University of Zurich[UZH]

Extracting cultural landscape properties of UNESCO cultural heritage sites in Switzerland

GEO 511 Master's Thesis

Author
Fabienne Christen
13-933-064

Supervised by
Prof. Dr. Ross Purves

Faculty representative
Prof. Dr. Ross Purves

16.06.2021
Department of Geography, University of Zurich

Abstract

The constant interaction between human and nature forms cultural landscapes. Thereby, the perception of landscapes plays an important part on how a landscape is valued and treated. The objectives of this study were to extract characteristic landscape properties in the regions considered to be UNESCO cultural heritage property in Switzerland: the Vineyards Terraces in Lavaux and the Rhaetian Railway in the Albula / Bernina Landscapes (RhB) (RQ1). Different to past studies, this study includes not only one data source and method of extraction, but uses a combined approach of text and image analysis to get a more diverse perception of landscape. Thereby the differences of the two methods were examined seeking a better understanding about the outcome based on the chosen method (RQ2). The resulting properties were compared to the criteria of the UNESCO world heritage list (RQ3).

Text data was gathered from E-Periodica, processed and analysed using concordance analysis. For the image analysis, geo-referenced Flickr data was used and analysed with the help of the machine learning algorithm Google Vision. The resulting words and tags were annotated into differrent categories. While the image analysis revealed a high amount of biophysical properties, properties regarding the history and identity of a region were better captured using text analysis. By combining the two approaches, different aspects of the landscapes could be captured. However, for future research the methodologies in text and image analysis should be further improved. In text analysis this includes the use of context to prevent word ambiguity whereas in image analysis the training data of the machine learning algorithm should be taken into consideration.

Keywords: *landscape perception, user generated content, natural language processing, geographic information retrieval, visual recognition, volunteered geographic information, Flickr, E-Periodica*

Acknowledgements

By submitting my Master's Thesis, my studying years at the University of Zurich come to an end. I would like to thank my friends and family for their support during this time and all the people who accompanied me along the way. I would like to give a special thanks to the following people:

Prof. Dr. Ross Purves Thank you for your support and assistance during my thesis. I highly value all your inputs and questions which helped me during my work.

Fabian Biland Thank you for interesting discussions, carrying different perspectives and for taking time to annotate labels and words.

Patricia Schnidrig and Eric Christen Thank you for proof-reading my thesis and giving inputs from a different point of view.

Contents

List of Figures

List of Tables

List of Listings

Chapter 1

Introduction

Cultural landscapes are formed in the interaction of human and nature and profit humans in various ways. An example for such a benefit are cultural ecosystem services (CES) which describe amongst other things recreational or spiritual values that increase the health and well-being of people (Chan et al., 2012). Different organizations aim to protect cultural landscapes. The Millenium Ecosystem Assessment (Millenium Ecosystem Assessment, 2005) emphasizes on the health and well-being benefits for humans which result from nature and tries to quantify values of CES. Concentrating on cultural landscapes, the United Nations Educational, Scientific and Cultural Organization (UNESCO) wants to sustain cultural and natural heritage. For that purpose a representative, balanced and credible world heritage list was created and constantly updated (UNESCO World Heritage Centre, 2008a).

As the value of CES increased in the past, many studies are aiming for a better understanding of these landscapes (Cooper and Gregory, 2019; Gliozzo et al., 2016; Van Zanten et al., 2016). However, landscapes are complex constructs, which combine many different aspects. It is therefore a challenging task to extract properties of a landscape which combine the landscape's entire perception. In past studies one type of data was included leading to an incomplete image of the landscape.

However, with the rising of new technologies, different types of data are more easily accessible widening the variety of data that can be analysed (Miller and Goodchild, 2015). Combined with automated approaches of machine learning algorithm, it is possible to include a great amount of data. Thereby, not the perception of a selected amount of people, but of much more people perceiving a landscape can be included. In this thesis the focus will lie on the combination of text and image data, aiming for a profound caption of all landscape properties. As areas of interest, the two UNESCO cultural heritage sites in Switzerland are analysed: the vineyard terraces of

Lavaux and the Rhaetian Railway in the Albula / Bernina Landscapes (RhB).

Based on the example of the RhB and Lavaux, landscape properties will be extracted using a combined approach of text analysis and image processing. The goal is to gain a deeper understanding in the perception of landscapes by extracting and comparing landscape properties of the two sites. Especially, characteristic properties of UNESCO cultural heritage sites will be elaborated such that areas of great cultural value can be more easily identified and protected.

Chapter 2

Background

This chapter clarifies definitions and concepts. In a first section, the importance of landscapes and their values is shown. Cultural ecosystem services and the UNESCO are introduced including their principles and goals. In a second section the term *landscape* is explained, showing different definitions of the term as well as differences on its perception. Thereby it is elaborated how the literature influences the value of a landscape. Furthermore, natural language as data source is discussed revealing differences between unstructured and structured text and discussing common steps in working with natural language. In a final section, the value and possibilities of user generated content is discussed.

2.1 Value of landscapes

The relation between human and landscapes is of interest in many research fields due to topics such as ecology or land change. Because of the Millenium Ecosystem Assessment (2005) the concept of ecosystem services became the dominant focus in research (Bieling, 2014). Thereby three types of services are differentiated: provisioning services, regulating services and cultural services. Provisioning services define regions, from which products, e.g. food, can be derived from the ecosystem. Regulating services include ecosystems which have a regulating function such as climate regulation. The final type, cultural ecosystem services, describes regions from which benefits for human derive, e.g. recreation (Millenium Ecosystem Assessment, 2005).

2.1.1 Cultural ecosystem services

The importance of landscapes rose with the adoption of the European Landscape Convention (Council of Europe, 2000). The convention is the first official attempt to qualify landscape properties for landscape planning (Butler, 2016). Although all aspects of landscapes should be captured, Butler (2016) criticizes that the landscape is often only considered as surface and not perceived in its entirety. The concept of cultural ecosystem services (CES) includes the interaction of human and nature and results as collection of human perception (Gliozzo et al., 2016). Thereby, all aspects contributing to the human health or well-being such as recreational or spiritual values are considered (Chan et al., 2012). A further characteristic of CES is the non consumptive use of natural resources (Millenium Ecosystem Assessment, 2005). However, the assessment of non-material benefits is challenging which is why it is often neglected in studies despite its need for capturing all aspects of a landscape (Schaich et al., 2010).

2.1.2 UNESCO

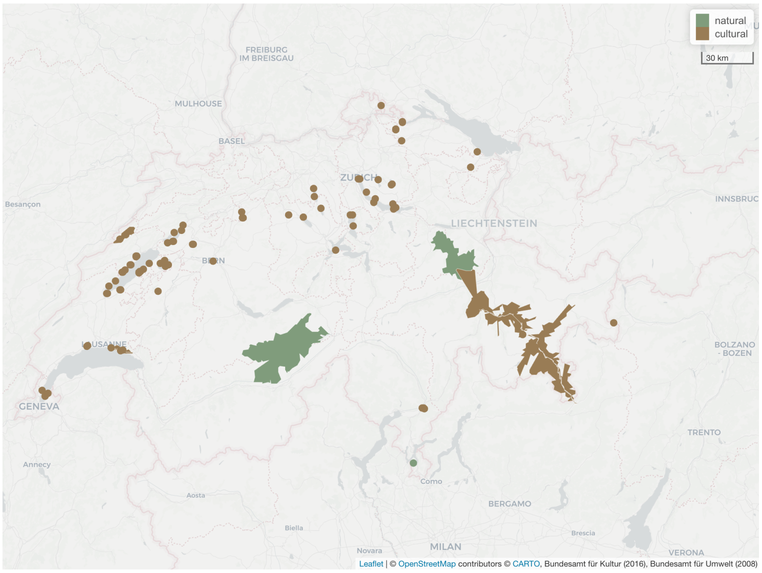

Figure 2.1: UNESCO world heritage properties in Switzerland, divided into natural and cultural world heritage. The data is provided by Federal Office for the Environment FOEN (2008) and Federal Office of Culture FOC (2016)

.

The UNESCO wants to sustain the culture and nature heritage. For that purpose a representative, balanced and credible world heritage list was created and is constantly updated (UNESCO World Heritage Centre, 2008a) . In this thesis the focus lies on cultural landscapes by the definition of the UNESCO. Cultural landscapes combine both human works and nature and should be representative for the development of human society with the influence of constraint posed by nature's influence (Article 47, UNESCO World Heritage Centre (2008a)). According to the UNESCO, there are three different categories of cultural heritage: monuments, groups of buildings and sites. All of these categories include a technical aspect which show the progress of humanity. However, the list should be diverse. Consequently, objects similar to already existing world heritage sights, will not be taken into account as possible candidates. The following criteria were defined to be able to qualify as UNESCO cultural heritage property (UNESCO World Heritage Centre, 2008a):

i) The property needs to be designed and created intentionally by humanity

ii) The common works between nature and humanity needs to be displayed

iii) It should be an outstanding example of human progress

iv) Authenticity and integrity of the property needs to be guaranteed

v) The list should be balanced, containing diverse properties

In Switzerland, a total of twelve sites are included in the World Heritage list of UNESCO (see figure 2.1). Out of these, only two classify as cultural landscapes: the vineyard terraces of Lavaux and the Rhaetian Railway in the Albula / Bernina Landscapes (RhB). As this thesis focuses on the UNESCO world cultural heritage, the two areas are shortly described.

Vineyard terraces in Lavaux

The vineyard terraces of Lavaux are situated in the canton of Vaud and date back to the 11th century. They were included to the UNESCO list in 2007 due to the exemplary interaction between humanity and its environment which helped the local economy (UNESCO World Heritage Centre, 2007). The vineyard terraces are next to the lake of Geneva and the world heritage property has a size of 890 ha. The surrounding buffer zone increases the region by 1408 ha. The region of Lavaux includes residential areas and has approximately 722 inhabitants per square kilometer (Federal Statistical Office FSO, 2019).

Figure 2.2: Flickr image taken of the vineyard terraces in Lavaux. Image taken by 'flederma' (Flickr user 54305039@N03), licensed under Creative Common (https://live.staticflickr.com/757/31207119350_8f92ecd425_c.jpg [15.05.21]

Rhaetian Railway in the Albula / Bernina Landscape (RhB)
The RhB was included one year later in 2008 as an outstanding example of technical progress. Due to the opening of RhB in 1904 isolated regions became more easily accessible which helped the human development (UNESCO World Heritage Centre, 2008b). The world heritage property covers the area of 152 ha, however the surrounding buffer zone is much bigger, covering an area of 109’394 ha. The region in which the RhB is situated is in the middle of the Alps with only little villages, the average population density is only about 41 inhabitants per square kilometer (Federal Statistical Office FSO, 2019).

Figure 2.3: Flickr image taken in the region of the RhB. Image taken by ’thomas alan’ (Flickr user 9021833@N06), licensed under Creative Common (https://live.staticflickr.com/3895/18880762485_ffdfd6ed2b_c.jpg [15.05.21])

2.2 Defining landscapes

To explore landscapes further, the term landscape needs to be clarified and defined. Furthermore, different perceptions of landscape will be elaborated and the concept of place writing is introduced.

2.2.1 The concept of landscape

Landscape as a concept includes different aspects which need to be considered when trying to capture its characteristics. A landscape can either be treated as one continuous field, which needs to be characterized as a whole or it can be looked at as various small divisions which together represent the landscape (Mark et al., 2011). Furthermore, the importance of human perception needs to be discussed. The characterization of a landscape does not only include objective properties but also the individual perception of space (Warnock and Griffiths, 2015). Johnson and Hunn (2010) studied landscapes as set of features with the help of generic terms. As the landscape is divided into individual features, the content of language as well as the perception of the place are emphasized in the analysis. Considering landscape as a continuous field on the other hand results in more objective properties (Bunce et al., 1996; Gerçek et al., 2011).

