedestrians should be considered and integrated in the design process.

2.3.3 Human-Centred Pedestrian Navigation Systems

In the last decade, numerous researchers have been interested in improving navigation systems and significant progress has been achieved concerning the design of human-centred PNSs. Concerning the route planning module of the navigation systems, new algorithms and approaches for finding optimal routes meeting other criteria than just distance and durations have been proposed.

In order to meet the needs of simplicity, Winter (2002) proposed the computation of “best” route in terms of routes with minimal number or turns or routes with minimal angle, while Haque et al. (2007) proposed the calculation of routes minimizing the number of complex intersections with turn ambiguities. For the same purpose, Jiang and Liu (2011) proposed a fewest-turn route algorithm. Their approach to calculate routes relies on the connectivity of natural roads, which are joined road segments that perceptually constitute good continuity. After testing their approach, they suggested that the routes derived possess fewer turns than the ones derived based on conventional street segments graphs and they are generally more effective and favoured by users due to increased comprehensiveness.

The level of safety a user might experience, has also been the focus of several studies in pedestrian navigation. Miura et al. (2011) consider that the illumination of streets is an important aspect concerning the level of safety. Thus, they developed a system that computes pedestrian routes considering the illumination of streets sensed by a network of wireless sensor devices and evaluated their system through computer simulations. Similarly, Bao et al. (2017) proposed a pedestrian routing algorithm that considers the street illumination and the width of the sidewalks. In their algorithm, each street segment receives a score based on the two before mentioned variables and their associated weights (the weights differ depending on the time of the day and were defined based on a survey with 25 participants).

Recently, Nowack et al. (2018) presented their approach of quantifying the attractiveness of a route in terms of presence of green areas, social places and noise. They created a system that enables the users to define to what extent they prefer one of the three qualities mentioned above and generates customized pedestrian routes entirely based on data from OpenStreetMap. The evaluation of their systems, among others, concluded that 90% of the users that took part in their experiment claimed that they preferred the pleasant and interesting alternative routes over the conventional shortest ones.

Most of the approaches that aim to improve pedestrian navigation application are based on objective, physical parameters. However, as it was previously mentioned and attested, people affectively perceive and evaluate their surroundings and by doing so they develop subjective relations with the environments.

The “AffectRoute” project (Huang et al., 2014) acknowledged the importance of considering these human-environment subjective relations and suggested that in order to meet users’ needs concerning safety or attractiveness, people’s affective responses to environment should be integrated in the route computation. Through a crowdsourcing approach, affective data regarding the level of comfort the participant experienced in certain locations was collected. Subsequently, they aggregated (averaged) the subjective spatial ratings of similar users to model/approximate current user’s subjective relations to different street segments in the street network. The results of their experiment and empirical evaluation of the algorithm showed that routes (“AffectRoutes”) generated considering people’s affective responses to environments were preferred over the conventional shortest ones.

These studies, among others, represent important steps that have been made in the direction of tailoring pedestrian navigation to people's needs and proved that considering the way humans perceive the environment affectively is of great significance for providing human-centred PNSs. Even with the growing popularity of crowdsourcing in the era of Web 2.0, collecting enough data for a successful implementation of the routing algorithm at a larger scale is quite challenging since the collection requires intense user involvement. Consequently, it is important to find additional meanings of collecting this subjective data and this represents the purpose of this thesis.

3 PROPOSED METHODOLOGY

This chapter provides a detailed description of the methods proposed by this master thesis to initially extract people's affective responses to the environment from the metadata of Flickr and Instagram photographs and subsequently aggregate the extracted data and create an affective data layer to be further used for pedestrian navigation purposes.

The methodology relies on the assumption that the content that people share on Flickr and Instagram often contains emotional expressions with regard to the visited or transited places and builds upon the theory that environments are perceived in terms of their affective qualities, which are "emotion-inducing qualities that persons verbally attribute to that place" (Russell & Pratt, 1980, pp. 312). Consequently, lexicon-based sentiment analysis will be performed on the metadata collected to decode the affective responses and a proximity analysis on the extracted data will be used to aggregate (in this case, by averaging) the affective responses in relation to a navigation graph. A detailed description of the conceptual work is presented in the following sections.

3.1 Workflow

To generate the generic solution and address the research objectives outlined in the first chapter, the problem statement was divided in three different phases, as indicated in Figure 8.

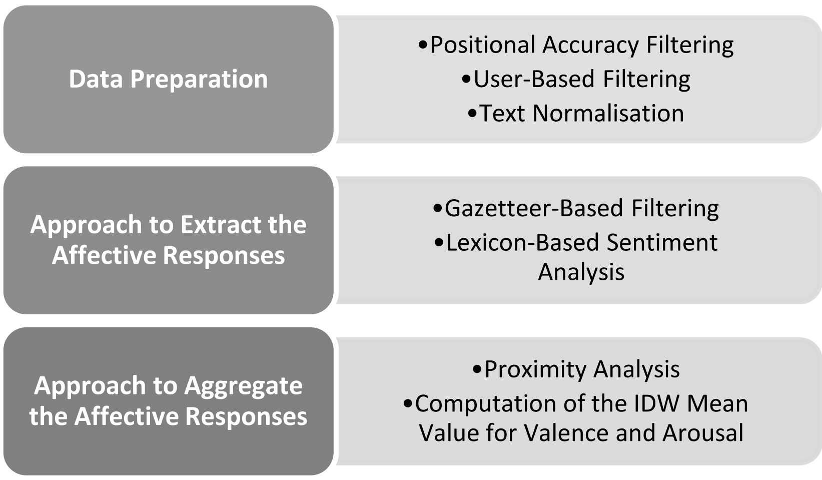


Figure 8 Methodology Overview

The first phase comprises the data preparation and it is represented by a series of filtering processes the data needs to be subjected to before being analysed further. The second phase is represented by the extraction per se of the affective responses with the aid of sentiment analysis methods, tools and resources, whereas the final phase constitutes the approach to aggregate the extracted affective responses and integrate them with a routing graph.

3.2 Data Preparation

As discussed in chapter 2, LBSM data comes with different levels of positional accuracy according to the geotagging method that was used. Since the purpose of this work is to integrate the extracted affective responses with a routing graph, a street positional accuracy level is required. Thus, the entries of the data that do not comply with these requirements, should be filtered out.

Furthermore, after inspecting the raw data it could be observed that user have the tendency to upload for the same location a set of images and use identical metadata (title, description, tags) for all of them. To avoid user-biased valence and arousal values, duplicate metadata should be removed.

As previously mentioned in chapter 2, one of the disadvantages in using LBSM data is that the textual content has a low degree of formal semantic and syntactic accuracy. In most of the content provided, the language register of the users is informal or even vulgar. Users create their own words and spelling shortcuts, often misspell or type the words erroneously, they use slang, genre terminology, abbreviations and words from multiple languages, include hyperlinks, emoticons and other users' names (as mention tags) and concatenate multiple words to create hashtags. Thus, text normalisation strategies need to be applied and such text demands to be corrected before performing any data analysis. In order to provide only significant information for lexicon-based sentiment analysis, in general, the metadata of a geotagged photograph should not contain hyperlinks, mentions (i.e. @username), emoticons, numbers and other characters besides letters and punctuation. Furthermore, the language identification is crucial to be able to process the data further. According to the language identified, the misspelling and typing errors need to be corrected. In this thesis textual data in English and German language was analysed.

"Hashtagging" has become very popular in the last few years and users use it extensively to describe the media they upload. Thus, the hashtags are important metadata and they should be included in the data analysis. Typically, hashtags contain no delimiters between words (i.e. #dresdenmagiccity, #amazingcastle, #schloessernachtdresden etc).

Subsequently, word segmentation is an important step in natural language processing and data preparation, otherwise it is difficult to derive accurate meaning and sentiment classification from a piece of text without first determining the words that it comprises. For the extraction of single words from multi-word hashtags, an algorithm that recursively break the substrings of the input and searches for the longest matching word in a list of words (English or German depending on the language detected in a previous step) until it returns a complete set of meaningful words, is employed. The hashtags are then segmented into meaningful English or German phrases. The English collection of words contains 127156 entries and was assembled by combining the WordNet®[9] lexical database (a database of English nouns, verbs, adjectives and adverbs that are grouped into sets of cognitive synonyms) with the gazetteer compiled for place name identification (which will be mentioned in the next section). For the compilation of the German words' list, the same gazetteer was merged with a list of German words with slightly more than 1.9 million entries (including inflected forms) available online and created by Jan Schreiber (2017)[10].

3.3 Approach to Extract the Affective Responses to the Environment

The approach proposed for the extraction of the affective responses is based on two assumptions:

(1) the content that people share on Flickr and Instagram often contains emotional expressions with regard to visited or transited places and these expressions are not explicit, but rather coded with affective connotations of the words people choose to describe the content shared. Thus, a photograph of a place the user likes or perceives as attractive and comfortable is described and tagged using words with a positive affective connotation (e.g. 'beautiful', 'safe', 'clean', 'peaceful'), whereas the words to describe a less attractive and uncomfortable environment would have a rather negative affective connotation (e.g. 'dangerous', 'loud', 'dirty', 'ruin', noise, 'desolated').

(2) people often include a place name (e.g. 'Zwinger', 'Frauenkirche') or a place noun (e.g. 'museum', 'church', 'street') in the title, descriptions or tags of a photo when they are sharing content about the environment.

[9] https://wordnet.princeton.edu/
[10] https://sourceforge.net/projects/germandict

This approach considers only metadata entries that refer to specific places situated in the study area and assumes that users do not have to explicitly describe places by correlating adjectives with place names or place nouns (i.e. 'safe street', 'nice place', 'horrible building') and that as long as in the metadata of an uploaded photograph a place name is included, an affective state towards that place or an affective state experienced at that place could be expressed (e.g. 'Feeling amazing at #dresden #neustadt', 'Baustelle BlausesWunder#Dresden#Sachsen#Saxony#BlauesWunder#','Kuschelig! #dresden #st ädtetrip #reisejule #flaneuse #spazierganggeschichten'). In the interest of filtering out the entries that are not related to places or environment, a rudimentary version of a gazetteer was compiled by creating a list of the place names (both in English and German) associated with the study area (city of Dresden). The gazetteer is a collection of street and district names, names of points of interest (amenities such as public institutions, food courts and bars; different landuse categories such as parks, gardens, cemeteries; shops and shopping malls; historic sites; touristic places such as hotels and museums) and places of worship (churches, shrines, mosques, synagogues, etc.). The list containing in total 18863 entries was compiled gathering data from OSM, Geonames11 and Wikitravel 12 and taking into consideration the presence of hashtags, the multi-word names repeat themselves by appearing the first time in their original form as single words entries and a subsequent time as concatenated multi-word lexical items. Accordingly, a data mining algorithm is applied on each photograph's metadata and if one of the words or hashtags matched at least one entry of the gazetteer, the affective polarity of the entire metadata along with an average value of the arousal is computed.

To extract the affective responses from the metadata of the geotagged photographs, the lexicon-based sentiment analysis or keywords spotting method was used. Lexicon-based sentiment analysis employs NLP techniques to tokenize, tag the parts-of-speech and lemmatize text into a list of words that are further looked up in affective word lexicons to determine their affective valence and arousal (in some cases). Tokenising is the process of dividing strings into lists of substrings and the tools that are used to perform this task are called tokenizers. For example, the sentence tokenizer splits a phrase into individual sentences, while the word tokenizer can be used to find the list of words in a string. Part-of-speech (POS) tagging, also known as word classification or lexical categorization, is

[11] https://www.geonames.org
[12] https://wikitravel.org

used to classify words according to their grammatical properties and label them accordingly. To exemplify the method, a random photograph was chosen with the following metadata: “A look into the past sins #blackandwhite#blackandwhitephoto#bnw#blacknwhite #blackandwhitephotography#ruins#moodygrams #dresden #dresdencity #streetlife#streetphotography #travelphotography #schloss”. After performing the data preparation processes, the result of POS tagging is the one illustrated in Figure 9.