To capture all aspects of landscapes, both objectives as well as perceptive features need to be included in the analysis. To understand the entire extent of place, Agnew and Livingstone (2011) divided place into three different aspects: location, locale and sense of place. Location includes the exact position of place such as coordinates or the name of a certain place. Objective features such as geological or physical features as well as landforms describe the locale of place. Furthermore, locale includes cultural and perceptual features such as infrastructure or sounds. Finally, sense of place describes personal attachment to the respective place or landscape, which could be expressed through memories or feelings.

2.2.2 Landscape perception and place writing

The importance of human perception of places is underlined by Smith (2013), as he sees places as creations made by humans due to their environmental conditions. As places are creations of people's perceptions, the value of places can differ and be influenced. Cooper (2019) states that places get their meaning often through literature. He introduces the concept of place writing which is characterized through a language which leads to the enchantment of a certain landscape. Furthermore, the perception of landscapes is not a single sensation but multi-sensory as it includes olfactory, auditory and visual senses (Mark et al., 2011). Therefore, the inclusion of people's perception is necessary when quantifying landscapes. Although the perception of landscapes is multi-sensory, the visual perception often dominates when perceiving landscapes (Koblet and Purves, 2020a). Partly, this can be explained with the help of the etymology of the word. Secondly, senses are not evenly distributed in language. In conversations, vision is the sense which is discussed most often across all languages (San Roque et al., 2015).

2.2.3 Extracting Landscape properties

The European landscape convention has the aim to consider and capture all aspects of landscapes (Council of Europe, 2000). Especially, the diversity of landscapes should be captured. Therefore, not only scientists and specialists should be included in the discussion, but the public's perception of landscape should be considered as well (Jones and Stenseke, 2011). However, by including the perception of numerous people, new challenges rise. By doing qualitative interviews for example, only a limited number of people can be reached. Furthermore, this methods covers a small geographic width, as most often local people are questioned (Plieninger et al., 2013; Bieling et al., 2014). To get perceptions of a larger amount of people, surveys or participatory GIS (PPGIS) are used. However, external visitors who look at the landscape from outside, are often not included in such methods limiting the diversity of the perceived landscapes (Butler, 2016). To include more people, user generated content can be used as data source (Van Zanten et al., 2016; Gliozzo et al., 2016).

Depending on the methodology and the data set chosen, different landscape properties can be extracted as shown by Wartmann et al. (2018). The data source influences how well the three different aspects of place - location, locale and sense of place - can be identified and to which extent a category is represented in the landscape. In tags often properties of the categories biophysical, cultural or toponym can be found, whereas in free listing biophysical properties dominate and toponyms are hardly present. When considering text data, properties related to the sense of place are captured well. Differences in the extracted properties can be explained with the data structure. Text data carries more information and context to a place whereas tags describe a single feature. The content of information is more superficial when using tags leading to a loss of detail. Jolicoeur et al. (1984) defined three different levels for the classification of tags: *subordinate level*, *basic level* and *superordinate level*. The *basic level* represents the level on which people often identify an object in a first instance. The bottom level on the other hand, *the subordinate level*, shows more detail to a property allowing to be more specific about an object. On the other end, the *superordinate level*, tags which show a higher level on generality are placed. An example would be "larch" on a *subordinate level*, "tree" as the respective basic level *basic level* and "vegetation" on the *superordinate level*.

Wartmann et al. (2018) emphasize the importance of the data source, as landscape descriptions are more similar when using the same type of data source independent of the place the landscape is situated.

It was shown, that depending on the data source some properties are more easily extracted than others. In conclusion, it is essential to combine different data sources to capture all aspects of a landscape.

2.3 Natural language

Through the use of text data, natural language is included in studies. In this section, a short introduction in characteristics of the natural language is given followed by the use of text data as data source.

The content and extent of language depends on its use. For example, differences in the use of language can be found depending on the climate of the region where a language is spoken. Regier et al. (2016) showed that the frequency of using concepts of snow and ice is much higher in colder than in warmer regions and also the differentiation between these two concepts is more likely. As a consequence, the language chosen influences the result of a study due to its natural extent.

Furthermore, as shown by Warriner and Kuperman (2015) positive words dominate in the English language. Not only are they used more frequently than negative words, but also the amount of words which are correlated with a positive sensation is higher than negative words. As a result, a positive bias exists in the natural language caused by the extent of the language and the distribution of the words within (Dodds et al., 2015).

When automatically retrieving the meaning of a word, two aspects need to be considered: the input representation and the semantic representation (Gaskell and Marslen-Wilson, 1997). The input representation defines the structure of the word, which includes orthography or phonology. The semantic representation on the other hand defines the meaning of the word in the context. Often models with the aim of word recognition reduce the semantic representation to one meaning. However, words are not always well-defined with a unique meaning (Rodd et al., 2004). Ambiguity describes the property of words which are written the same but have a different meaning (Britton, 1978). The word *Zug* describes a train in German but is also a city as well as a canton in Switzerland. The task of finding the correct semantic meaning based on its context is called word sense disambiguation. Humans often capture the correct semantic meaning of a word by reading the entire text. A computer on the other hand needs to process the unstructured text into structures, which makes the task of word sense disambiguation difficult (Navigli, 2009).

2.3.1 Text corpus

Text data has been used as data source for a long time. However, the rise of new technologies opened up new possibilities of data gathering and increased the amount of data available, widening the area of research.

As an example, a great number of archives and libraries are digitising their inventory which leads to an increase in available data online (Hu, 2018).

To analyse large amount of text data with computational techniques, the data is gathered into a text corpus (Gregory and Hardie, 2011). Thereby, the nature of the language within this corpus is investigated. In the 1960s one of the first corpora was created which was meant to represent present-day American English (Francis and Kucera, 1964). Since then many text corpora have been built to cover specific topics, languages or time periods. Examples are the corpus Text + Berg by the Swiss Alpine Club (Text + Berg, 2021) or the Reuters corpus (NIST, 2004).

Further progress in publicly available data sources was made due to the growth of Web 2.0. Web 2.0 describes web application that enable the sharing and collaboration (Stefanidis et al., 2013). Famous examples of Web 2.0 platforms include Flickr, YouTube or Wikipedia. Thereby, advantages of Web 2.0 technologies such as asynchronous JavaScript and XML (AJAX) and the use of application programming interfaces (API) lead to further progress of Web 2.0 (Haklay et al., 2008). The use of AJAX enables users to gather information from a remote server without the need of refreshing the entire web page (Zucker, 2007). APIs are interfaces which are rather easily accessible making it attractive for a large amount of people to use for the creation and sharing of different types of data (Haklay et al., 2008). This can include the integrating of data from other sources but also gathering data. A common method to create a text corpus from data available online is web-scraping. However, when using web-scraping techniques legal and ethical issues should be clarified in advance (Zimmer, 2018).

2.3.2 E-Periodica

E-Periodica is a service provided by the library of the ETH Zürich, in which Swiss magazines are made available online and free of charge. The service's aim is it to enable digital access to magazines and originated with the idea to electronically archive the library. The oldest documents available date back to the 18th century while new articles are still being added. At first only scientific papers were collected, however nowadays articles covering various topics are available (E-Periodica, 2020). The use of articles originating from E-Periodica is free as long as it is not commercial and confined to private or scientific purposes.

2.3.3 Text data

By using spatial references, texts can be linked to geographic locations and used as data source in geographic research. In many texts a georeference is included Hu (2018). Georeferences could be geotags, which attach coordinates, but also mentioning of places in the text. By linking a text to an exact location, the text can be spatially placed and spatial relevance to an area of interest can be ensured.

Text data can be divided into two groups: unstructured and structured text data. Unstructured text data originates from web-scraping. Different from structured text data, unstructured text data is missing metadata or a clear data-model, making it hard to organise or index it (Hu, 2018).

Simple methods to computationally analyse unstructured text are the use of frequency counts or the search for concordance of a specific instance in the text (Gregory and Hardie, 2011).

Koblet and Purves (2020a) introduced several methods to analyse unstructured text. Bag-of-words does not consider the order of the words, but the frequency of each word and their co-occurrence with each other is of relevance. N-grams take a sequence of text into account, this could be sequences of two up to several tokens. For both these methods the use of the word's part of speech can be used as additional information, but is not necessary. For detecting patterns on the other hand the word's part of speech is important. Patterns are often based on occurrences of part of speech. An example of a pattern could be 'adjective + noun' (Kisilevich et al., 2010). A final method is looking for syntactic dependencies in which the lowest hierarchical levels for different part of speeches are examined.

To analyse the corpus further, the method of annotation can be used. Thereby, each element in the corpus is categorized (annotated) into the corresponding category. The categories are pre-defined (Garside et al., 2020).

2.3.4 Annotation

Corpus annotation adds additional value the text corpus, which bring several advantages (McEnery et al., 2006). The use of annotation enables people and machines to analyse and understand text data, although they might not completely master the respective language (Garside et al., 2020). Furthermore, the use of corpus annotation enables to use the text data for several studies which then can be compared easily (McEnery et al., 2006). However, although annotation simplifies research, it is important to still consider the entire text without annotation tags (Hunston, 2002). Furthermore, the use of annotation does not always represent the original data, but is simply an interpretation of it (McEnery et al., 2006). Even though annotation is a powerful method to analyse text data, the original text and especially the context of the data need to be considered.

2.4 User generated content (UGC)

Technical progress influenced various aspects in human life, including the importance of public information (Goodchild, 2007). The Organization of Economic Cooperation and Development (OECD) defined three main criteria for user generated content (UGC): the availability of the content in the internet for public access, the content shows creativity and the collection of the content was not done professionally (Vickery and Wunsch-Vincent, 2007).

In the context of geographic information systems, UGC includes gathering georeferenced data or recording the exact location of stuff with mobile devices (Turner, 2006). Methods like crowdsourcing or citizen science projects are examples of UGC (Bubalo et al., 2019). However, within UGC the distinction between volunteered geographic content (VGI) and contributed geographic content (CGI) needs to be explained. VGI defines data which is intentionally created and distributed. CGI on the other hand, includes data which is collected without the awareness or permission of the user (Goodchild, 2007; Elwood, 2008; Bubalo et al., 2019).

The development of UGC evolved due to the growth of Web 2.0. Enabling the sharing and collaboration of data, resulted in large amounts of UGC (Stefanidis et al., 2013). The rise of Web 2.0 and UGC includes collections of spatial data and geographic information. The public availability and easy access to UGC makes its use attractive for research (Bubalo et al., 2019; Van Zanten et al., 2016; Gliozzo et al., 2016).

2.4.1 Flickr

Flickr is a social networking and image sharing platform, on which georeferenced images with their respective metadata is found. The aim of Flickr is to offer an easy way to share images and enable the exchange with other people. Additionally, it offers various options of organizing even a large amount of image or video data. It was founded in 2004 and since then a large community evolved with about 100 million registered users up to date (Flickr, 2021). The large amount of data is made accessible among others with the Flickr application programming interface (API), which handles more than seven billion requests monthly (Flickr, 2021). The large amount of UGC made accessible on Flickr has been used in the past to extract landscape properties. In this study the use of Flickr data is meant to supplement the text analysis.

2.4.2 Copyright and protection of privacy

When using UGC, the copyright of the data and the protection of the user's privacy needs to be clarified. On the platform Flickr most users use a creative common license. The Creative Commons (CC) is a non-profit organization founded in 2001. Licences with creative common give the creator the opportunity to share their work with indication on how the data can be used (Thomee et al., 2016). Different levels of restrictions exist in creative common which limit the use of images to certain areas. All images used for the image analysis in this thesis are creative common and can be used for academic purposes.

Although UGC can be voluntarily or passive (VGI vs. CGI), most contributors are not aware of the digital footprint they leave (Mooney et al., 2017). When doing research with UGC the aim is to gather a great variety of data. Therefore, people are often encouraged to participate in VGI. Nevertheless, the privacy and ethics of participating people need to be guaranteed and the identity of the user should be protected (Mooney et al., 2017; Torra and Navarro-Arribas, 2014). In this study, the individual user is not of interest. The content of the image data is analysed independent of the user.

2.4.3 Machine learning

Using UGC as data source leads to a large amount of data which needs to be processed. To reduce the time needed to analyse a large amount of data, automated approaches and the use of artificial intelligence are used. The idea behind machine learning is the ability for a computer to learn rules based on training data and then to predict the respective outcome of a new data set (Wang et al., 2021).

In machine vision the focus lies on extracting important information from images. First, the algorithm searches for feature vectors which represent the image, afterwards the machine learning algorithm classifies the images (Abdullah et al., 2009) into categories. There are two types of classifying algorithms: supervised and unsupervised classification. In supervised classification, the data is compared to a training set and put into categories which were prior defined by a supervisor. When using unsupervised classification, the images are grouped according to their similarity without any pre-defined groups (Das, 2017). In supervised classification the data is usually divided into two parts: a training data set and a test data set. The training data is used to train the algorithm and to find matching values for the classifiers. The test data uses the trained model to get classification results and to check the validity of the model (Wang et al., 2021).

In this thesis the image analysis was carried out with the machine learning algorithm from Google Vision API. The Google Vision API is a commercial tool which supports understanding images with a machine learning algorithm. It processes the content of an image and extracts information by returning labels (Google Cloud, 2020). The Perficient Digital Agency (2019) showed in a comparison of several image recognition algorithms the high accuracy of Google Vision with approximately 81%. Combined with the low costs for the use of the algorithm, Google Vision is attractive for analysing large data sets and has been used in the development of applications or studies (Meiliana et al., 2017; Chen and Chen, 2017). Despite the advantages of Google Vision, the algorithm has its limitations. A great restriction of the algorithm is the missing information about the training data. This issue is further elaborated in chapter 6.5.

2.4.4 UGC in landscape research

Due to the easy access of UGC, the use of UGC has been used in numerous studies regarding landscape research and landscape perception (Bubalo et al., 2019; Van Zanten et al., 2016). In the case of image data, there are in general two common approaches that were considered. One method is to use the georeferenced image and analyse its tags for image analysis (Jeawak et al., 2017). In this case, the actual content of the image is not included, but only its metadata (Oteros-Rozas et al., 2018; Van Zanten et al., 2016). The second approach considers the content of the image for the extraction of landscape properties. In this case the image analysis is either done manually as it is done by Tenerelli et al. (2016) or Tieskens et al. (2018). Other researches analysed the images with the help of a machine learning algorithm (Seresinhe et al., 2017; Richards and Tunçer, 2018). The advantage of using a machine learning algorithm is the amount of data which can be processed. Seresinhe et al. (2017) analysed hundreds images to find landscape features which people consider as beautiful.

2.5 Research gap

The research field of extracting cultural ecosystem properties using different data sources covers many aspects. However, some research gaps could be discovered which will be introduced in this section.

First of all, many studies concentrate on cultural ecosystem services as area of interest. However, looking at UNESCO world heritage sites as area of interest was neglected in the past.

On a methodological level, past research mostly concentrated on one data source. Studies to extract landscape properties using participatory mapping techniques such as interviews or surveys have been done by Plieninger et al. (2013) and Pleasant et al. (2014) to include the opinion and knowledge of local people. Although participatory GIS is valuable and the data often of high quality, the data collection is time-consuming. Less time consuming than interviews is free listing which was done by (Bieling et al., 2014). Nevertheless, free listing should be done in the field with people experiencing the landscape which is why its reach is limited. Other research included people's opinion by using geo-tagged photographs collected by social media (Gliozzo et al., 2016; Van Zanten et al., 2016). However, tags are often limited in length which is why a lot of information about the perception of the landscape is lost. By including natural language, a deeper understanding in the values and properties of cultural landscapes could be reached (Van Zanten et al., 2016).

In addition to natural language, image processing shows promising results in extracting landscape properties. Both manual as well as automated approaches were made to analyse image content (Tenerelli et al., 2016; Oteros-Rozas et al., 2018).

As manual analysis of images is time consuming, automated approaches were made. Richards and Tunçer (2018) used an online machine learning algorithm which generates automated tags for each image.

Although the field of research is not new, the combination of several data sources has been neglected in the past. Although each data source shows promising results in capturing specific aspects of landscape, the combination of data sources to receive a more profound image of the landscape's perception has not been done.

By combining both text analysis with natural language and image processing, an improved method for extracting landscape properties is tested.

2.6 Research Questions

The following research question will be in focus of the research. *RQ1* focuses on the content which will be extracted using both text and image analysis. The aim is to find both similarities as well as differences between the two cultural heritage sights in Switzerland. *RQ2* aims to review the methodologies chosen and find differences between text and image analysis. The final objective *RQ3* wants to find the level of agreement between the found properties and the criteria to be on the UNESCO world heritage list.

RQ1. What distinct landscape properties can be extracted using text and image processing that qualify to be UNESCO world heritage?

  a) At the examples of Lavaux and RhB?

  b) What common properties can be extracted when comparing Lavaux to RhB?

RQ2. What differences in properties can be found when using text vs. image processing?

RQ3. To what extent match the extracted properties to the criteria of the UNESCO world heritage list?

Chapter 3

Data

The analysis includes two main parts: the text and the image analysis. In both parts data was collected from different sources which will be further elaborated in this chapter.

3.1 Data used for text analysis

The text data used in the analysis was gathered from the service E-Periodica by the ETH Zurich. The returned articles were locally saved and further processed as described in chapter 4.2.

3.1.1 E-Periodica

To query the large amount of documents stored on E-Periodica, a built-in search function is available. Due to the advanced search options, even detailed information about the data such as authors can be queried. The magazine articles are made available in image format and can be downloaded as pdf or jpg. Furthermore, is it possible to call each page of a journal with the help of an asynchronous JavaScript and XML (AJAX) call, which returns the page as JavaScript Object Notification (JSON). In each JSON response not only the text itself, but also metadata about the article are present. In listing 1 an example of a JSON response and its respective metadata is shown. Thereby the following attributes were of importance in the study:

- *journalTitle*: serves as unique identifier for each magazine, which was used to count the number of articles published per magazine
- *volumeNumYear*: volume and year of the article published. The year of the published article was used to remove all articles published after a region was added to the UNESCO list.
- *pid*: unique identifier for each page published in a magazine

```
{
    "movingWallInfoLink": "/digbib/jinfo?UID=adg-001",
    ↪ "available": true,
    "articles": [{"authors": "", "doi": "", "doiLink":
    ↪ null, "persons": "", "supplement": "", "title":
    ↪ "Kurzberichte", "chapterTitles": [], "pdfLink":
    ↪ "/digbib/../cntmng?pid=adg-001%3A2005%3A0%3A%3A179",
    ↪ "risLink": null, "journalTitle": null,
    ↪ "hasAuthors": false, "hasDoi": false,
    ↪ "hasPersons": false, "hasSupplement": false,
    ↪ "hasChapterTitle": false, "chapterTitle": null,
    ↪ "hasPdfLink": true, "hasRisLink": false}],
    "references": [],
    "collectionLink": "/digbib/browse5_3",
    ↪ "collectionTitle": "Geschichte und Geographie",
    ↪ "journalId": "adg-001",
    "journalLink": "/digbib/volumes?UID=adg-001",
    ↪ "journalTitle": "Jahresberichte des
    ↪ Archäologischen Dienstes Graubünden und der
    ↪ Denkmalpflege Graubünden",
    "journalEditor": "Archäologischer Dienst Graubünden;
    ↪ Denkmalpflege Graubünden",
    "journalPublisher": "",
    "journalMovingWall": 0,
    "volumeLink": "/digbib/view?pid=adg-001%3A2005%3A0",
    "volumeTitle": "Band - (2005)",
    "volumeShortTitle": "Band -",
    "volumeNumYear": "- (2005)",
    "issueLink": null,
    "issueNumber": null,
    "issueTitle": null,
    "label": "82",
    "gotoLink": "/digbib/view",
    "imageLink": "https://www.e-periodica.ch/iiif/2
    ↪ /e-periodica!adg!2005_000!adg-001_2005_000_0085.jpg
    ↪ /full/800,/0/default.jpg",
    "imageId": "adg!2005_000!adg-001_2005_000_0085.jpg",
    "nextPageLink":
    ↪ "/digbib/view?pid=adg-001%3A2005%3A0%3A%3A89",
    "prevPageLink":
    ↪ "/digbib/view?pid=adg-001%3A2005%3A0%3A%3A87",
    "nextPagePid": "89",
    "prevPagePid": "87",
    "viewerLink":
    ↪ "/digbib/view?pid=adg-001%3A2005%3A0%3A%3A88",
    "overlay": false,
    "overlayLink": null,
    "pdfLink":
    ↪ "/digbib/../cntmng?pid=adg-001%3A2005%3A0%3A%3A88",
    "pid": "88",
    "showGotoPage": false }
```

```
{"fulltext": "Kurzberichte Alvaneu, Bahnhof LK 1216,
↪ 769 420/171 288, 1005 m ü. M. Im Sommer lieferte
↪ Hansueli Tinner-Guler, Landquart, dem ADG ein
↪ Bronzeschwert ab, das er einige Zeit zuvor im
↪ Bereich des Bahnhofs Alvaneu gefunden hatte.110
↪ Tinner-Guler, der Lokomotivführer bei der
↪ Rhätischen Bahn ist, musste im Bahnhof Alvaneu
↪ einen entgegenkommenden Zug abwarten. Um die
↪ Wartezeit zu überbrücken, nutzte er die
↪ Gelegenheit für einen kurzen Spaziergang längs der
↪ Geleiseanlagen. Dabei fiel ihm im östlichen
↪ Bereich des Bahnhofs in der nördlich gelegenen
↪ Bahnböschung ein grünes Objekt auf, das 4-5 m
↪ oberhalb der Geleise in der Böschung lag.
↪ Anlässlich eines Besuches des Rätischen Museums in
↪ Chur (RM) erkannte Tinner-Guler, dass es sich beim
↪ Objekt um ein bronzezeitliches Schwert handelt,
↪ worauf er unverzüglich Meldung an das RM und an
↪ den ADG erstattete. Tinner-Guler teilte dem ADG
↪ mit, dass ihm völlig unklar sei, wie das Schwert
↪ an seine Fundlage gelangt sei; er könne sich
↪ vorstellen, dass es bei
↪ Geleise-Sanierungsarbeiten, die nach dem Unwetter
↪ im Herbst 2002 durchgeführt werden mussten und die
↪ vorwiegend in der Nacht stattfanden, durch
↪ Baggerarbeiten ans Tageslicht gelangte und
↪ anschliessend von einem Arbeiter die Böschung
↪ hinaufgeworfen wurde. Abklärungen bei der
↪ Verwaltung der RhB in Chur ergaben aber, dass
↪ diese Hypothese kaum wahrscheinlich ist, da jene
↪ Geleisesanierungen westlich des Bahnhofes von
↪ Alvaneu stattgefunden hatten, während im
↪ Ostbereich keine solchen Arbeiten vorgenommen
↪ wurden. Hingegen waren im Jahr 1990 im ganzen
↪ Bahnhofareal grössere Umbau- und Ausbauarbeiten
↪ durchgeführt worden, wobei auch einzelne Geleise
↪ neu angelegt und Abb. 55: Alvaneu, Bahnhof.
↪ Bronzeschwert vom Typ Stätzling (12. Jahrhundert
↪ v. Chr.). Mst. 1:3. 82",
"width": 800,
"height": 1053,
"words": [],
"hasReferences": false,
"hasJournalEditor": true,
"hasJournalPublisher": false,
"hasDetails": true,
"hasIssueLink": false}}
}
```


Listing 1: Returned article in JSON format as response of the AJAX call (https://www.e-periodica.ch/digbib/ajax/pageinfo?pid=adg-001%3A2005%3A0%3A%3A88 [30.05.21])

3.2 Data used for image analysis

The image analysis included georeferenced images, which originated from the Flickr platform. To limit the data to the areas of interest (Lavaux and RhB), a spatial intersection was performed. To do so, a dataset of the UNESCO heritage regions was used.