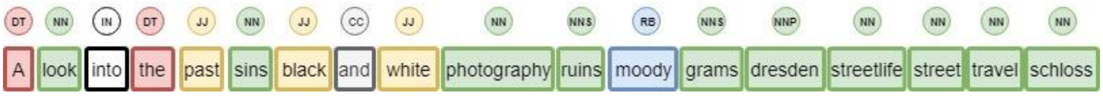


Figure 9 POS tagging explained

As it can be observed, the part-of- speech tags are divided into multiple categories and encoded with specific abbreviations. All possible categories are displayed in the following table.

CC	Coordinating conjunction	NNS	Noun, plural	UH	Interjection
CD	Cardinal number	NNP	Proper noun, singular	VB	Verb, base form
DT	Determiner	NNPS	Proper noun, plural	VBD	Verb, past tense
EX	Existential there	PDT	Predeterminer	VBG	Verb, gerund or present participle
FW	Foreign word	POS	Possessive ending	VBN	Verb, past participle
IN	Preposition or subordinating conjunction	PRP	Personal pronoun	VBP	Verb, non-3rd person singular present
JJ	Adjective	PRP$	Possessive pronoun	VBZ	Verb, 3rd person singular present
JJR	Adjective, comparative	RB	Adverb	WDT	Wh-determiner
JJS	Adjective, superlative	RBR	Adverb, comparative	WP	Wh-pronoun
LS	List item marker	RBS	Adverb, superlative	WP$	Possessive wh-pronoun
MD	Modal	RP	Particle	WRB	Wh-adverb
NN	Noun, singular or mass	SYM	Symbol		
		TO	to		

Table 2 Meaning of Part-of-Speech Codes 13

The approach delivered by this study computes the affective connotation of adjectives, nouns and adverbs. The reasoning behind this choice is that although adjectives usually represent the affective qualities of places (i.e. beautiful building, quite place, noisy area), single nouns could also be regarded as an expression of the affective qualities perceived (e.g. this place is a ruin). Furthermore, adverbs in verb-adverb constructions (e.g. I feel comfortable, I feel happy) represent as well expressions of the affective qualities or of the affective state itself. Subsequently, after classifying the words according to the part-of-speech role they fill in the sentences, words that do not fall under the categories of nouns, adjectives or adverbs are filtered out and the rest of the words are lemmatised. Lemmatisation, in linguistics, is the process of reducing inflectional forms and sometimes derivationally related forms of a word to a common base form, called lemma (Bird

[13] Source: https://pythonspot.com/nltk-speech-tagging/

et al., 2009). After tokenisation, part-of-speech tagging and lemmatisation, the words with no affective connotation need to be removed in order to keep only relevant data.

The affective connotations of the single words are then quantified with values for the dimensions of valence and arousal retrieved from an affective lexicon, these being the lexicon created by Warriner et al. (2013) for the English words and the BAWL-R lexicon for the German words. The emotional valence and arousal of the English words within the Warriner et al. lexicon is rated on a 9-point positive scale (ranging from 1 to 9) while for the German words, BAWL-R's emotional valence is rated on a 7-point scale ranging from -3 (very negative) through 0 (neutral) to +3 (very positive) and arousal is rated on a 5-point scale ranging from 1 (low arousal) to 5 (high arousal). We apply the same polarity classification approach as BAWL-R. Thus, to be able to combine and compare the data extracted, the valence and arousal ratings of Warriner et al. needed to be rescaled.

3.4 Approach to Aggregate and Model the Extracted Affective Responses

Typically, the edges of a routing graph are associated with weights representing the street segments' length and speed limits, the best routes being computed through route cost functions that take into consideration these weights assigned. The basic idea behind this approach is to aggregate the extracted affective responses to model the collective affective perceptions of the environments and encode them as a weighting variable for each street segment of a navigation graph. In this way, the encoded collective affective ratings could be integrated in the cost function that computes the best route.

3.4.1 Spatial Extent of the Affective Responses and Aggregation Method

The approach proposes as an aggregation method the computation of the arithmetic mean value for valence and arousal values of all the affective responses relevantly located for each street segment.

Pippig et al. (2013) propound the idea that all geospatial features occupy three different kinds of spatial extents: a tangible extent, a perceptual extent and an awareness extent. They defined the tangible extent of a feature as the space occupied within its physical boundary and the perceptual extent as an idealised space in which a specific feature is sensually (visually, audibly or olfactory) noticeable (Pippig et al., 2013, pp. 227). To exemplify, if one of the affective responses extracted is related to the sight of an abandoned building or the occurrence of disturbing noises, the tangible extent is represented by the exact location of the object or the exact location where the noise originates from while the perceptual extent is represented by the space from which the building can be seen or the noise can be heard.

Following this idea, in this approach the perceptual spatial extent is defined as the space, whose qualities could be affectively perceived by the pedestrian when located on a specific street segment. Out of practical reasons and to enable the implementation of the approach, in the proposed methodology, the perceptual spatial extent is regarded as the space represented by the close proximity and is limited to a 100m buffer area around each street segment.

Nevertheless, since the choice is not based on empirical evidence, the same reasoning, on which the Inverse Distance Interpolation method is based, was applied and the approach considers that each affective response extracted has a local influence that diminishes with distance. Consequently, in the computation of the arithmetic mean value of the valence and arousal values for each street segment, the discrete values closest to the street segments receive greater weights and the weights diminish as a function of distance.

Thus, the aggregation approach proposed computes the inverse distance weighted arithmetic mean for the valence and arousal ratings of all extracted affective responses situated within a distance of 100m from each street segment, as illustrated in Formula 1 and Formula 2:

$$mean\_valence = \frac{\sum_{i=1}^{n} \frac{v_i}{d_i}}{\sum_{i=1}^{n} d_i} \qquad mean\_arousal = \frac{\sum_{i=1}^{n} \frac{a_i}{d_i}}{\sum_{i=1}^{n} d_i}$$

Equation 1 Weighted Average Valence   Equation 2 Weighted Average Arousal

(where v_i is the valence and a_i the arousal value of the extracted affective response i and d_i represents the distance between affective response i and the street segment to which is assigned) and stores the values as additional attributes of the street segments.

3.4.2 Quantification and Classification of the Affective Ratings

For the quantification and classification of the affective ratings, a bipolar framework is approached, which is organized around the dimensions of bipolar valence and arousal and is largely based on the circumplex model of affect (Russell, 1980). The circumplex model of emotion suggests that affective states are distributed in a two-dimensional circular space, defined by the arousal and valence dimensions (see Figure 10).

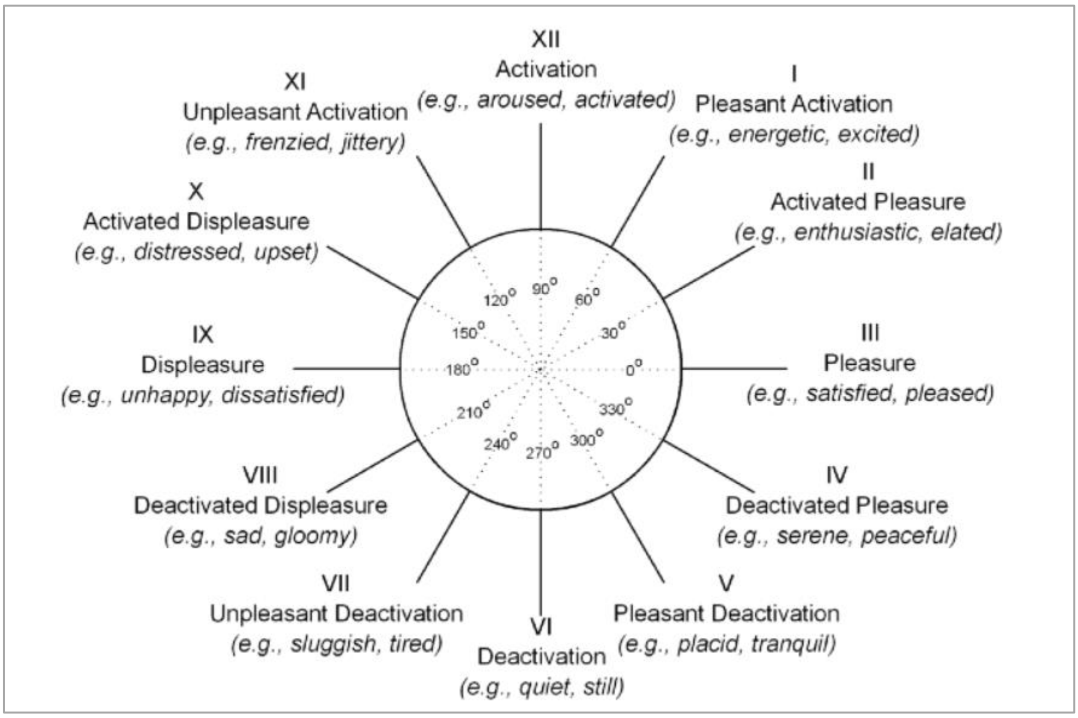


Figure 10 *The 12 – Point Affect Circumplex Model* (Yik et al., 2011)

The framework employs a single bipolar scale ranging from positive to negative to measure subjective valence. The changes in reported valence are subjectively different from changes in reported arousal, so this framework also employs an additional subjective arousal scale ranging from low (e.g. sleepy/inactive) to high (e.g. excited/active). The emotional states (affective ratings) can be represented at any level of valence and arousal, or at a neutral level of one or both of these factors.

Thus, the collective affective rating computed for each street segment is then plotted on the two-dimensional affective space to determine the affective class to which the street segment belongs. To rank each of the street segments contained by the routing graph, five possible affective classes (depicted in Figure 11) were conceptualised as follows:

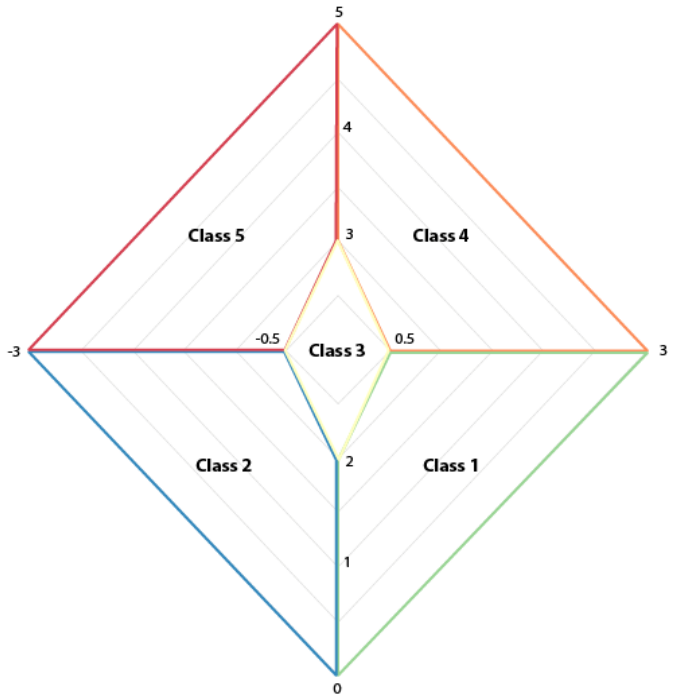


Figure 11 Affective Classes (the horizontal axis represents the valence dimension, whereas the vertical axis represents the arousal dimension)

(1) positive valence and low arousal class (valence > 0.5 and arousal < 2.5). In this case the affective responses could reflect a pleasant, comfortable, calm, attractive environment such as gardens and parks.