3.2.1 Flickr data

line_nr	94210831
photo_id	7236383786
id_hash	a2fd63612a4ecc46936f26a49b55f274
user_nsid	14900985@N07
user_nickname	Steve+Groom
date_taken	20.05.12 09:24
date_uploaded	1337546848
capture_device	NIKON+CORPORATION+NIKON+D4
title	Sonloup+overlooking+the+Lavaux
description	Narcissus+Trail+2012
user_tags	les+avants,narcissus+trail
machine_tags	-
lng	6.938604
lat	46.457115
accuracy	16
page_url	http://www.flickr.com/photos/14900985@N07/7236383786/
download_url	http://farm8.staticflickr.com/7071/7236383786_058bc17334.jpg
license_name	Attribution-NonCommercial-ShareAlike License
license_url	http://creativecommons.org/licenses/by-nc-sa/2.0/

Table 3.1: Example of the datastructure of a Flickr image

The Flickr data used in this study is called Yahoo Flickr Creative Commons 100 Million Dataset (YFCC100M), which was collected as part of the Yahoo Webscope program by Thomee et al. (2016). In total, 100 million media items were collected which were all published with a creative common license. As the main focus of this study lies in Switzerland, a smaller dataset containing only images geo-referenced in Switzerland was used. This reduced the data to about 700'000 images, taken between 2002 and August 2019. The creation of this subset was done by the Geocomputation Unit of the University of Zurich.

Each image contains metadata, such as the name of the user, the time it was taken or uploaded, the camera used and more information. In table 3.1 the data structure of the Flickr images is shown based on one exemplary image.

As the data was already pre-processed and used, many more attributes are available for each image. They are not listed in the example, as they were not used for the image processing.

Of importance in this study were the following attributes:

- *user_nsid*: serves as unique identifier for each user, which was used to count the number of images taken per user
- *lng* and *lat*: the coordinates of the image taken was used to identify whether the respective image was taken in one of the two regions of interest (RhB and Lavaux).

3.2.2 UNESCO heritage regions in Switzerland

The Federal Office of Culture maintains a dataset containing all regions and sites which are classified to be UNESCO world heritage (Federal Office of Culture FOC, 2016). This data is constantly updated and made available online as OGC service or as ESRI shapefile. The data is divided into two data sets: one containing the UNESCO world natural heritage and one containing the UNESCO world cultural heritage. All the sites are divided into world heritage property, which contains the main area of the region and a buffer zone, which is an extended zone around the property. In figure 3.1 both the world heritage property of Lavaux and the RhB are mapped including its buffer zones.

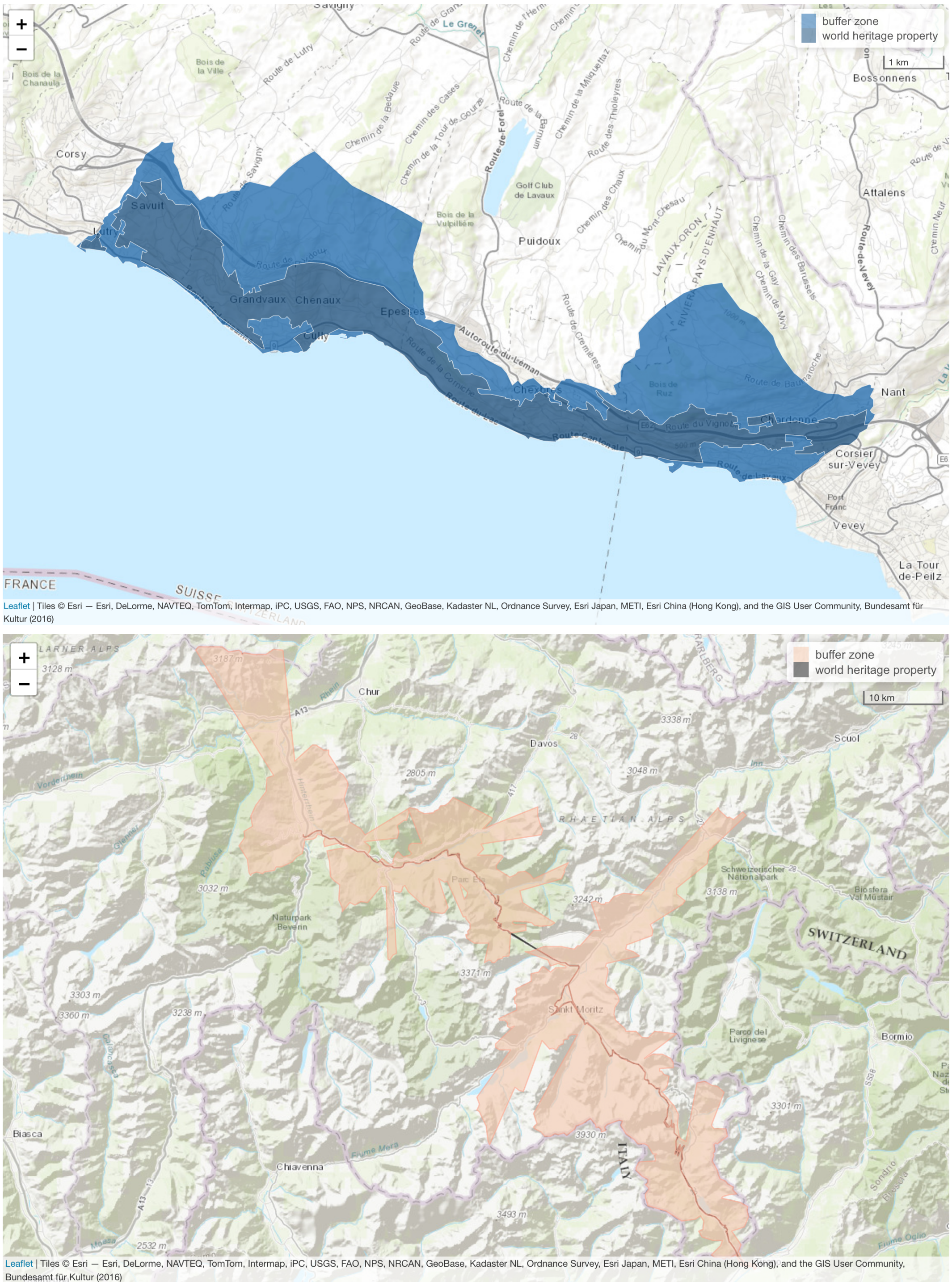

Figure 3.1: World heritage properties of Lavaux and RhB. Top: UNESCO region of Lavaux, divided into world heritage property and buffer zone. Bottom: UNESCO region of RhB, divided into world heritage property and buffer zone. Data provided by Federal Office of Culture FOC (2016).

Chapter 4

Methodology

The procedure included three main parts: the text analysis, the image analysis and the annotation of the resulting word lists. Each part included further processing steps. The workflow of the procedure is visualised in figure 4.1 and the individual steps described in more detail in the following chapter.

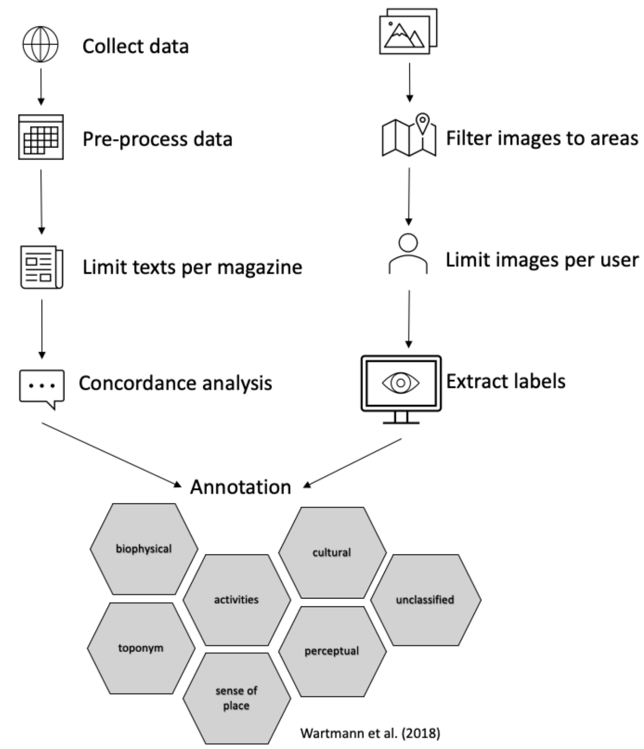

Figure 4.1: Overview of the workflow. The steps for the text analysis are shown on the left side, the steps of the image analysis on the right side.

4.1 Software and Scripting

Due to the wide options of packages and possibilities, Python (PyCharm) was used for most of the analysis. Namely, this included the web-scraping for text data acquisition, processing of the data as well as the concordance analysis as part of the text analysis. Furthermore, Python was used for the spatial intersection of the image data and the use of the Google API as well as the calculation of Cohen's Kappa.

Due to personal knowledge in visualizations and the creation of maps, the plots and maps were created using R (RStudio). More information about the exact software and the used packages can be found in the appendix.

4.2 Text analysis

In order to analyse text data, a corpus needed to be created first and data gathered. With the resulting text data, an analysis about its content was made. In the following chapters the methods of both building the corpus as well as the following text analysis are further explained.

4.2.1 Building a corpus

To build a corpus for analysis, text files were extracted from E-Periodica which were then further filtered according to the language they were written as well as the year the article was published. Further pre-processing steps included the removing of similar articles as well as the limitation of articles coming from the same magazine.

Data Extraction

The data was extracted using Web 2.0 technologies such as AJAX calls and the use of the server's API. E-Periodica was queried using an AJAX call, which returned each articles related to the search term. Search terms used for the two regions are *Lavaux* respectively *RhB*. By using the place name, the returned articles are georeferenced to the region (Hu, 2018). Each text including its metadata can be accessed with a unique identifier *pid*. When querying the site, all returend *pids* are written in a list. To locally save the JSON files for each text, each *pid* was called to return the respective text and metadata as JSON file. An example of the AJAX call is shown in listing 1.

However, the data is not structured. Each page of a magazine has its own *pid*. In the case that an article is longer than one page, the article is split into several *pids*. To get one file per article, the files needed to be joined according to their *pid*. Furthermore, only a subset of all the metadata information were kept, namely the articles, the journal title, the volume and the year of its publication as well as the fulltext. For each article one JSON file resulted.

Precision
Precision defines the fraction of returned results that are actually relevant to the information needed (Manning, Raghavan et al., 2008). To find whether the resulted articles are relevant to the topic, precision was considered. The precision was calculated by looking at the first 10 search results delivered by the search engine of E-Periodica following the procedure of Girit et al. (2012). Each article was read and classified as either *relevant* or *non-relevant*. In this case only relevance to the area of interest (Lavaux or RhB) was considered, not thematic relevance. Precision was calculated by dividing the number of true positives (retrieved and relevant) by the total number of retrieved, which in this case is 10. As visible in table 4.1, the obtained values showed high precision for both regions.

Lavaux	RhB
1.0	0.9

Table 4.1: Precision values for the retrieved articles from E-Periodica for each region.

4.2.2 Pre-processing data

Filter according to year of publication and language

In a next step, the articles were filtered according to their year of publication. Thereby, articles published after the year the respective landscape was voted to be of cultural heritage were filtered and no longer part of the data set. In the case of Lavaux, this was the year 2007 (UNESCO World Heritage Centre, 2007), the region of the RhB was added to the list in 2008 (UNESCO World Heritage Centre, 2008b). With this step, the bias of text containing elements of placemaking could be reduced. Furthermore, the language of the articles was checked. In the original data set, articles in all official Swiss languages were present as well as English articles. However, for this study the data was reduced to only German articles.

Check for duplicity

Next, the content of the articles was checked such that no duplicate or very similar articles are present in the corpus. This was done such that the resulting landscape properties are not falsified by including the same article article multiple times (Koblet and Purves, 2020b). In a first attempt, the fuzzy string method was applied to check duplicity (Zachara and Pałka, 2016). However, this turned out to be a very time-consuming task, as all string tokens are compared to another. As an alternative approach the Jaccard coefficient was calculated to measure similarity of text (Hajishirzi and Ave, 2010). Thereby, the articles are divided into *n*-grams. *n*-grams are groups of *n* consecutive words in the article. In this case the size was set to five, meaning the articles were divided into groups of five consecutive words. Afterwards, the *n*-grams are compared using the Jaccard coefficient to measure the similarity between the shingles.

The Jaccard coefficient lies between 0 and 1, whereas 1 indicates a perfect agreement. A value of 0 on the other hand shows that no similarities are detected. Articles were considered to be duplicates, if their Jaccard coefficient was higher than 0.8. In this case, one of the two articles were removed from the data set.

Limit number of articles per magazine

The boxplots in figure 4.2 show the distribution of number of articles per magazine present in the data for each study site. The site of Lavaux shows a high range with several outliers and a maximum value of 496 articles by one magazine. However, the median is only 2 articles per magazine. In addition, the interquartile range does not show a great range but is rather low (4). The site of the RhB on the other hand shows a smaller total range with a maximum value being 107 articles by one magazine, but both the median (3) and the interquartile range are higher (7). Moreover, no outliers are visible.

In order to not get a biased result due to a high number of articles contributed by only one magazine, the number of articles per magazine was limited to 25. To do so, the number of articles per magazine was counted. For magazines which contributed more than 25 articles a random subset of 25 articles was chosen. This step reduced the number of articles taken into account to 1023 for the site of Lavaux and 896 for RhB.

The final data was written into a comma-separated values (csv) file, one file for each region containing the remaining articles and the respective metadata.

4.2.3 Text analysis

The remaining data was used for the text analysis. Therefore, the Natural Language Toolkit (NLTK) provided by Pyhton was used. In order to analyse the text data, different steps were applied.

Tokenizing

To analyse text computationally, the data needs to be divided into smaller units of text (Chesnokova et al., 2019). The smallest units in natural language are called tokens which could be words, punctuation but also lemmas or

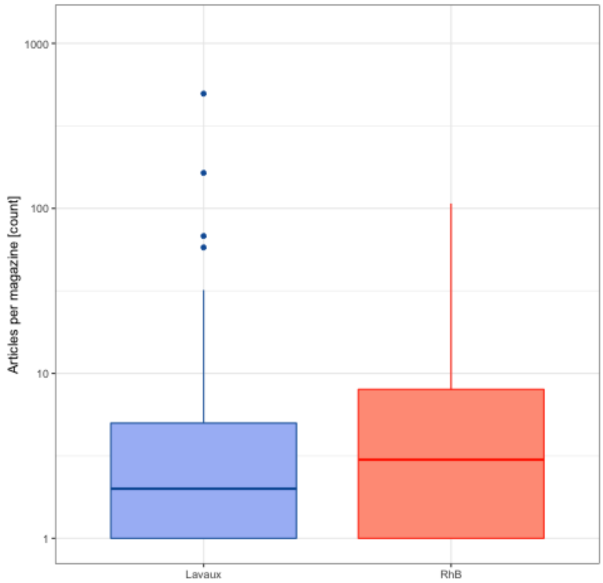

Figure 4.2: Boxplots showing the article count per magazine by region. The y-axis is log10 for better readability.

stems (Manning, Schütze et al., 2002). Furthermore, the text data was normalised to lowercase. These steps are necessary to successfully compare text units to another.

Removing stop words

In a next step, stop words were removed from the data. Stop words include words which occur very often in a natural language, but do not provide any additional information (Manning, Raghavan et al., 2008). Examples of stop words in German would be *sein, und, der, die, ein, ...*. By removing stop words and punctuation, the data remains more informative.

Lemmatization

After tokenizing and removing of stop words, the remaining tokens were lemmatized. Lemmatization reduces single words to their root (Manning, Schütze et al., 2002). Words with the same semantic meaning are thereby reduced to the same lemma. An exmple hereby is the word *spielte* which is transformed to *spielen* after lemmatization. By lemmatizing the data, words with the same lemma are not counted separately but together.

The Python library *spacey* was used for lemmatization, as it provides lemmatization for German words.

```
[ConcordanceLine(left=['umfahrungstunnel', 'flims',
↪ 'felsdreieck', 'nah', 'strasse', 'lichtdioden',
↪ 'reflektierend', 'schicht', 'glasperlen',
↪ 'überziehen', 'tag', 'nacht', 'nebel', 'ausschneiden',
↪ 'schimmern', 'bergfeuer', 'lässt', 'grenzen',
↪ 'kunst'], query='landschaft', right=['zerfliessen',
↪ 'verbinden', 'natur', 'technik', 'fläche',
↪ 'quadratmetern', 'oben', 'wald', 'unten', 'strasse',
↪ 'begrenzen', 'brigitte', 'kowanz', 'feuer',
↪ 'erinnern', 'bergsturz', 'einen', 'beletage'],
↪ offset=436, left_print='ern bergfeuer lässt grenzen
↪ kunst', right_print='zerfliessen verbinden natur
↪ techn', line='ern bergfeuer lässt grenzen kunst
↪ landschaft zerfliessen verbinden natur techn')]
```

Listing 2: Example of the concordance list. The listing *left* shows the words occurring on the left side of the query term, the listing *right* the words on the right side. The *offset* describes the offset location where the query term was found, whereas the *line* shows the context by representing the exact line. This example shows the concordance list for the article *Ein Pakt mit dem Piz: Naturpark im Mittelbünden* by Marco Guetg published in the journal *Hochparterre: Zeitschrift für Architektur und Design* in 2005

Concordance analysis

The final data was used for concordance analysis. Therefore, the corpus is searched for a query term and displayed with its immediate context (Gregory and Hardie, 2011). Words concording with the term *landschaft* were searched. Therefore the python library NKLTK was used, which provides numerous functions for natural language processing. As a result, a list is returned which contains objects concording with the query term (see listing 2). For further analysis, all the words originating from the left side of the query word as well as the words from the right side were structured into a dictionary. If the word is already present in the dictionary, the count of the word is increased by one. However, if the word is not yet contained in the dictionary, a new entry is created and the count set to one. Furthermore, the line in which the query term was found was written as a whole into a text file. Thereby, the context of the query term can be analysed enhancing the understanding further.

Precision
To quantify the relevance of the articles, precision was calculated. Following the procedure of Girit et al. (2012), the first 10 results containing the term

landschaft were looked at and classified as *relevant* or *non-relevant*. In this case, articles were considered relevant when they contained information about the landscape of the region. The retrieved precision values are lower than the precision of the search query on E-Periodica. Especially, the value for Lavaux decreased strongly (see table 4.2). Reason for this can be found in the type of magazines and the ambiguity of the word *landschaft*. Further elaboration on the representativity of the data set can be found in chapter 6.1.

Lavaux	RhB
0.8	0.6

Table 4.2: Precision values for the articles retrieved with relevance to region's landscape.

4.3 Image analysis

The image data set contained Flickr images which are georeferenced in Switzerland. Pre-processing the data included the use of a point in polygon algorithm such that the data is reduced to the areas of interest and limiting the number of images per contributor which were taken into account. For the analysis itself, Google Vision API was used to gather automated tags. All these steps are explained with more detail in this section.

4.3.1 Pre-processing of images

The image data set contained Flickr images in Switzerland from 2002 to 2019. In total, 716'468 data entries are present. All the images are geo-tagged at a specific location. In this analysis, only the images taken in the two UNESCO world hertiage areas, RhB and Lavaux, were taken into account.

Filter regions

The Federal Office of Culture published the data set *UNESCO World cultural heritage* in which all the classified UNESCO world heritage areas are present (Federal Office of Culture FOC, 2016). The regions are divided into *World Heritage properties* und *Buffer zones*. This data was filtered such that one shapefile for each area of interest was extracted containing both the *World Heritage property* as well as its respective *Buffer zone*.

Points in polygon

To do a spatial query between the Flickr images and the bounding polygons, the data need to have the same projection. The projection of the polygon data is LV03, whereas the flickr data was projected in WGS84. Before performing a point in polygon algorithm, the polygon data was reprojected to WGS84. Afterwards the geo-tagged Flickr images were used in a spatial query, such that images which were taken in the area of interest could be found and used for further analysis. To do so, a binary mask, *pip_mask*, was created. This mask contained the boolean information whether the respective point was in- or outside the polygon in question. In a next step, the *pip_mask* was used to find all points which were inside the polygon. To get a overview of the images taken in the regions, the data was plotted (see figures 4.3 & 4.4) and a kernel density estimation was mapped. The spatial intersection lead to a reduction of data. In total, 8294 images were taken in the region of RhB and 3310 in the region of Lavaux. However, the density of images taken is much higher in the region of Lavaux than in the region of the Rhb.

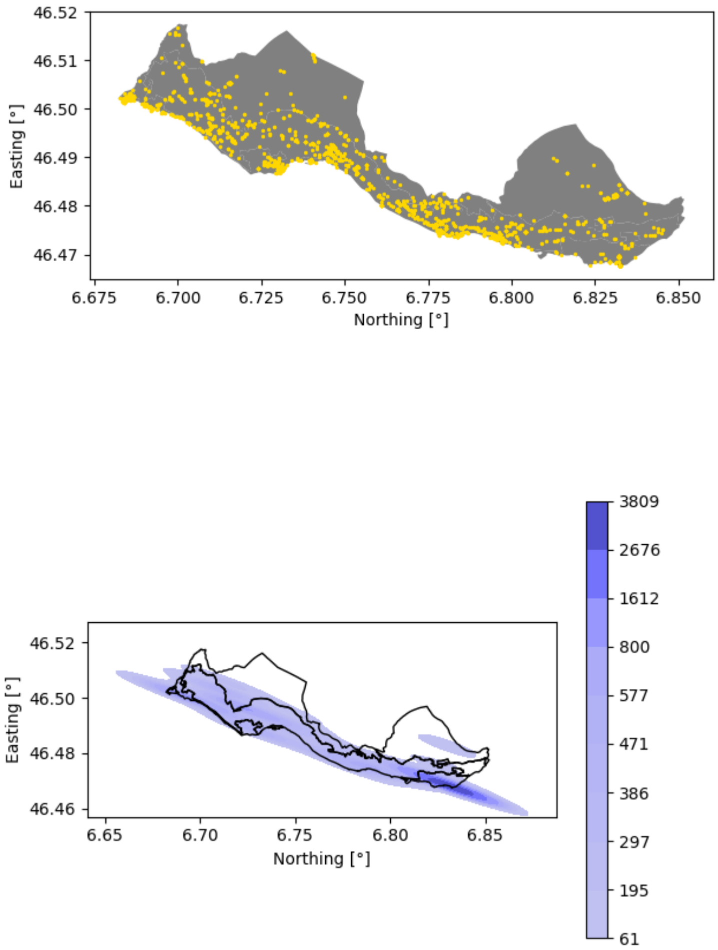

Figure 4.3: Flickr images taken in the region of Lavaux. Top: Flickr images taken in the region of Lavaux. Bottom: Kerndel density estimation of the images taken in the region of Lavaux. Data provided by Federal Office of Culture FOC (2016).