(2) negative valence and low arousal class (valence <- 0.5 and arousal < 2.5). This class would include environment that evoke sentiments of sadness and dissatisfaction. For example, war memorials or places associated with sad events.

(3) neutral valence and neutral arousal class (-0.5 <= valence <= 0.5 and 2<= arousal<=3). The particular values for the valence and arousal were chosen after observing the data extracted; due to the numerous affective ratings that have slightly positive or negative words (values between 0 and +/- 0.5), this approach proposes the mentioned interval [-0.5,0.5] as neutral valence values. Moreover, it was observed that most of the affective responses that fall in the category of neutral valence, have arousal values between 2 and 3. This class would describe rather ordinary environments that do not particularly evoke any emotions in the pedestrians. An example of such an environment could be residential areas.

(4) positive valence and high arousal class (valence > 0.5 and arousal > 2.5). This would be the case of energetic and pleasant environments such as touristic areas of a city, sport centres, shopping areas.

(5) negative valence and high arousal class (valence < -0.5 and arousal > 2.5). This class would describe areas and environment that evoke feelings of fear, terror and in general high discomfort such as infamous suburbs where the rate of criminality is high, sites of unpleasant events such as attacks and revolts or noisy areas with intensive traffic.

4 IMPLEMENTATION OF THE METHODOLOGY

This chapter reveals how the proposed methodology was implemented for a trial area, describes the data basis, the algorithms developed for each approach and the software resources used.

4.1 Data Basis

For the realisation and the implementation of the methodology, data collected for a trial area, namely the city of Dresden, was used. The dataset comprised of 137,421 Flickr and 708,123 Instagram geotagged photos and has been provided by the Institute of Cartography within the Technical University Dresden. The Flickr data was retrieved for the time period between 2007 and 2018, whereas the Instagram data for the time span between 2011 and 2018. The Instagram and Flickr data was retrieved using the API available.

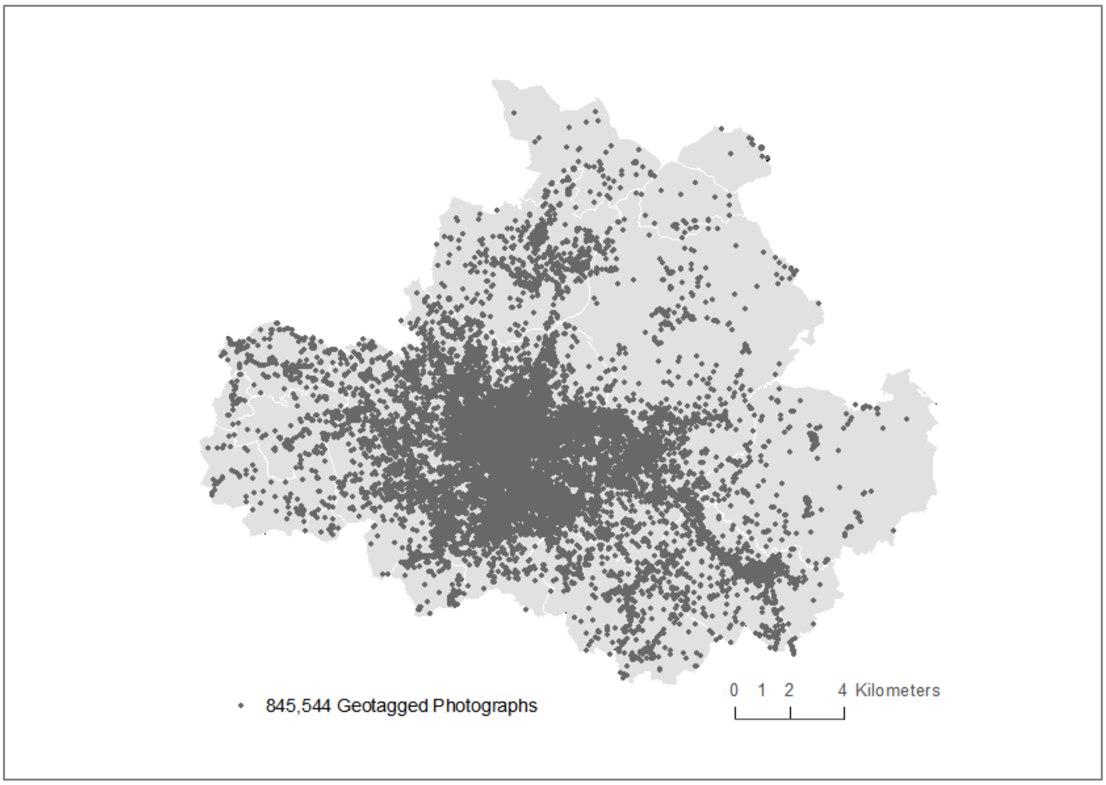


Figure 12 Spatial Distribution of the Geotagged Photographs

Both services required registration. Regarding the privacy and user identity issues that user generated content normally raises, according to the provider of the data, special measures have been taken to ensure the compliance with legal requirements regarding data privacy. Thus, the data was processed with a series of methods to ensure the privacy of users. First of all, the Global Unique Identifiers (GUIDs), such as UserIDs, PlaceIDs, PostIDs, have been Hashed with sha256 and only information that was explicitly made public by users was provided. Furthermore, other identifiers and private information such as usernames, biography etc. references to the original data (such as URLs to Online Posts/User Profiles) were removed. Finally, the data provided contains only metadata and summaries (e.g. Total Like Counts, View Counts, etc.), not the actual images themselves.

Moreover, for the implementation of the aggregation and modelling methodology a routing graph was required. The underlying network elements are based on OSM and the graph was downloaded and converted to a Feature Class[14] (an ESRI's format for a collection of geographic features with the same geometry type, such as node, line or polygon, having the same attributes and the same spatial reference) using the OSMnx python tool developed by Geoff Boeing (2017)[15].

4.2 Data Preparation and Extraction of the Affective Responses

4.2.1 Developed Algorithm

For the implementation of the methodology presented in the previous chapter, an algorithm that pre-processes the data and extracts the affective responses was developed using Python[16] encompassing a series of steps that could be grouped into two phases, i.e. the *Data Preparation* phase and the *Extraction of the Affective Responses* phase (depicted in Figure 13).

[14] GIS dictionary: Feature Class - http://support.esri.com/other-resources/gis-dictionary/term/feature%20class
[15] https://github.com/gboeing/osmnx
[16] https://www.python.org/

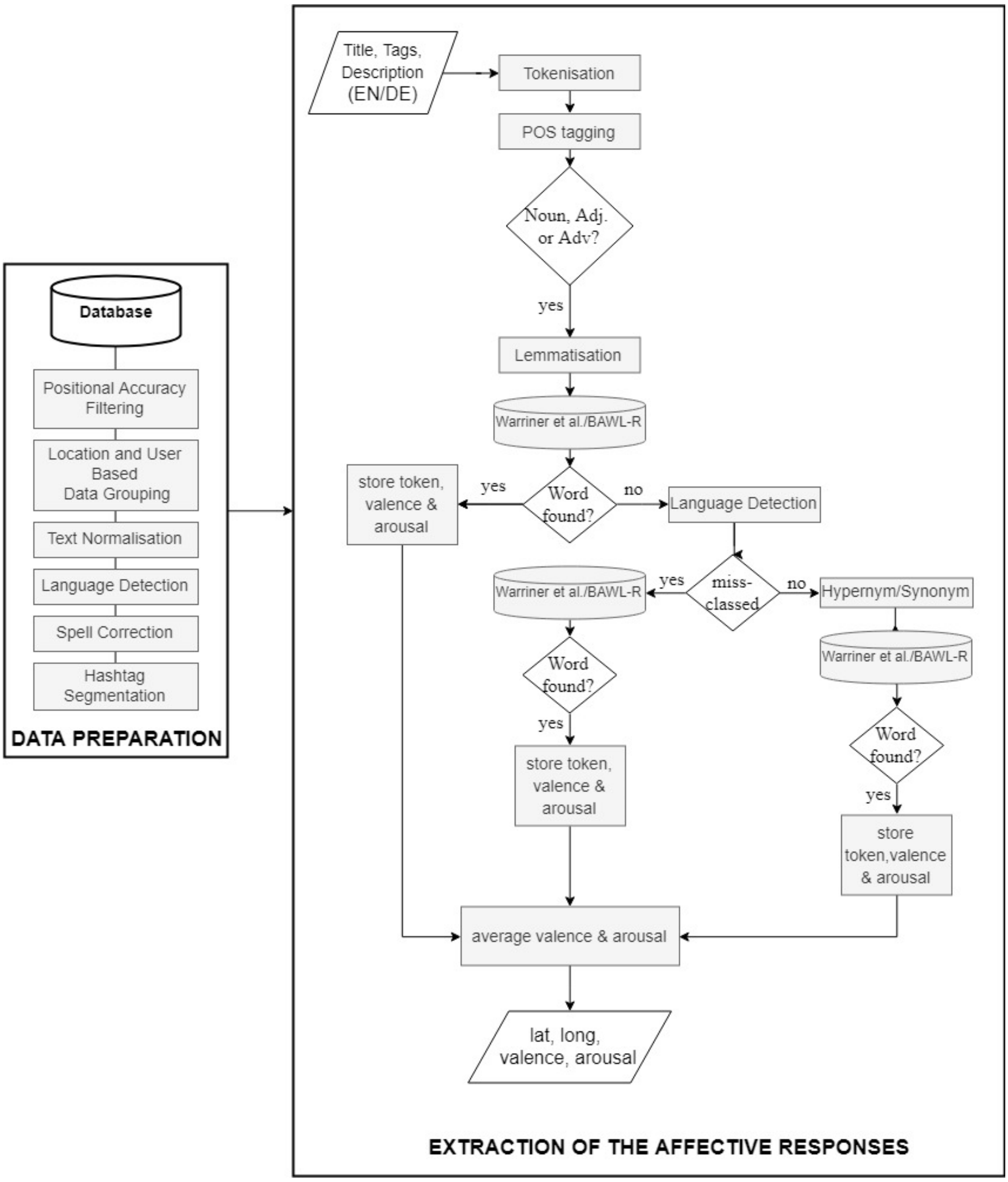


Figure 13 Implemented Algorithm for the Extraction of the Affective Responses

In the first phase, it was crucial to select only the metadata of the photographs with a street level of positional accuracy. In the case of the Flickr data (where the level of accuracy was specified in a field of the provided data), only geotagged images with a level of positional accuracy higher than 14 or classified under the name "street" were selected to be further analysed. Regarding the Instagram data, the process of filtering the data according the level of positional accuracy was more complicated because there is no such data provided by the API of the service. In the search for a practical solution of avoiding large volumes of inaccurate data, the author of this thesis, made use of the *place_name* field provided by the API and manually filtered out all the entries, whose *place_name* values represent large areas such as the entire city ("DRESDEN"), district of the city ("LOSCHWITZ", "STREHLEN" etc) and obviously did not comply with the positional accuracy requirements. This solution provides however just a superficial evaluation of the positional accuracy of the geotagged photos and might lead to inaccuracies in the analysis. Consequently, it needs to be revised. After this step, when parsing the data in the algorithm, the metadata provided by the same user at one single location is grouped under the UserID and the duplicate data is removed. Further, the text normalisation, language detection and hashtag segmentation are performed. The result of this phase is a database containing for each location single entries of the metadata provided by individual users divided into English and German text.

The second phase of the algorithm represents the implementation of the lexicon-based sentiment analysis method. In the manner being indicated, for each entry (metadata of one geotagged photo) of the pre-processed data, for the textual data belonging to each language, the text is tokenised. Afterwards, POS tagging is performed to identify the parts of speech that are of interest in this approach (nouns, adjectives and adverbs) which are then reduced to their base forms (i.e. plural forms to singular, comparative and superlative forms to standard forms). Then each lemmatised word is looked up correspondingly in one of the two affective lexicons used (Warriner et al. for English lemmas and BAWL-R for German lemmas). If the word matches one of the entries in the affective lexicon, it is the stored together with the corresponding valence and arousal values. Otherwise, due to the facts that the language detection tool was applied to sentences and not to single words (out of performance reasons – detecting the language of single words has lower levels of accuracy) and that users tend to use multiple languages in their metadata, it was possible that the word did not match any entry in the affective dictionary because it is of a different language. Subsequently, the language of the single word is detected and in case it results that the word belongs to the other language we consider (in case the word was looked up in the English affective lexicon and it turned out that actually it is a German word), this is looked up in the corresponding affective lexicon and in case it is found it is then stored along with the corresponding valence and arousal values. If the result of the language detection shows that the initial language detected was the correct one, a list of synonyms and hypernyms is generated and each of them is looked up in the affective dictionary until a match was found. In case no match was detected, then the word is skipped or else it is stored in the data base with its emotional values. After iterating through all the tokens of an entry, for the location and user associated with it the average valence and average arousal is calculated based on the following formulas:

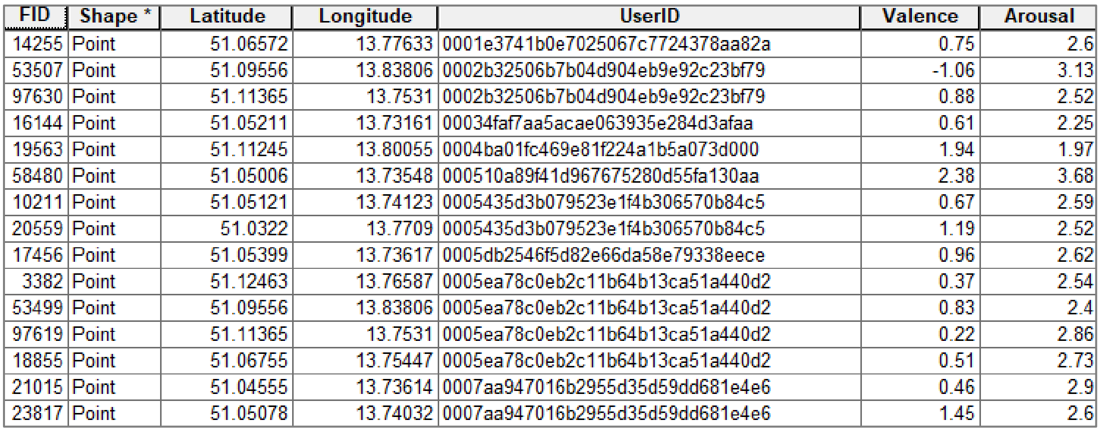
$$Avg\_Valence_{(user)} = \frac{\sum_{i=1}^{n} Pos_{valence(wi)} - Neg_{valence(wi)}}{NumberOfAffectiveWords}$$

Equation 3 Average Valence of User's Textual Entry

$$Avg\_Arousal_{(user)} = \frac{\sum_{i=1}^{n} Arousal(wi)}{NumberOfAffectiveWords}$$

Equation 4 Average Arousal of User's Textual Entry

(where $Pos_{valence(wi)}$ are positive valence values of the identified affective words (wi) and $Neg_{valence(wi)}$ are the negative valence values, $Arousal(wi)$ is the arousal corresponding value of the identified affective word wi and $NumberOfAffectiveWords$ is the total number of words in the textual entry found in the affective lexicon). The values are afterwards stored in a data base together with the coordinates (as depicted in the Figure 14)



FID	Shape *	Latitude	Longitude	UserID	Valence	Arousal
14255	Point	51.06572	13.77633	0001e3741b0e7025067c7724378aa82a	0.75	2.6
53507	Point	51.09556	13.83806	0002b32506b7b04d904eb9e92c23bf79	-1.06	3.13
97630	Point	51.11365	13.7531	0002b32506b7b04d904eb9e92c23bf79	0.88	2.52
16144	Point	51.05211	13.73161	00034faf7aa5acae063935e284d3afaa	0.61	2.25
19563	Point	51.11245	13.80055	0004ba01fc469e81f224a1b5a073d000	1.94	1.97
58480	Point	51.05006	13.73548	000510a89f41d967675280d55fa130aa	2.38	3.68
10211	Point	51.05121	13.74123	0005435d3b079523e1f4b306570b84c5	0.67	2.59
20559	Point	51.0322	13.7709	0005435d3b079523e1f4b306570b84c5	1.19	2.52
17456	Point	51.05399	13.73617	0005db2546f5d82e66da58e79338eece	0.96	2.62
3382	Point	51.12463	13.76587	0005ea78c0eb2c11b64b13ca51a440d2	0.37	2.54
53499	Point	51.09556	13.83806	0005ea78c0eb2c11b64b13ca51a440d2	0.83	2.4
97619	Point	51.11365	13.7531	0005ea78c0eb2c11b64b13ca51a440d2	0.22	2.86
18855	Point	51.06755	13.75447	0005ea78c0eb2c11b64b13ca51a440d2	0.51	2.73
21015	Point	51.04555	13.73614	0007aa947016b2955d35d59dd681e4e6	0.46	2.9
23817	Point	51.05078	13.74032	0007aa947016b2955d35d59dd681e4e6	1.45	2.6

Figure 14 Excerpt from the Resulting Affective Responses Feature Class

4.2.2 Software Resources

Besides the native libraries that Python includes, the following external libraries and tools were used to develop the algorithm:

- tweet-preprocessor library for text normalisation[17]
- for language detection: Polyglot 16.7.4 (a natural language pipeline that supports

[17] https://github.com/s/preprocessor

multilingual applications) [18]

- pyspellchecker library for spell correction.[19]
- Natural Language Toolkit (NLTK 3.4)[20] and German language support for TextBlob by Steven Loria (textblob-de)[21] for tokenisation, POS tagging and lemmatisation.

4.3 Aggregation of the Affective Responses

The purpose of this phase of the methodology was to aggregate the extracted affective responses and incorporate them into the routing graph. Incorporation in this case refers to the classification of the graph's segments according to the valence and arousal values of the aggregated affective responses. The data base resulted from the previous implementation was exported as a Point Feature Class, in which for each individual user identified, the locations of the photographs updated are stored along with the valence and arousal values of the textual metadata provided. For the implementation of the aggregation method conceptualised, the *Generate Near Table* analysis tool of ESRI's ArcGIS desktop software was use together with a Python algorithm, designed to calculate the weighted mean valence and arousal for each street segment (depicted in Figure 15).

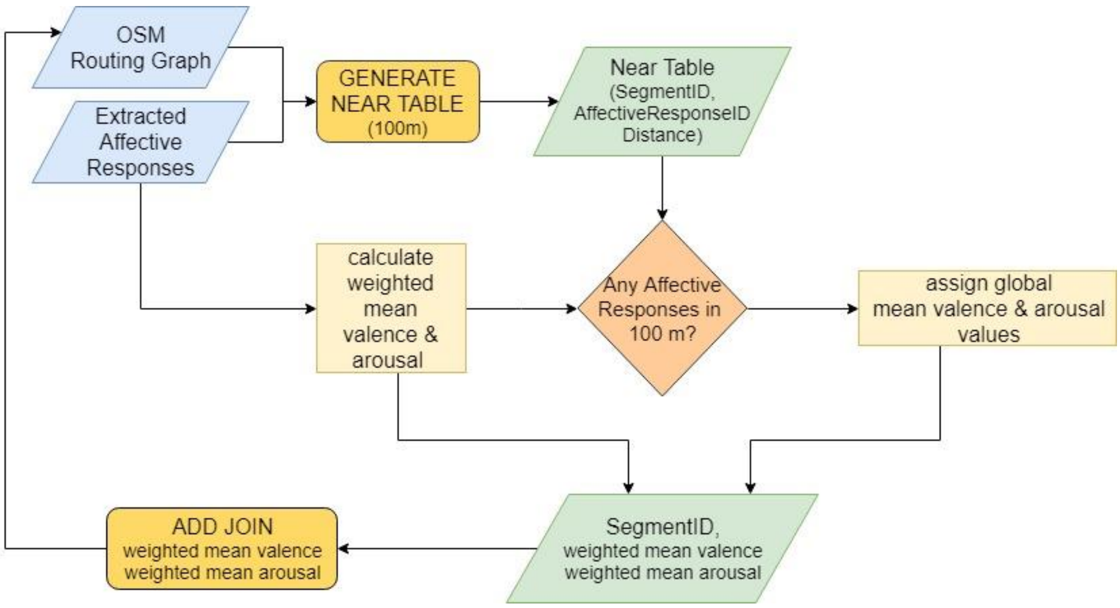


Figure 15 Implemented Algorithm for the Aggregation of the Extracted Affective Responses

[18] https://pypi.org/project/polyglot/
[19] https://github.com/barrust/pyspellchecker
[20] https://www.nltk.org/
[21] https://pypi.org/project/textblob-de/

As an initial step, with the help of the *Generate Near Table* Analysis-tool, for each graph segment all the affective responses situated within a radius of 100m were identified and the distance between the segment and each affective response was measured and stored. Thereupon, an algorithm was developed to calculate the weighted mean valence and arousal values for each street segment, the weighting factor being the inverse value of the distance between each affective rating and the street segment ($\frac{1}{distance}$).

Since for some of the street segments there were no identified affective ratings, to maintain the continuity of data, the approach proposes the calculation and assignment of local average valence and arousal values. Thus, for each district of the trial area the average valence and arousal values of all the affective responses located within the borders is computed and assigned to the graph segments with no available data situated in the corresponding district.

The last step was to integrate the aggregated affective dimensions with the routing graph and classify the street segments accordingly. To effectuate this integration the *Add Join* Data Management tool was used and the classification was performed based on the five previously established classes.

5 RESULTS

Within this chapter the results of the implemented methodology are presented along with a brief performance assessment of the developed algorithms to extract and model the affective responses.

5.1 Results of the Implementation

After applying the algorithm developed to extract the affective responses, 243,882 ratings of the affective qualities of the places, whose names could be found in the compiled gazetteer, or of the affective states experienced by the users in these places were identified. In total, 98,560 distinct users generated the content. As it can be observed in the map below (Figure 16) the affective ratings detected are unequally spatially distributed, most of the data being located in the area of the central, southern and south-western districts of the city, where most of the places of interest and touristic attractions of the city of Dresden are concentrated. This fact influenced the classification of the street segments in the way that some streets had associated numerous affective ratings, whereas others had no associated ratings at all.