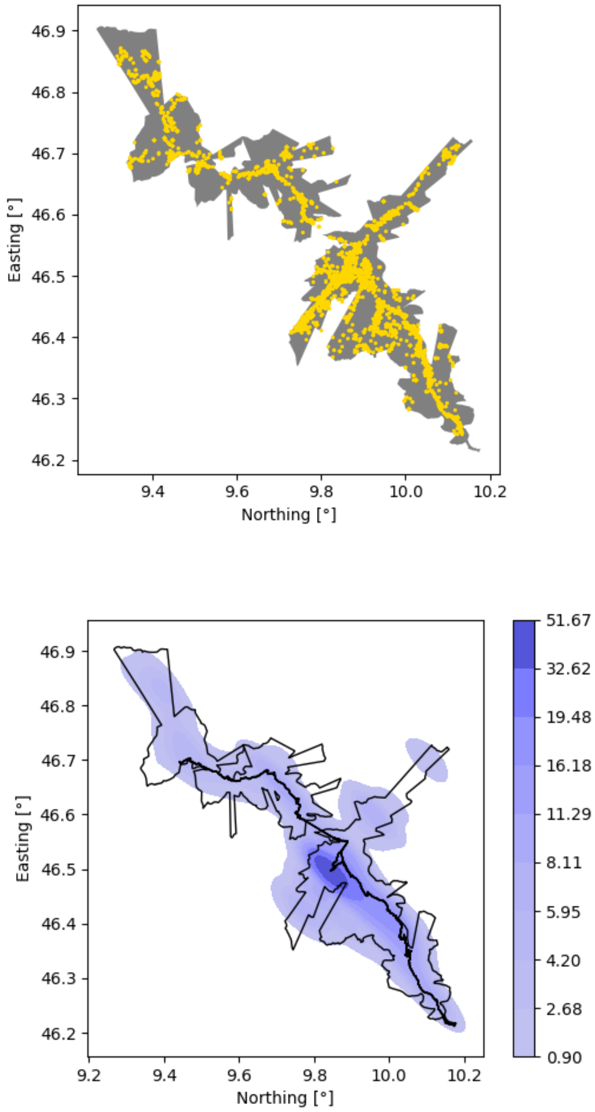

Figure 4.4: Flickr images taken in the region of RhB. Top: Flickr images taken in the region of Lavaux. Bottom: Kerndel density estimation of the images taken in the region of RhB. Data provided by Federal Office of Culture FOC (2016).

Limit number of images per user

As shown by Nielsen (2006) users of online communities do not contribute evenly to its content. Most often, only a small amount of users are responsible for most of the contributions. In figure 4.5 the distribution of number of images taken per user can be seen for each studysite. It is visible, that the interquartilerange is around 7 for Lavaux and 17 for the RhB respectively. However, one outlier shows in the site of Lavaux with more than one thousand images. In order to not get a bias due to these users with a high number of contributions the number of images per user was limited to 100 images per user. This step reduced the number of images to 6990 in the region of RhB and 2146 for Lavaux.

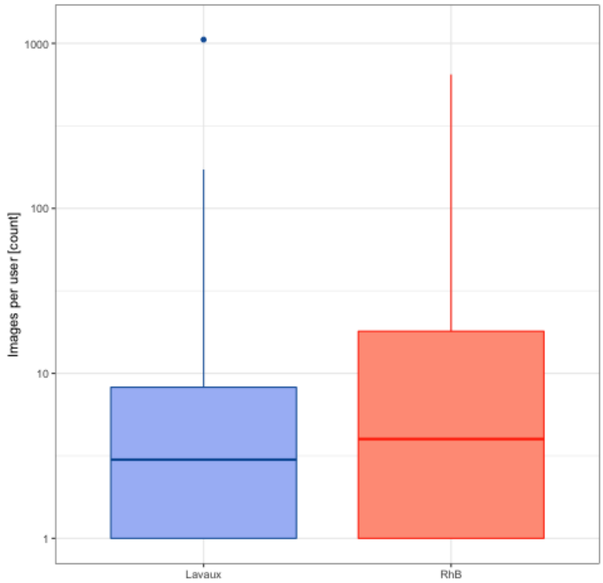

Figure 4.5: Boxplots showing the image count per user by region. The y-axis is log10 for better readability.

4.3.2 Google Vision

To analyze the content of the images and extract individual landscape properties, the Google Vision API was used. Google Vision API is a machine-trained tool which returns matching tags to the images based on what is visible on the image. Due to the machine-trained tool, the labels generated are reproducible and independent of the viewer. As a result no bias based on the objective opinion of a single viewer is generated. Each image was called using the Google Vision API which then returned annotating labels. An example can be seen in figure 4.6 and the corresponding table 4.3. The image was taken by a flickr user in the region of the RhB. By using Google Vision, the labels listed in table 4.3 were annotated with the image.

Thereby, several labels were returned for each image. All the labels were collected and counted according to their frequency. In the end, a csv file resulted containing all the annotated labels and their respective frequency.

Figure 4.6: Impression of the region of the RhB. Image taken by 'Marco Bianchi' (Flickr user 126883340@N08), licensed under Creative Common (https://www.flickr.com/photos/126883340@N08/35904453446/ [15.05.21])

train	plant	vehicle	cloud
mountain	tree	sky	green
rolling stock	highland	vegetation	grass
track	natural landscape	rolling	slope
railway	locomotive	landscape	railroad car
railroad car	passenger car	public transport	grassland
forest	hill station	valley	hill
mountain range	spring	travel	bridge
pasture	field	electricity	plantation
road	tourism	massif	transport
mountain pass			

Table 4.3: Labels annotated to the image in figure 4.6 by Google Vision

4.4 Annotation

Both the concording words from the text analysis as well as the labels from the image analysis were classified into the following categories according to Wartmann et al. (2018): *biophysical, cultural, perceptual, activities, sense of place* and *toponyms*. Furthermore, the category *unclassified* was added. To get a more detailed result, the categories were further divided into subcategories (see figure 4.4. Thereby Wartmann et al. (2018) and the Landscape Character Assessment were taken as guideline for the definitions of the subcategories. As suggested by (Gregory and Hardie, 2011) the aim is to annotate the elements in the corpus to find which properties are present when describing landscapes.

Category	Subcategories
biophysical	geology, soil, flora, fauna, climate, landform, land cover, hydrology
cultural	land use, settlement, infrastructure, domesticated animals, anthropogenic objects, time depth
perceptual	color, touch, feel, weather, atmospheric conditions, pattern
activities	-
sense of place	meaning, feelings, associations, attachments, identity, history
toponyms	-
unclassified	-

Table 4.4: Categories chosen for annotation according to Wartmann et al. (2018)

The annotation of the words from the text analysis and the received tags from the image processing was done manually. As a consequence, a bias due to the objective perception of the annotating person is created. By introducing subcategories, the classification aims to be more reproducible. However, the bias cannot be fully reduced. For both sites, Lavaux and the RhB, the 150 most returned words from text analysis and the 150 most returned tags from image analysis were annotated. The amount of annotated words are marked yellow und can be seen in figures 4.8 and 4.7 as well as figures 4.9 and 4.10.

4.4.1 Verification

Annotation is a challenging process, especially when using text data. In order to verify the annotation results, a cross-validation was completed and the inter-annotator agreement of Cohen's Kappa calculated as proposed by Chesnokova et al. (2019). Cohen's Kappa measures the reliability between several annotators and includes the possibility of agreeing by chance (Landis and Koch, 1977). Therefore, a fellow geography student annotated both the text data as well as the returned labels. In table 4.5 the agreement between the two annotators is listed. According to Landis and Koch (1977) the level of agreement for most areas is moderate. However, the annotation of the labels for the site of the RhB is higher showing substantial agreement. Furthermore, when looking at the level of granularity, the level of agreement is even higher when only considering the categories.

Granularity	Site	Text data	image data
Category	Lavaux	0.57	0.69
	RhB	0.66	0.78
Subcategory	Lavaux	0.58	0.43
	RhB	0.59	0.61

Table 4.5: Cohen's Kappa values for the level of agreement between the two annotators

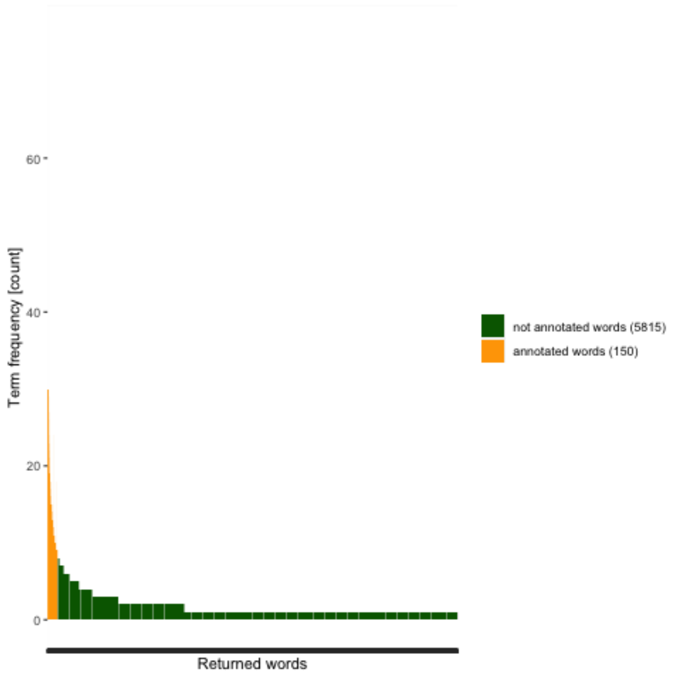

Figure 4.7: Frequency of returned words from the text analysis in the region of Lavaux. The 150 most common returned words are colored in yellow and were used for annotation.

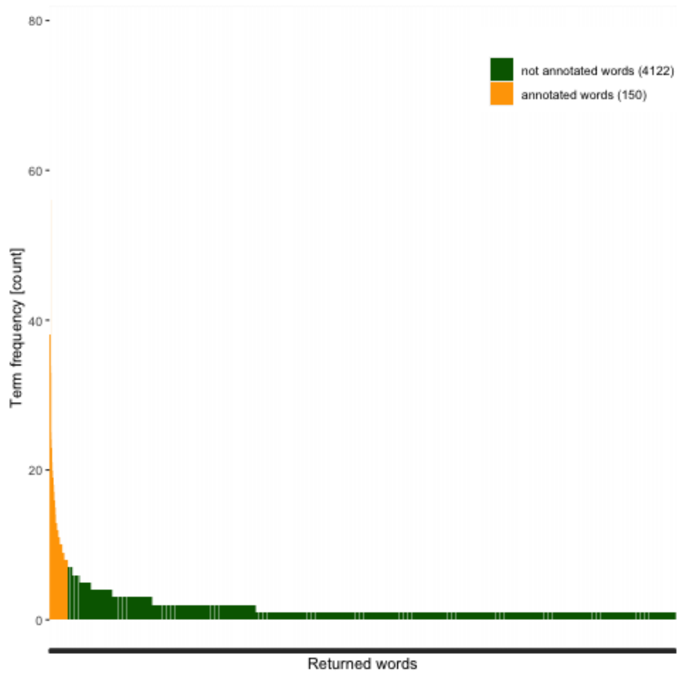

Figure 4.8: Frequency of returned words from the text analysis in the region of the RhB. The 150 most common returned words are colored in yellow and were used for annotation.

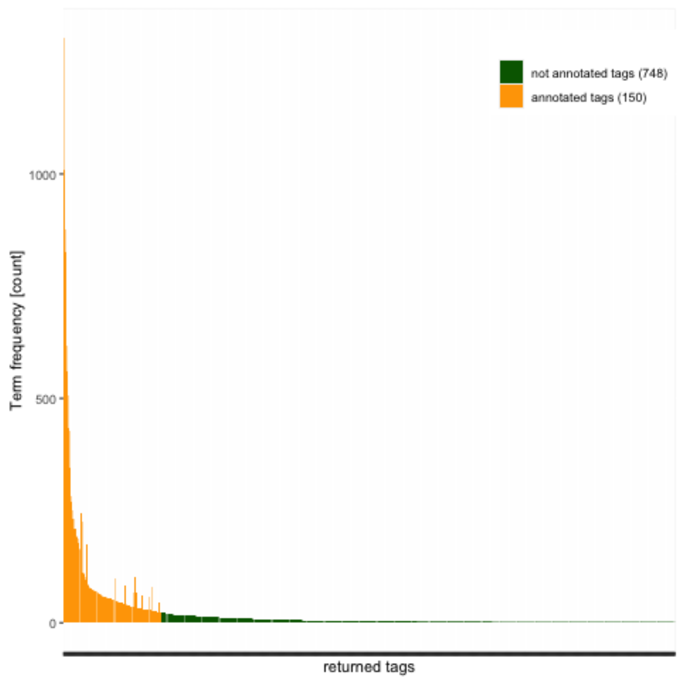

Figure 4.9: Frequency of returned tags from the image analysis in the region of the Lavaux. The 150 most common returned tags are colored in yellow and were used for annotation.