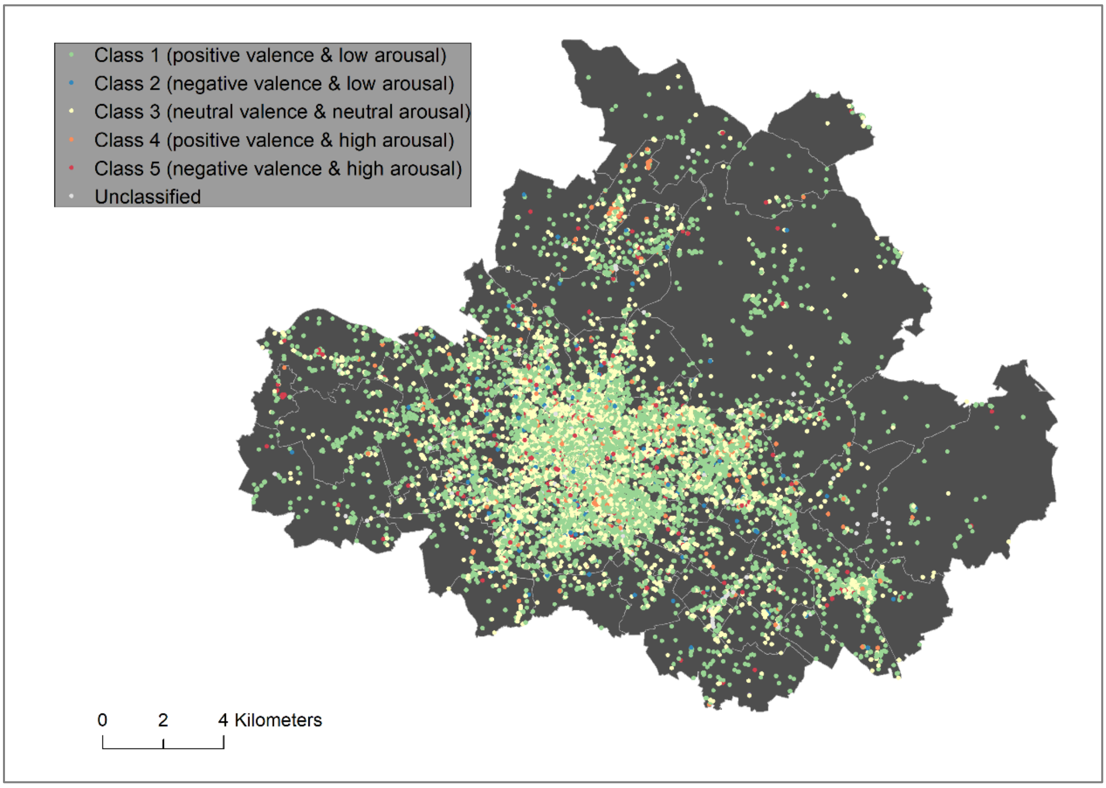


Figure 16 Spatial Distribution and Classification of the Extracted Affective Responses

To gain insight from the data extracted, the individual affective ratings were classified according to the five classes conceived in the methodology chapter. One percentage of the affective rating's valence and arousal values do not match any of the classifications rule and are represented as unclassified. Figure 16 shows the spatial distribution of the classified extracted affective responses and it can be clearly observed that most of the extracted affective ratings (72% of the total – see Figure 17) belong to the first affective class which is represented by positive valence and low arousal values and would generally describe attractive, comfortable or serene places. The second most prevalent affective class (registering 19% of the affective responses) is the class with neutral valence and arousal corresponding values, the other three classes summing up a total of just 8% of the whole extracted affective responses. This could imply the fact that social media users have a general tendency to share positive things including pictures of pleasant places that evoke positive emotions and that the second most common pictures shared on social media platforms are the ones that depict common environments which users encounter on a daily basis (such as the areas where they live, work or study).

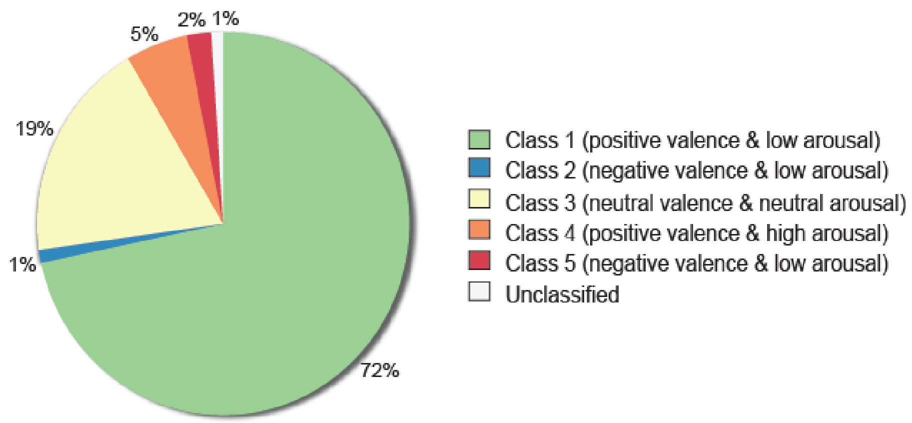


Figure 17 Classification of the Extracted Affective Ratings (according to the 5 defined classes)

Appendix 2 depicts the classified navigation graph according to the values of the collective affective responses obtained through the aggregation of the data. As anticipated, considering the results regarding the classification of the discrete affective ratings, the prevailing affective class is represented by positive valence and low arousal values, with slightly over 50% of the graph's edges (see Figure 18). At a small difference, the next class that comprises a large part of the edges is represented by the third class of neutral valence and arousal values (nearly 49% of the network); however, only for 10% of the edges the aggregated valence and arousal values rely on the discrete data located within a 100m radius from the street segments, the rest of 39% having assigned local average values representative for the district in which the edge is situated. The other three affective classes sum up a total of less than 1%, being isolated cases, whereas only 0.36% of the edges remained unclassified. These results backup another time the supposition that among the users of social media, there is a general disposition to share content with a positive or neutral affective association.

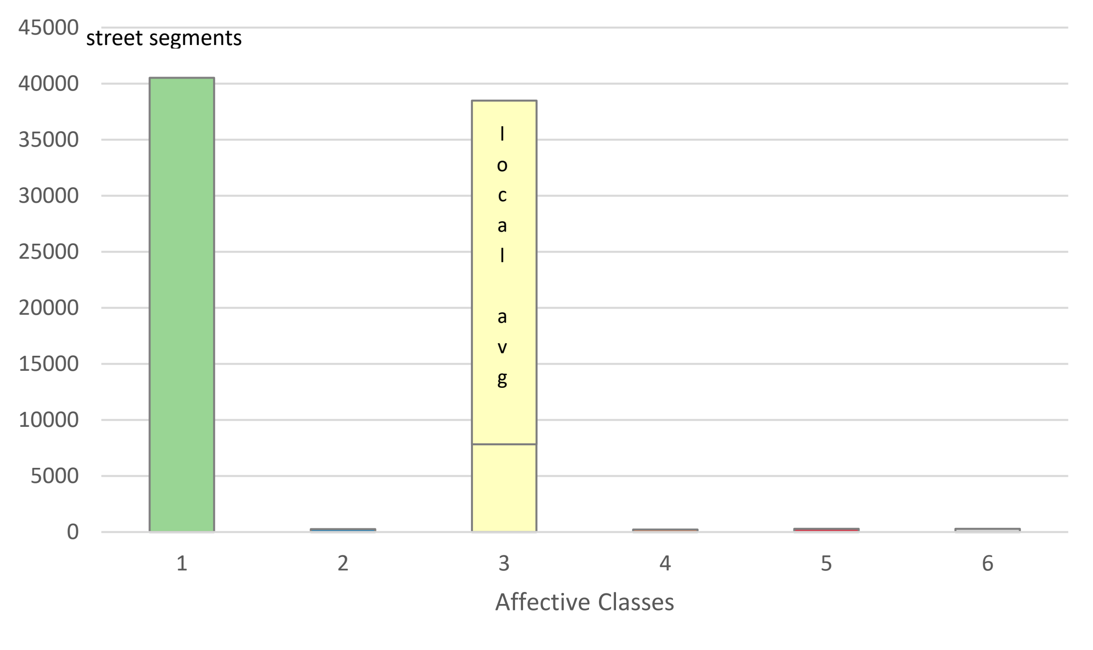


Figure 18 Results of the Classification of the Routing Graph's Edges after the Aggregation of the Affective Ratings

5.2 Evaluation of the Developed Methodology

5.2.1 Assessment of the Data Preparation

In the interest of evaluating the proposed methodology to extract the affective responses to the environment from the metadata of geotagged images, the developed algorithm was applied to a set of 100 randomly selected images from the database (split into 50 Flickr and 50 Instagram images) and each step of the execution was closely monitored.

After parsing the 100 pictures into the database, 75 unique entries were registered containing 34 distinct locations and 62 distinct users that input metadata at a single or multiple locations. Out of the 700 hashtags detected (492 classified as English text and 208 classified as German text), 208 (29.71%) were segmented incorrectly (see Figure 19). As previously mentioned, social media users are inconsistent in their choice of language when tagging the content, they upload. Thus, language detection tools cannot be 100% proficient especially when applied at sentence level and often, if for instance a sentence is composed of 90% English words, it would be classified as English text even though 10% of the words belong to other languages. This issue had repercussions on the following processes such as the hashtag segmentation, finding synonyms/hypernyms and, subsequently, on the sentiment analysis because the words were looked up in the wrong words-collection for hashtag segmentation, synonyms and hypernyms dictionary or affective lexicon. Consequently, as observed during the evaluation, the erroneous results of the hashtags segmentation were mainly caused by the language misclassification (106 cases). Another issue concerning the performance of hashtag segmentation, in particular, and the sentiment analysis, in general, was represented by the lack of domain-specific vocabulary such as social media acronyms, abbreviations, and neologisms, vocabulary specific to photography domain, company names, brands and personalities names (e.g. 'hdr', 'flickrstruerefflection award', 'instamood', 'instacool', 'lowfiphotography', 'bokeh', 'kurtvonnegut', 'nikon' etc.). This was the case of almost 32% of the incorrectly segmented hashtags. Adding to the previous issues, since there was no possible way to detect only the multiple-words-hashtags, all hashtags (including the single-word ones) have been treated equally and went through the hashtag segmentation process (in this case for example, the German word "schloss" belonging to an English sentence became 'loss' after segmentation as well as '#neustadt' became 'tad', 'instacool' became 'taco', 'instagram' became 'stag' and 'gram' or 'dresdenagram' became 'dresden' and 'gram'). Misspelling was also one of the reasons for which the hashtag segmentation performed erroneously (36 cases). Lastly, some users create hashtags that simply contain too many words which makes the algorithm incapable of performing without errors. An example from the sample of data used for evaluation is the following: *#dasgÃ¤nseblÃ¼mchen-bellisperennisauchâ€žmehrjÃ¤hrigesgÃ¤nseblÃ¼mchenâ€œmasliebchentausend-schÃ¶noderschweizerischmargritliâ€žkleinemargeriteâ€œgenannt1isteinepflanzenar-tausderfamiliederkorbblÃ¼tlerasteraceaedaesauffastjederrasenfl.*

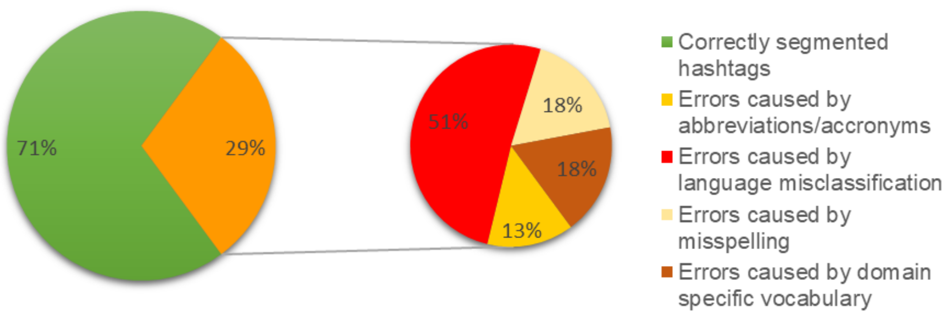


Figure 19 Hashtag Segmentation Evaluation

The identification of entries that contained at least one place name performed very well. Out of total of 117 identified sentences, 57 matched at least one toponym in the compiled gazetteer, number which coincides with the obtained value after manually counting the sentences that abide by this rule.