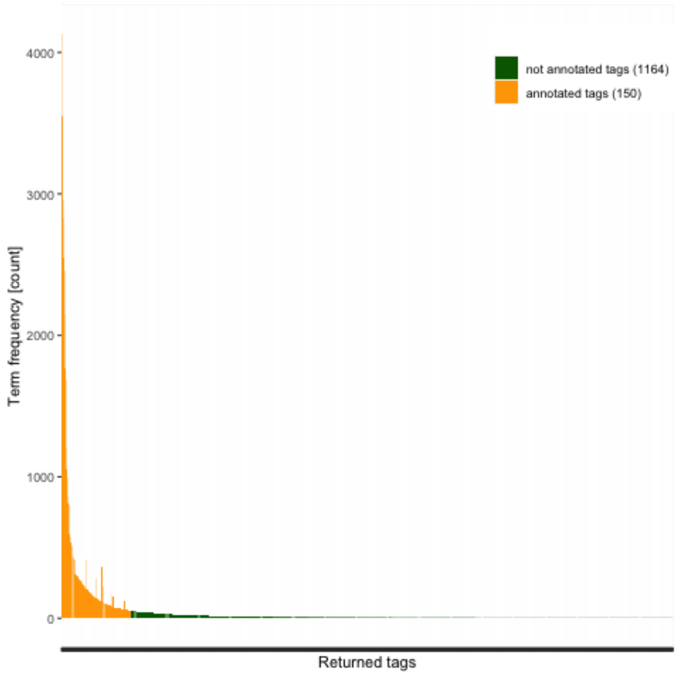

Figure 4.10: Frequency of returned tags from the image analysis in the region of the RhB. The 150 most common returned tags are colored in yellow and were used for annotation.

Chapter 5

Results

In this chapter the results of the analysis will be presented. In a first section, the results gained from the text analysis will be presented and differences as well as similarities between the two sites Lavaux and RhB will be pointed out. In a second part, the outcome from the image processing will be discussed. Again, the two different world heritage properties will be compared. In a final part, the two approaches, text analysis and image processing, will be discussed.

5.1 Results text analysis

After pre-processing the articles, a total of 1023 articles were used for the world heritage property Lavaux and 896 journal articles for RhB. The resulting words were annotated in the following categories: biophysical, cultural, activities, perceptual, toponyms, sense of place and unclassified. However, no words were annotated to the category *activities*, which is why this category is missing in the figures.

Lavaux	**RhB**
NIKE-Bulletin	HTR Hotel Revue
Schweizer Heimatschutz	die Schweiz: offizielle Reisezeitschrift der Schweiz
HTR Hotel Revue	Tec21
Bibliographica scientiae naturalis Helvetica	Bünder Monatsblatt
die Schweiz: offizielle Reisezeitschrift der Schweiz	Hochparterre

Table 5.1: Magazines contributing the highest amount of articles per region

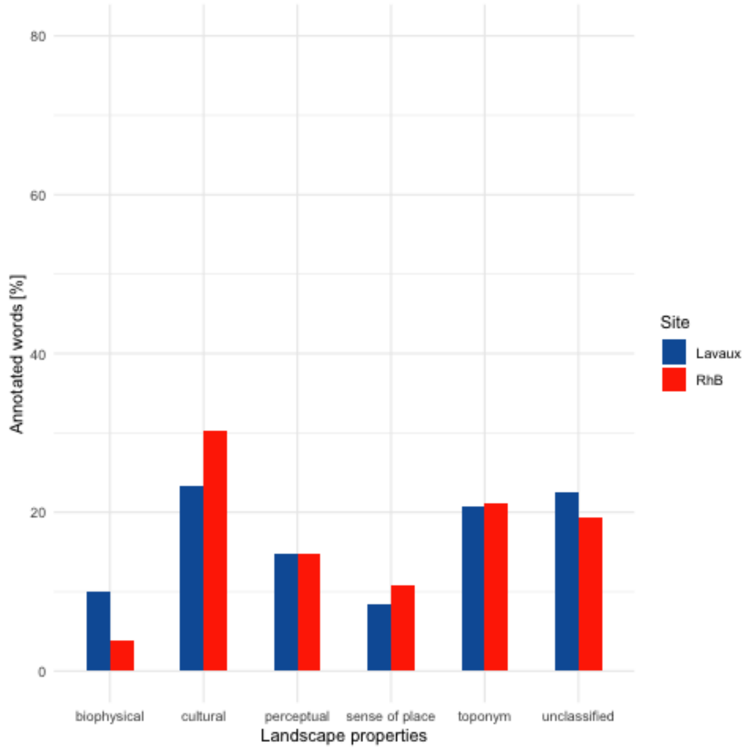

Figure 5.1: Fraction of different landscape properties (Wartmann et al., 2018) for each region extracted by text analysis.

5.1.1 Lavaux

In the area of interest of Lavaux a total of 930 were analysed in which the term *Landschaft* occurred 1538 times. In total, 5986 different words were extracted using concordance analysis.

The words used in the text analysis result form articles. To understand the text corpus better, the magazines in which the articles were published needs to be looked at. In table 5.1 the magazines are listed, from which the most articles are from. For the site of Lavaux, the topics of these magazines do have an emphasis on the identity and the cultural background of the region. The *NIKE-Bulletin* focuses on the national, cultural heritage, the magazine *Schweizer Heimatschutz* publishes articles about the Swiss heritage. The *HTR Hotel Revue* and the magazine *die Schweiz.* on the other hand have a higher emphasis on promoting the touristic aspects of the region.

The annotation of the resulted words can be seen in figure 5.1. Most of the annotated words can be placed in the categories *cultural*, *toponym* and *unclassified* which contribute more than 20% of the annotated words. The categories *perceptual*, *biophysical* and *sense of place* contribute 14.7%, 10.1%

Figure 5.2: Word cloud representing the most common returned words with text analysis in the region of Lavaux.

and 8.5% respectively. No word was annotated to the category *activities*.

The high amount of toponyms present are also visible, when taking a closer look at the words, which occurred most. In the region of Lavaux, five toponyms are among the ten most returned words (see figure 5.2). An explanation for the high amount of toponyms can be found in the type of magazines in which the articles were published. Magazines focusing on heritage or aiming to lie en emphasis on tourism, include toponyms for the reader to add a geographic context. Furthermore, one of the magazines included in the analysis was the *Schweizerisches Handelsamtblatt*, which publishes short messages about newly founded companies. In general, these messages are really short containing information about the type and address of the companies. Due to the data structure of E-Periodica, only entire pages of a magazine were considered. In the case of the Handelsamtblatt, one page included up to ten different short articles containing a lot of toponyms. Furthermore, due to the Swiss canton Basel-Landschaft, the query term *landschaft* returned a lot of concordances.

Cultural properties found in the text analysis include aspects of infrastructure, time depth, settlement and land use (see figure 5.5). Especially, the aspects of land use and settlement are characteristic in the region. This can be explained with the high population density (722 inhabitants/km^2) but also with the cultivation of the vineyard terraces, which is of high importance in the region. However, within the most returned words which were annotated, only the word *Wein* directly shows a connection to the vineyard terraces. Partly, this can again be explained with the analysed articles, as thematic relevance of the articles cannot be guaranteed (for further elaboration see section 6.1). On the other hand, it can be explained by the use of the region. The region of Lavaux is not solely a world heritage property, but first and foremost it is a residential area in which people live.

The biophysical properties of the landscape are dominated by the landform and land cover (see figure 5.4), but also hydrological properties due to the lake of Geneva can be extracted.

One of the most returned words in the text analysis is the word *landschaft* itself. This is due to the parameters chosen for the concordance analysis. The window searching for concording words is set to 2 lines. However, most of the articles are rather short, which is why two lines often cover a large part of the text. Furthermore, if an article describes the landscape, the term *landschaft* occurs more than once. Together with the searching window this leads to a high amount of the term *landschaft*.

Furthermore, the properties linked to the history of the place emphasizing the identity of the region can be seen in the high amount of properties which were annotated in to the category sense of place.

5.1.2 RhB

869 articles were used in the region of the RhB. The term *landschaft* occurred 759 times and 4280 words were extracted with concording analysis.

In the region of the RhB the two magazines *HTR Hotel Revue* and *die Schweiz.* focusing on the touristic aspect of the region can be found as well. The magazine *Bündner Monatsblatt* examines cultural aspects in the entire canton Grisons. Other magazines contributing a high amount of articles to the analysis have a strong focus on architecture and construction (*Tec21, Hochparterre*).

Solely based on the field in which the magazines are situated, it can be expected, that a high amount of words were annotated in the categories *cultural* and *sense of place*. In figure 5.1 distribution of the annotated words into the different categories can be seen. More than 30% of the returned words were annotated to the category *cultural*. On the other hand, the category

Figure 5.3: Word cloud representing the most common returned words with text analysis in the region of the RhB.

biophysical only contains 3.8% of the annotated words. The categories *toponym* and *unclassified* contain a similar amount of annotated words which is around 20%. 16% of the returned words are of the category *perceptual* whereas the category *sense of place* holds 8% of the resulted words.

The high amount of cultural properties corresponds the to type of magazines the articles are from. The main focus on the texts lies on the architecture and infrastructure of the railway leading to a high amount of words which were annotated in the subcategory *infrastructure.* This can also be seen when looking at the words directly (see figure 5.3). Words such as *bauen*, *technik*, *architektur* emphasize the technical aspect of the landscape.

Although the natural landscape is an important part of the region, the category *biophysical* shows the smallest share of words. Characterising aspects of the landscape such as *berg* or *tal* are missing in the most returned words and as a consequence these words were not annotated. Based on the text corpora, the technical aspect as well as the development and history of the place are more crucial aspects of the landscape. Therefore more words were annotated in these categories.

Furthermore, both the categories *perceptual* and *sense of place* show a high amount of annotated words. This includes words with an emphasis on the development and history of the region such as *bedeutung, geschichte* or *verändern* but also feelings towards the region (*wichtig*). When writing about a region, the background and history of the region plays an important part. Especially when writing in a touristic journal, the history of the region is often elaborated to give the reader more insight.

5.1.3 Comparing Lavaux and RhB

At a first glance, the distribution of the annotated words into the different categories is similar. Cultural properties can be extracted the most, whereas only little biophysical properties were found. Nevertheless, these are also the two categories which show the highest differences between the two regions, with more than 5%. Furthermore, the distribution within these categories differs strongly (see figures 5.4 and 5.5).

The main differences in the biophysical properties of the two regions can be explained with the differences in landscape and landform. Lavaux is situated next to the lake of Geneva with a hilly landscape, these properties can be found in the amount of words that were annotated in the subcategories hydrology, land cover and landform. The region of the RhB of the other hand is influenced by the climate of the Alps, which can be seen at the amount of words annotated in the subcategory *climate*.

The distribution of the cultural category follows a similar pattern for both sites. However, some differences can be seen. The most apparent difference can be found at the amount of words annotated in the subcategory *infrastructure*. In the region of the RhB more properties realted to the infrastructure can be extracted than in Lavaux. This is not surprising when the region is considered. The region of the RhB includes the railway aiming to transport people through the region, whereas in Lavaux the landscape as a whole is more part of the region.

In general, the differences found in the results of the text analysis between the two sites can be traced back to the differences of the regions themselves. In addition, the diversity of returned words needs to be mentioned. The region of Lavaux was described in a more diverse way than it was the case of the region of the RhB. The total amount of returned tags is higher in the region of the Lavaux, but also frerquencies of the returned tags are higher. In the region of Lavaux 431 tags have a frequency of five or higher, whereas in the region of the RhB only 240 have a frequency of at least five.