After the previously performed steps, the filtered metadata contained 857 words (656 English and 201 German words) that could represent an affective state or quality (as discussed in chapter 3, adjectives, nouns and adverbs); 462 words or a corresponding synonym/hypernym had a match in one of the affective lexicons, of which 40 words were German words misclassified as English and 64 words were English words misclassified as German.

5.2.2 Assessment of the Sentiment Analysis and of the Affective Classes Proposed

To evaluate the sentiment analysis performance and the classification of the extracted affective responses, during the execution of the implemented algorithm two lists (one for each language considered) of all the words or the respective synonyms/hypernyms, that matched an entry in one of the affective lexicons, were compiled. Altogether, 14,760,021 English and 156,200 German words were stored, that were subsequently categorised into the five affective classes (based on the values of valence and arousal) in order to determine the most frequent terms for each of the proposed affective classes and establish their relevance in relation with the validity of the classification criteria.

As it can be observed in Table 3, the difference between languages is substantial. The higher number of detected English affective terms was foreseen from the moment the affective lexicons were chosen, since the Warriner et al. lexicon is much more complex than the BAWL-R one, including data from more domains and also informal and vulgar registry words. In addition to this fact, the language misclassification issues, discussed in the previous section, led to this considerable difference between the performances of English and German word extraction.

	Number of English Words found in Warriner et al.	Number of German Words found in BAWL-R
Class 1	708,777	95,188
Class 2	57,872	4,021
Class 3	333,500	16,413
Class 4	248,432	22,251
Class 5	42,158	6,624

Table 3 Difference between the performances of English and German affective words extraction

In Appendix 1, for each affective class and for each language word clouds with the 50 most frequently encountered affective terms are depicted. Thus, for the prevailing positive valence and low arousal class, the most frequent words are represented by items such as ‘tree’ or ‘baum’, ‘garden’ or ‘garten’, ‘schloss’ (‘palace’ in German, ‘nature’ or ‘natur’, ‘architecture’, ‘art or ‘kunst’, ‘park’, whereas for the neutral class, words such ‘building’, ‘street’, ‘winter’, ‘kirche’ (‘church’ in German), ‘arbeit’ (‘work’ in German), ‘platz’(‘place’ in German), ‘strasse’ (‘street’ in German),’urban’ were the most common ones. The most frequently encountered words with low valence and high arousal in both the languages are mainly related to the bombing of the city during the Second World War (e.g. ‘war’, ‘krieg’, ‘bomb’, ‘nazi’, ‘fire’, ‘rage’, ‘attack’, ‘fight’ or ‘kampf’ ‘battle’) or words such as ‘laut’ (‘loud’ in German), ‘tod’ (‘dead’ in German), ‘grafitti’, ‘loss’, ‘addict’, etc. The affective responses to vibrant and pleasant environments were identified through words such as ‘travel’, ‘beautiful’, ‘wunder’ (‘wonder in German),’love’, ‘party’,’ live’, ‘music’, ‘freude’ (‘happiness’ in German), ‘freiheit’ (‘freedom’ in German), ‘sport’, whereas the affective responses to sad and unpleasant places could be pinpointed through terms such as ‘cold’, ‘omen’, ‘industrie’(‘industry’ in German), ‘kitsch’, ‘traurig’, ‘bad’, ‘sad’, etc. After inspecting the word clouds, the presence of relatively high number of representative affective terms confirms that the valence and arousal values ranges used for the classification were suitable for the validity of the classes.

Nevertheless, besides the relevant affective terms mentioned before, a fairly large number of incorrect processed affective terms was observed (see Table 4). Most of the terms are the result of incorrect hashtag segmentation or/and language misclassification. Moreover, a series of nouns, whose affective connotations are not necessarily relevant, in the way that they could represent affective qualities of the environment or affective states towards them, were also taken into consideration. This fact, of course, affects the quality of the data aggregation in the further steps; thus, the selection of the parts of speech relevant to the sentiment analysis should be reviewed and refined.

Incorrectly processed affective term	Frequency	Incorrect Hashtag Segmentation	Incorrect Language Detection	Affective class to which it was falsely assigned
err	1,722	x	x	2
euro	2,503	x		3
gut	4,744		x	2
gram	9,294	x		3
loss	1,394	x	x	5
mar	1,010	x	x	2
mess	606	x	x	2
ram	7,573	x		3
rat	636			5
ravel	3,948	x		4
roe	448	x	x	2
sap	365	x	x	2
shot	4,708	x		5
shoot	1,301	x		5
stag	15,460	x		3
tad	1,426	x		3
tar	246	x		2
wunder	1,664	x		4

Table 4 List of Incorrectly Processed Affective Terms

The results of the evaluation procedures revealed a set of issues that could hinder the performance of the methodology proposed. These issues shall be presented and discussed in the following chapter along with possible solutions and directions for future work that needs to be conducted.

6 DISCUSSION AND FUTURE WORK

In this chapter the issues and weaknesses of the proposed methodology are discussed. Furthermore, possible technical improvements and directions for future work are suggested according to the findings of this research.

6.1 Weaknesses and Limitations of the Proposed Methodology

Throughout the conducted research a considerable number of issues concerning the nature of LBSM data and, as a consequence, an important number of difficulties regarding the analysis of such data could be identified.

It is important to mention that the first limitation of the proposed methodology is represented by the fact that within the sentiment analysis phase, grammatical special cases such as negations or degree words were not considered. Thus, in order to tackle the quality issues of the data analysis, it is imperative to consider this aspect as a first step in the improvement of the methodology.

Furthermore, as revealed in the previous chapter, the main issues, when working with textual-based social media data, are related to language processing. The implications of the erroneous language processing on the extraction of affective responses from the metadata of geotagged photographs are twofold. Due to incorrect language detection, on one hand, the computation of the valence and arousal scores is not possible since the words are looked up in the wrong affective lexicon and on the other hand, the hashtag segmentation algorithm performs erroneously as well. The improvement of language detection is a challenging task since the multilingual nature of texts. The language of single words would need to be detected, process that generally would raise other validity issues because of the lack of context. To maximize the performance of the sentiment analysis, these issues need to be further studied.

Despite the abundance of noise inherent in social media data, trending hashtags often contain important information and provide insight into user's affective evaluation of objects, concepts or even environments and the affective states experienced. Hashtag segmentation is, consequently, an important first step in NLP processing, since it would be difficult to derive accurate meaning from a piece of text without initially determining the words that it consists of. A series of factors: typos, abbreviations and acronyms, social media or other domain specific vocabularies, online slang and the inconsistencies in user's choice of language or just the general occurrence of linguistic rule-breaking in hashtags, make the processing and segmentation of this particular type of data a very difficult task prone to produce errors.

Even if the results of the evaluation of the proposed approach to determine whether people refer to the environment in the provided metadata were satisfactory (57 out of 57 entries in the evaluated sample data were identified as containing at least one place name), users are generally not that straightforward as implied by this approach and just name the places they are referring to. Furthermore, to quantise individual affective ratings, the approach averaged the valence and arousal values of all the nouns, adjectives and adverbs in the metadata of one or multiple photos provided by one user for a specific location; however, not all the words might refer to the affective qualities of the place identified or the affective states the user experienced in that place and subsequently, the computed values of the affective dimensions could be inaccurate. On top of that, there is no certainty that the geolocation of the photo actually matched the place the user refers to. Taking into account all of the above, the approach needs revision and improvement by considering alternative ways of determining whether the users express their affective responses towards environments and by addressing the issue of location accuracy.

With regard to the aggregation of the extracted affective ratings, two possible limitations of the approach need to be addressed. Firstly, the proposed approach assumes that affective responses from a large number of users can be aggregated (averaged) to approximate the collective affective evaluation of the area surrounding one specific street. Notwithstanding, people's affective responses to environments are subjective in nature and abstract values; in consequence, it is unclear to what extent averaging abstract subjective data is suitable for the identification of collective affective responses and this aspect should be subject to further investigations. Secondly, to enable the implementation of the approach and following the idea of a perceptual spatial extent exposed in chapter 3, the affective responses located within a radius of 100m surrounding each street segment were considered as influential in the calculation of the valence and arousal values of the collective affective ratings. Nevertheless, this aspect needs further investigation since the choice is not based on empirical evidence and might lead to further data validity issues.

6.2 Future Work

In accordance with the series of weaknesses and limitation presented in the previous section of this chapter, some possible technical improvements along with some important directions in future work are suggested as follows.

First of all, the identification and correct processing of grammatical special cases in the sentiment analysis is compulsory for the improvement of the proposed methodology. For this step, the methodology proposed and the data base compiled by Hauthal (2015), which considers six cases of grammatical special (based on simple negation words or negation conveying prefixes and the comparative and superlative forms of adjectives) for both English and German languages, could be regarded as the underlying theoretical framework for the extraction and processing of the grammatical special cases.

To improve the segmentation of hashtags, one possible solution could be represented by the compilation of more complex and complete, domain specific words collection that would cover the categories mentioned in the previous section such as slang, abbreviations, social media specific vocabulary etc. and the creation of a list of most popular hashtags based on a social media corpus. The idea of a list with the most popular hashtags would partially solve the issue of language misclassification; however, as mentioned before it is quite difficult to fully tackle this issue.

With regard to the positional accuracy of the extracted affective responses, according to the literature findings presented in the second chapter, Flickr geotagged photographs have in general a higher positional accuracy than the Instagram ones. Moreover, in contrast with the Instagram API, the Flickr API provides specific information regarding the level of positional accuracy; thus, the nature of collected Flickr data allows a more precise (contestably) filtering of the data before performing any analysis on it. Nevertheless, as a further direction for future work, better strategies of assessing the positional accuracy of geotagged photographs need to be explored or developed.

On a further note, to enhance the performance of the algorithm that detects relevant affective information expressed towards the environment, as a closely-related improvement, the gazetteer compiled needs to be expanded by including further words that generally express the concept of place, environment or space based on domain specific corpus data. Another possible approach could be represented by employment of topic modelling tools; nevertheless, this proposed solution would also require domain specific annotated corpus data or more complex and sophisticated deep learning algorithms, that would require expertise in the realms of linguistics and computer science. Developing such algorithms would also enable the extraction of data regarding the context in which the users respond affectively to the environments, data that could be considered for the improvement of the affective responses' aggregation and modelling approaches.

The conducted evaluation of the implemented methodology is rather superficial and further research should be done to explore more realistic and detailed evaluation in more diverse environments. As a way of validating the methodology proposed, a pedestrian navigation application which implements a routing algorithm that incorporates the aggregated affective responses should be developed. The top-K algorithm used by Huang et al. (2014) in their approach to compute the "affect routes", namely Yen's algorithm (Yen, 1971), could be adapted and employed in this case as well. The computer algorithm, which is formulated based on weighted graphs works in the following manner: from a given starting node of a graph aims to find the best top-k paths applying a cost function based on the weighted attributes of the edges (e.g. length, duration, etc.). In this manner, the affective classes in which the graph segments fall, could be considered as weighted attributes. With the aid of such developed application, an empirical evaluation of the methodology could be conducted and would help us to better quantify the performance and limitations of the methodology.

7 CONCLUSIONS

The aim of this chapter is to present the conclusions of the work carried out by answering the underlying research questions and putting in a nutshell the findings of this research.

The conducted research sought to develop a methodology derived from the ideas and methods proposed by the "AffectRoute" study (Huang et al., 2014) and consider location-based social media data as an additional source for the extraction of the affective responses towards the environment. Furthermore, another goal of this research was to find a suitable way of aggregating the extracted data and deliver an affective-layer that could be further incorporated into a routing algorithm. In the conceptualisation phase of the methodology, three research questions, that needed to be answered in order to deliver the methodology, were identified. The questions and their answers are presented as follows.

The first research question posed at the beginning of this thesis was: *How can the affective responses towards environments be described in a structured way and how can they be extracted from location-based social media data?* According to the findings in literature, in the context of location-based social media data, the affective response towards environments could be described as the linguistics reactions (e.g. lexical and syntactical constructions, word choices) identified in the metadata of geotagged photos, that people use to express specific emotions they experience in a particular environment or to name certain affective qualities of environments that evoked specific emotions in them. To analyse these geotagged emotional expressions and to extract the affective responses from location-based social media data, the affective connotation of the word choices could be determined with the aid of the lexicon-based sentiment analysis method. The affective connotations of words are quantified in terms of the two dimensions of the core affect (valence and arousal) based on the dimensional approach to measure emotional states proposed by Russell (1980). In short, affective responses are georeferenced valence and arousal values associated with certain emotions or affective qualities.

With regard to the second question: *What are the main characteristics of text-based social media data and how should these be considered when extracting affective responses?*, the conducted research offers solid theoretical background of the nature of text-based social media data. To put it in a nutshell, textual-based social media data is mostly noisy and non-standard in nature, often being deviated from standard vocabulary in terms of syntax, or even in terms of semantic aspects. Thus, the employment of the following NLP procedures: text-normalisation (for the removal of noisy data such as non-alphabetic characters and for spell correction), language detection (for the segregation of text by language and removal of the irrelevant text) and hashtag segmentation (for the extraction of the vast amounts of data encoded as this social-media specific unique tagging format) is crucial before applying any sentiment analysis method meant to extract the affective responses.

Finally, the answer to the third research question: *The extracted affective responses will be initially stored as discrete data. How should be this data aggregated into an affective layer suitable for a further integration with the navigation graph used by the routing algorithm?* is based on the idea underlying the approach conceptualised to aggregate the extracted discrete affective responses. The approach proposed the idea that individual discrete affective responses can be aggregated into collective affective responses in order to assign them as weighting attributes to the edges of a routing graph. As an aggregation method, the computation of an inverse distance weighted arithmetic mean value for the valence and arousal of all the affective responses relevantly located for each street segment was proposed. Furthermore, to rank each of the street segments contained by the routing graph, five possible affective classes were conceived and they were, subsequently, validated by the results of the evaluation conducted.

The results of the implemented methodology showed that social media users do rate environments affectively in the content shared and on account of the vast amount of affective responses extracted. Accordingly, it can be asserted that location-based social media data is an important source of data that needs to be considered in order to provide human-centred pedestrian navigation applications. However, in view of the lack of empirical evaluation of the results of this research and the numerous quality issues that location-based social media data raises in general, it can be said that the implementation of a routing algorithm based solely on this type of data should be avoided and unless future empirical findings would suggest otherwise, it should be considered as an additional source of data to fill in the gaps in the traditionally acquired or crowdsourced datasets.

To conclude this thesis, it can be said that the conducted research offers a fairly solid theoretical and methodological framework for the extraction of the affective responses towards the environments from location-based social media data along with an innovative approach for the aggregation and modelling of the extracted data for pedestrian navigation purposes.

References

AGRAWAL A. (2018): *Enriching Affect Analysis Through Emotion and Sarcasm Detection.* Ph.D. thesis, York University, Toronto.

ANTONIOU V., MORLEY J. AND HAKLAY M. (2010): *Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon*. Geomatica, 64(1), pp. 99-110. Ph.D. thesis, York University, Toronto.

APA Dictionary of Psychology (2018) [online] https://dictionary.apa.org/ (accessed 24.12.2018)

BACCIANELLA S., ESULI A. AND SEBASTIANI F. (2010): *SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining.* In: *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010*, Valletta, Malta, 17-23 May 2010.

BAO S., NITTA T., YANAGISAWA M. AND TOGAWA N. (2017): *A safe and comprehensive route finding algorithm for pedestrians based on lighting and landmark conditions.* IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E100A (11), pp. 2439-2450.

BARRETT L.F., LINDQUIST K. AND GENDRON M. (2007): *Language as a context for emotion perception*. Trends in Cognitive Sciences, 11, pp. 327-332.

BIRD S., KLEIN E. AND LOPER E. (2009): *Natural Language Processing with Python Analyzing Text with the Natural Language Toolkit 1st*. O'Reilly Media. Available from: https://www.nltk.org/book/(accessed on 31.01. 2019)

BOEING G. (2017): *OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks*. Computers Environment and Urban Systems 65, pp. 126-139. Available from: https://www.researchgate.net/publication/309738462_OSMnx_New_Methods_for_Acquiring_Constructing_Analyzing_and_Visualizing_Complex_Street_Networks/citations (accessed on 24.12.2018)

BONDI L., DAVIDSON J. AND SMITH M. (2005): *Introduction: Geography's 'Emotional Turn'.* In: Bondi L., Davidson J. and Smith M. (Ed.), *Emotional Geographies,* pp. 1-16, Aldershot, Ashgate.

BORST H.C., DE VRIES S.I., GRAHAM J.M.A., VAN DONGEN J.E.F., BAKKER I. AND MIEDEMA H.M.E. (2009): *Influence of environmental street characteristics on walking route choice of elderly people.* Journal of Environmental Psychology, 29, pp. 477–484.

BRADLEY M.M. AND LANG P.J. (2010): *Affective norms for English Words (ANEW): Stimuli, instruction manual and affective ratings*. Technical Report C-2. The Center for Research in Psychophysiology. University of Florida, Gainesville, Florida.

CAMBRIA E., SCHULLER B., XIA Y. AND HAVASI C. (2013): *New avenues in opinion mining and sentiment analysis.* IEEE Intelligent Systems. 28 (2), pp. 15–21.

CORONA B., WINTER S. (2001): *Guidance of car drivers and pedestrians.* Technical report, Institute for Geoinformation and Cartography, TU Vienna.

CRAGLIA M., OSTERMANN F. AND SPINSANTI L. (2012): *Digital Earth from vision to practice: making sense of citizen-generated content.* International Journal of Digital Earth, 5 (5), pp. 398–416.

CVETOJEVIC S., JUHASZ L. AND HOCHMAIR H.H. (2016): *Positional Accuracy of Twitter and Instagram Images in Urban Environments.* GI Forum 1, pp. 191-203. Available at: https://www.researchgate.net/publication/305272606_Positional_Accuracy_of_Twitter_and_Instagram_Images_in_Urban_Environments (accessed 16.01.2019).

DOWNS R. AND STEA D. (1973): *Image and environment: Cognitive mapping and spatial behaviour*. Aldine Pub. Co., Chicago

DUCKHAM M. AND KULIK L. (2003): *"Simplest" paths: Automated route selection for navigation.* In: Kuhn W., Worboys M. and Timp S. (Ed.), *Spatial Information Theory. Foundations of Geographic Information Science*. *COSIT 2003. Lecture Notes in Computer Science,* vol. 2825, pp. 169–185, Springer, Berlin, Heidelberg.

DAAMEN W. AND HOOGENDOORN S.P. (2003): *Research on Pedestrian Traffic Flows in the Netherlands*. In: Proceedings Walk 21 IV, 1-3 May 2003, Portland, Oregon, pp. 101-117.

EKMAN P. AND FRIESEN W.V. (1971): *Constants Across Cultures In The Face And Emotion.* Journal of Personality and Social Psychology, 17(2), pp. 124-129.

GARTNER G., HUANG H., MILLONIG A., SCHMIDT M. AND ORTAG F. (2011): *Human-centred Mobile Pedestrian Navigation Systems.* Mitteilungen der Österreichischen Geographischen Gesellschaft, 153, pp. 237 – 250.

GOLDER S.A. AND HUBERMAN B.A. (2006): *Usage patterns of collaborative tagging systems.* Journal of Information Science 32 (2), pp. 198-208.

GOLLEDGE R.G. (1995): *Path Selection and Route Preference in Human Navigation: A Progress Report.* In: Frank A.U. and Kuhn W. (Ed.), *Spatial Information Theory – A Theoretical Basis for GIS*, pp. 207–222, Springer, Berlin, Heidelberg.

GOODCHILD M.F. (2007): *Citizens as sensors: the world of volunteered geography.* Geojournal, pp. 1-12.

HAUTHAL E. (2015): *Detection, Modelling and Visualisation of Georeferenced Emotions from User-Generated Content.* Ph.D. thesis. Technische Universität Dresden.

HAUTHAL E. AND BURGHARDT D. (2016): Mapping Space-Related Emotions Out of User-Generated Photo Metadata considering Grammatical Issues. The Cartographic Journal 53(1), pp. 78-90.

HAQUE S., KULIK L. AND KLIPPEL A. (2007): *Algorithms for reliable navigation and wayfinding.* In: Barkowsky T., Knauff M., Ligozat G. and Montello D.R. (Ed.), *Spatial Cognition V Reasoning, Action, Interaction*, vol 4387, pp. 308-326, Springer, Berlin, Heidelberg.

HOLLENSTEIN L. AND PURVES R. (2010): *Exploring Place through User-Generated Content: Using Flickr to Describe City Cores*. Journal of Spatial Information Science 1(1), pp. 21-48

HUANG H. AND GARTNER G. (2009): *Using Activity Theory to identify relevant context parameters.* In: Gartner G. and Rehrl K. (Ed.), *Location Based Services and TeleCartography II - from Sensor Fusion to Con-text Models,* pp. 35-45, Springer, Heidelberg, Berlin.

HUANG H., KLETTNER S., SCHMIDT M., GARTNER G., LEITINGER S., WAGNER A. AND STEINMANN R. (2014): *AffectRoute – considering people's affective responses to environments for enhancing route-planning services*. International Journal of Geographical Information Science, 28 (12), pp. 2456-2473.

HUANG H., GARTNER G. AND TURDEAN T. (2013): *Social media data as a source for studying people's perception and knowledge of environments*. MÖGG Mitteilungen der Österreichischen Geographischen Gesellschaft (155), pp. 291-302.

HUANG H. (2016): *Context-Aware Recommendations Using Geotagged Photos in Social Media.* ISPRS International Journal of Geo-Information 5 (11): 195. Available at: https://www.mdpi.com/2220-9964/5/11/195 (accessed 29. 12. 2018).

JANKOWSKI P., ANDRIENKO N., ANDRIENKO G. AND KISILEVICH S. (2010): *Discovering landmark preferences and movement patterns from photo postings.* Trans. GIS 14, pp. 833–852.

KAISER S. AND WEHRLE T. (1992): *Automated coding of facial behavior in human-computer interactions with FACS*. Journal of Nonverbal Behavior, 16, pp. 67-83.

KAPLAN R. AND KAPLAN S. (1989): *The experience of nature: a psychological perspective*. Cambridge University Press, New York.

KIPP M. (2007): *@toread and Cool: Tagging for Time, Task and Emotion.* In: Proceedings of Information Architecture Summit 2007: Las Vegas, Nevada, 22-26 March 2007.

KISILEVICH S., MANSMANN F. AND KEIM D. (2010): *P-DBSCAN: A density based clustering algorithm for exploration and analysis of attractive areas using collections of geotagged photos.* In: *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application*, New York, 21–23 June 2010.

KLETTNER S., HUANG H., SCHMIDT M. AND GARTNER G. (2013a): *Acquisition and Cartographic Applications of Subjective Geodata*. In: Buchroithner M.F. (Ed.), *Proceedings of the 26th International Cartographic Conference: August 25 - 30, 2013, Dresden*.

KLETTNER S., HUANG H., SCHMIDT M. AND GARTNER G. (2013b): *Crowdsourcing Affective Responses to Space*. Kartographische Nachrichten 63 (2/3), pp. 66-72.

LINDQUIST K.A., GENDRON M. AND SATPUTE A. (2016): *Language and emotion: putting feelings into words and words into feelings.* In: Barrett L.F., Lewis M. and Haviland-Jones J. (Ed.), *Handbook of Emotions, Fourth Edition*, Guildford Press, New York.

MACKERRON G. AND MOURATO S. (2013): *Happiness is greater in natural environments*. Global Environmental Change 23 (5), pp. 992-1000.

MEHRABIAN A. AND RUSSELL J.A. (1974*): An approach to environmental psychology*. M.I.T. Press, Cambridge.

MILLONIG A. AND GARTNER G. (2009): *Ways of Walking – Developing a Pedestrian Typology for Personalised Mobile Information Systems*. In: Gartner G. and Rehrl K. (Ed.), *Location Based Services and TeleCartography II - from Sensor Fusion to Con-text Models,* pp. 79-94, Springer, Heidelberg, Berlin.

MIURA H., TAKESHIMA S., MATSUDA N. AND TAKI H. (2011): A study on navigation system for pedestrians based on street illuminations. In: Kurosu M. (Ed.), *Knowledge-Based and Intelligent Information and Engineering Systems; KES 2011; Lecture Notes in Computer Science,* Volume 6883, pp. 49–55, Springer, Heidelberg.

MONTELLO D.R. (2005): *Navigation.* In: Shah P. and Miyake A. (Ed.), *The Cambridge book of Visuospatial Thinking*, pp. 257-294, Cambridge University Press, New York. Available at: http://dx.doi.org/10.1017/CBO9780511610448.008 (accessed 27.12.2018)

NOWACK T., WANG Z. AND ZIPF A. (2018): *A system for Generating Customized Pleasant Pedestrian Routes.* Sensors 2018, 18(11). Available at: https://doi.org/10.3390/s18113794 (accessed 29.12.2018).

PIPPIG K., BURGHARDT D. AND PRECHTEL N. (2013): *Semantic similarity analysis of user-generated content for theme-based route planning*. Journal of Location Based Services, 7:4, pp. 223-245. Available from: https://doi.org/10.1080/17489725.2013.804214 (accessed on 23.11.2018).

REHRL K., HÄUSLER E., LEITINGER S. (2010): *Comparing the effectiveness of GPS-enhanced voice guidance for pedestrians with metric- and landmark-based instruction sets.* In: Fabrikant S.I., Reichenbacher T., van Kreveld M. And Schlieder C. (Ed.), *GIScience 2010*, pp. 189–203, Springer, Berlin, Heidelberg.

REHRL K., GRÖCHENIG S., HOCHMAIR H., LEITINGER S., STEINMANN R. AND WAGNER A. (2012): *A conceptual model for analyzing contribution patterns in the context of VGI.* LBS 2012 – 9th Symposium on Location Based Services, Springer, Berlin.

RUSSELL J.A. (1980): *A circumplex model of affect.* Journal of Personality and Social Psychology (39), pp. 1161-1178.

RUSSELL J.A. AND PRATT G. (1980): *A description of the affective quality attributed to environments*. Journal of Personality and Social Psychology, 38(2), pp. 311–322.

RUSSELL J.A. (2003): *Core affect and the psychological construction of emotion*. Psychological Review, 110 (1), pp. 145–172.

SEN S., LAM S.K., RASHID A.M., COSLEY D., FRANKOWSKI D., OSTERHOUSE J., HARPER F.M. AND RIEDL J. (2006): *tagging, communities, vocabulary, evolution*. In: Hinds P. and Martin D. (Ed.), CSCW '06: *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work: Banff, Alberta, Canada*, *November 04 - 08, 2006*, pp. 181-190, ACM Press, New York.

ULRICH R.S. (1983): *Aesthetic and affective response to natural environment.* In: Altman I. and Wohlwill J. (Ed.*), Human Behavior and Environment*, vol. 1.6, pp. 85-125, Plenum, New York.

VALLI C. AND HANNAY P. (2010): *Geotagging Where Cyberspace Comes to Your Places.* In: International Conference on Security and Management (SAM '12) Las Vegas, 16-19 July 2012.

VÕ M.L.-H., CONRAD M., KUCHINKE L., HARTFELD K., HOFMANN M.F. AND JACOBS A.M. (2009): *The Berlin Affective Word List Reloaded (BAWL-R).* Behavior Research Methods 41 (2), pp. 534-538.

WARRINER A., KUPERMAN V. AND BRYSBAERT M. (2013): *Norms of valence, arousal, and dominance for 13,915 English lemmas*. Behaviour Research Method, 45(4), pp. 1191-1207.

WINTER S. (2002): *Modelling costs of turns in route planning*. GeoInformatica, 6(4), pp. 345–361.

YEN J.Y. (1971): *Finding the k shortest loopless paths in a network.* Management Science 17 (11), pp. 712-716. Available from: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.17.11.712 (accessed on 2.02.2019)

YIK M., RUSSELL J.A. AND STEIGER J.H. (2011): *A 12-point circumplex structure of core affect.* Emotion (11), pp. 705-731.

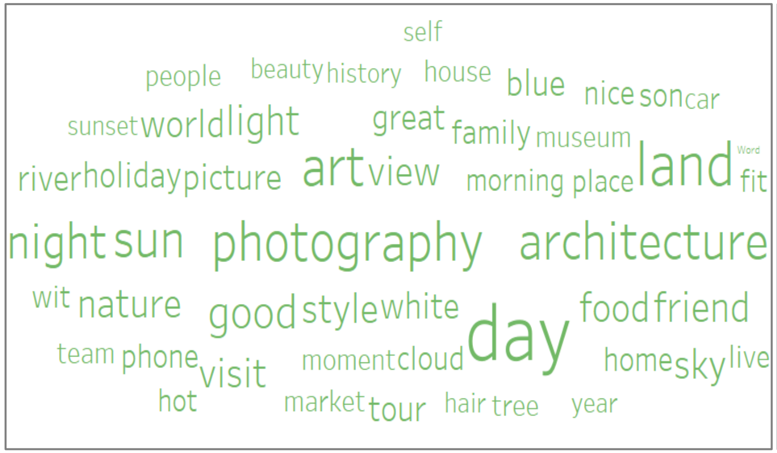
ZIELSTRA D. AND HOCHMAIR H. (2011): *Comparative Study of Pedestrian Accessibility to Transit Stations Using Free and Proprietary Network Data*. Transportation Research Record Journal of the Transportation Research Board 2217(-1). Available at: https://journals.sagepub.com/doi/10.3141/2217-18 (accessed 24.11.2018).

APPENDICES

Appendix 1 Top 50 Most Frequent Affective Words (English and German) Representative for each of the Five Affec-tive Classes VIII

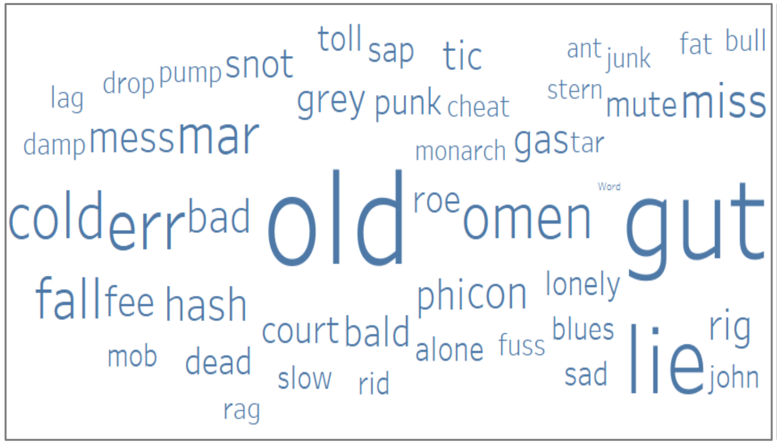
Appendix 2 Classification of the Street Segments Based on the Aggregated Affective Responses.. XI

Appendix 1 Top 50 Most Frequent Affective Words (English and German) Representative for each of the Five Affective Classes



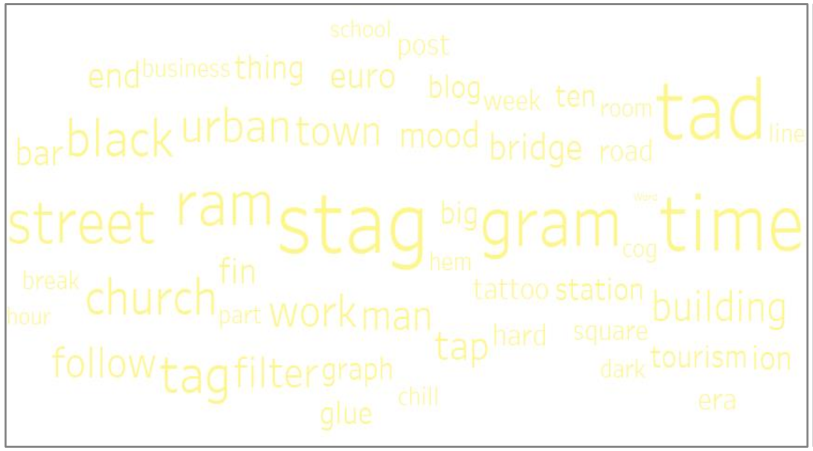
Class 1: Top 50 English Words

Class 1: Top 50 German Words



Class 2: Top 50 English Words

Class 2: Top 50 German Words

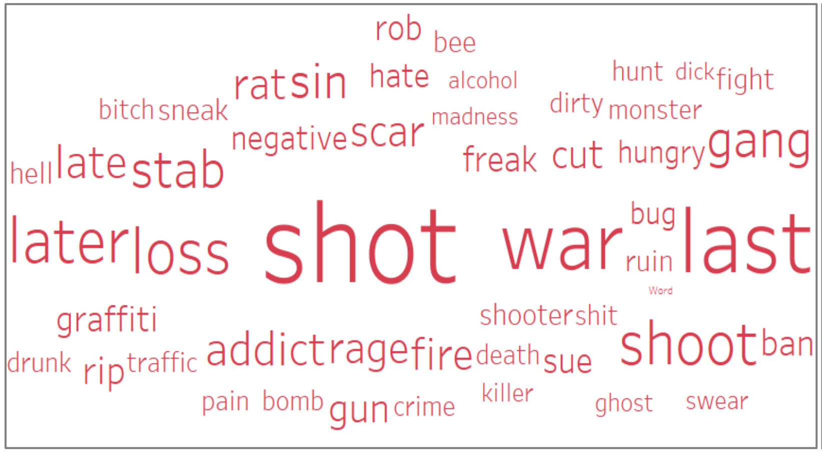


Class 3: Top 50 English Words

Class 3: Top 50 German Words

Class 4: Top 50 English Words

Class 4: Top 50 German Words



Class 4: Top 50 English Words

Class 4: Top 50 German Words

Appendix 2 Classification of the Street Segments Based on the Aggregated Affective Responses