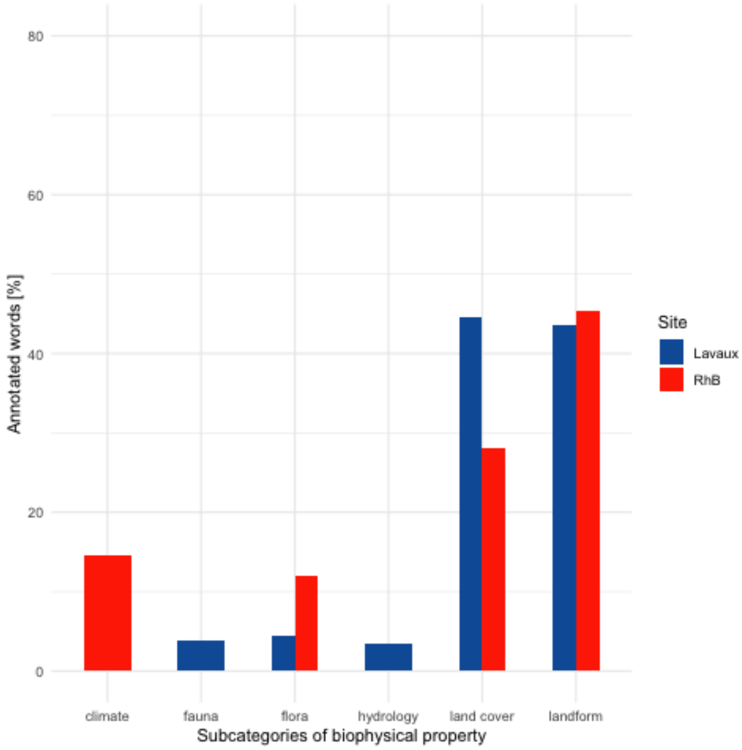

Figure 5.4: Fraction of subcategories composing the category *biophysical* extracted by text analysis for each region.

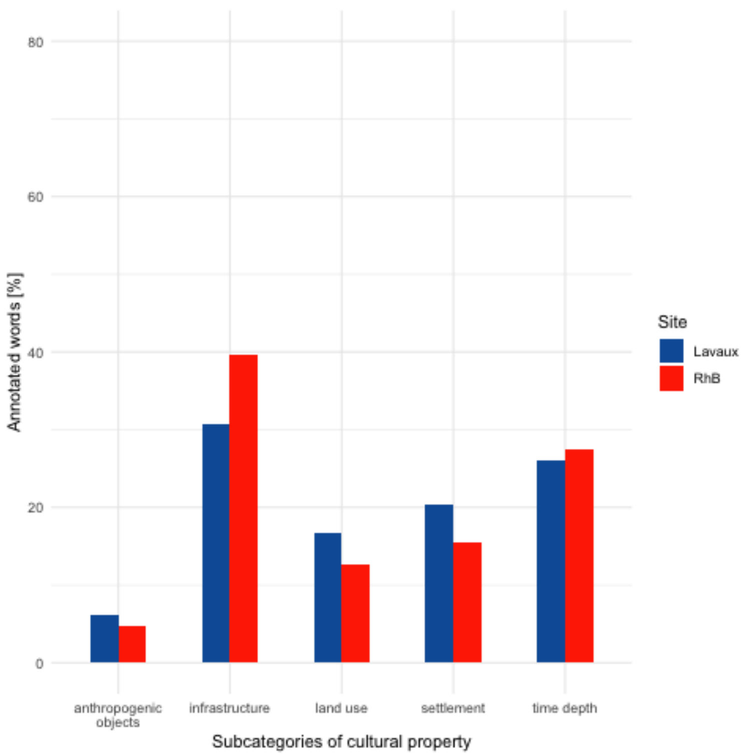

Figure 5.5: Fraction of subcategories composing the category *cultural* extracted by text analysis for each region.

5.2 Results image processing

The returned tags were classified into the same seven categories as before with the text analysis: biophysical, cultural, activities, perceptual, toponyms, sense of place and unclassified (Wartmann et al., 2018). However, among all the labels, no toponyms were returned. As a consequence, the category *toponym* is entirely missing leading to only six instead of seven categories.

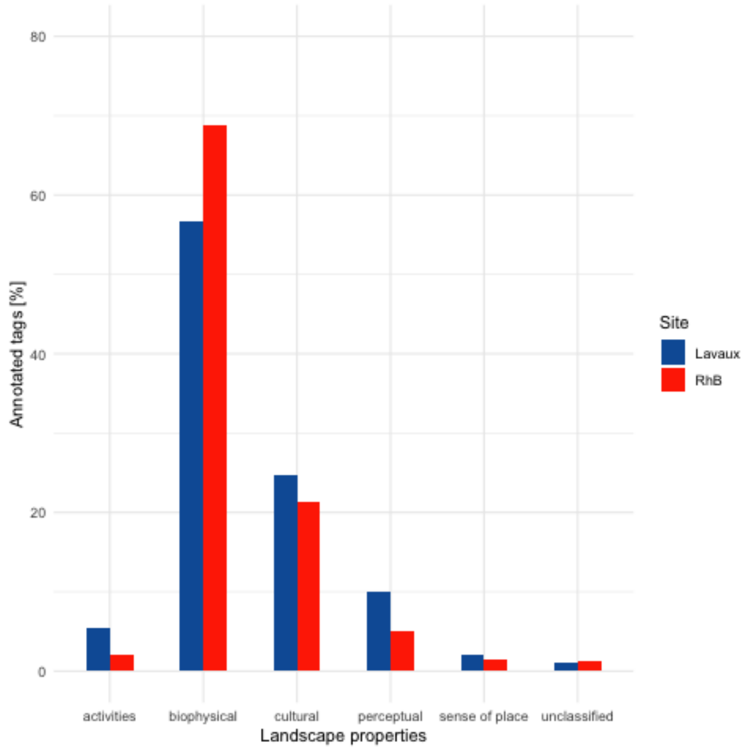

Figure 5.6: Fraction of different landscape properties (Wartmann et al., 2018) for each region extracted by image analysis.

5.2.1 Lavaux

In the region of Lavaux a total of 2146 images were taken into account. Each image returned multiple labels, which resulted in a total of 898 different tags.

When looking at figure 5.6, it becomes visible, that biophysical properties form the largest class among the returned labels. 55% of the classified returned tags are labelled *biophysical*. This can also be seen when looking at the wordcloud of the returned labels in figure 5.7. The most common returned tags are annotated in the biophysical category. Looking at the com-

Figure 5.7: Word cloud representing the most common returned words with image analysis in the region of the Lavaux.

position of the labels in the category *biophysical* (see figure 5.9), it is shown, that the subcategories *climate*, *flora*, *hydrology* and *landform* are the most present subcategories. Around 95% of all the tags classified as *biophysical* can be found in one of these subcategories. The share of the remaining categories is less than 5%, whereas labels concerning the geology of the landscape are entirely missing. When looking at the wordcloud, these subcategories can be clearly seen with tags such as *mountain*, *water*, *tree* or *coastal and oceanic landform*. These properties represent the natural landscape of Lavaux with the hilly terraces situated next to the lake of Geneva.

When taking a closer look at the class *cultural*, the subcategory *infrastructure* and *anthropogenic objects* are most present containing each about 40% of all the returned labels (see figure 5.10). This mainly includes labels related to buildings and residential areas. An explanation for this high amount of anthropogenic influence is the high amount of inhabitants inside the region of Lavaux. This includes the houses and buildings present (*building, window)* in the area as well as the transportation infrastructure (*vehicle, train)* or infrastructure for recreational causes(*swimming pool*).

Furthermore, a remarkable share of words were annotated in the category *activities*. Tags such as *swimming, sports* show that the region not only looked at as landscape with a high value, but the region is also inhabited and people are spending time within. Sport competitions, recreational activities or musical events are taking part within the region of Lavaux.

In the category of perceptional properties, the visual perception is showing a high amount of elements *green, azure, tints and shades*.

5.2.2 RhB

Figure 5.8: Word cloud representing the most common returned words with image analysis in the region of the RhB.

In the final subset of flickr images in the region of the RhB, 6990 images remained. Out of these the analysis with Google Vision returned 1314 different labels.

In the region of the RhB biophysical properties dominate. Out of all returned labels, about 68% are classified *biophysical*. When taking a closer look at this category, it can be seen that the largest part of it consists of the subcategories *landform*, *flora* and *climate* (see. figure 5.9) which combined make 86% of the labels in the *biophysical* category. The returned tags emphasize the alpine region (*mountain, snow, slope*). Furthermore, some properties annotated in the subcategory *geology* were found showing the importance of geological phenomenon in the region.

Looking at the composition of the labels classified as *cultural*, again two subcategories, *infrastructure* and *anthropogneic objects*, are strongly present. In the case of the RhB, a lagre amount of tags related to the infrastructure concern the transportation infrastructure, especially the railway infrastructure *train, railway, mode of transport*. Furthermore, some tags were annotated in the subcategory *settlement*, showing a relation to the villages present in the region. Additionally, some tags were annotated to the subcategory *domesticated animals* (*working animals*) showing the land use.

Only a small amount of activities were found in the returned tags for the region of the RhB. This indicates, that people do not spend time in the region of the RhB but use the train to pass through to get from one place to another.

The words annotated in the category *perceptual* include a high amount of visual elements.

5.2.3 Comparing Lavaux and RhB

When comparing the most common returned labels of Lavaux and RhB, a high similarity can be found. Seven out of the ten most common returned tags are listed in both sites. Reasons for this can be found in the amount of available tags with which the algorithm were trained (further elaboration can be found in section 6.5 limitations).

The general distribution of the categories presents itself similar for both regions. The landscapes are influenced strongly by biophysical properties, but also a fair amount of cultural properties can be extracted as well. However, the distribution of the subcategories composing the categories differ strongly. In Lavaux, a lot of the tags were annotated in the subcategory *hydrology* showing the influence of the lake of Geneva, whereas in the region of the RhB the geology and the landform are more relevant. These differences in the distribution of the categories represents the differences in the landscapes.

In the cultural category, the main differences between the different distributions can be explained with the purpose of the region. The region of Lavaux is inhabited, including residential areas. As a consequence, a high amount of tags were related to residential areas and infrastructure. In the RhB on the other hand, the mode of transportation and the railway itself is in focus explaining the high amount of tags annotated as infrastructure.

Furthermore, the extent of the tags needs to be elaborated. In the region of Lavaux, the images were more diverse leading to a higher amount of returned tags (Lavaux: 431 words mentioned at least 5 times, RhB: 240 words mentioned at least 5 times).

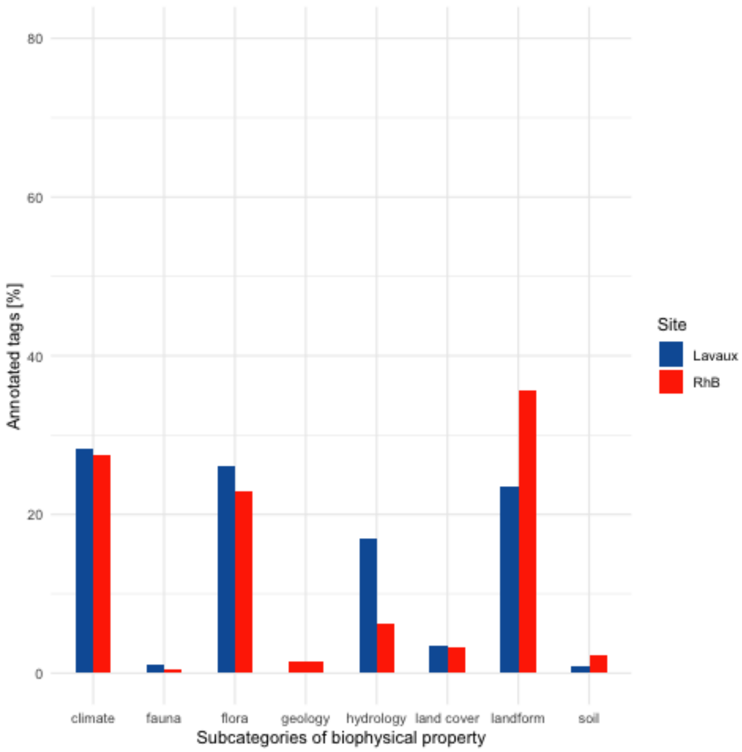

Figure 5.9: Fraction of subcategories composing the category *biophysical* extracted by image analysis for each region.

5.3 Comparison text vs. image analysis

When comparing the results of the text and the image analysis, it shows that part of *biophysical* labels is much higher in image analysis, whereas in text analysis the categories *toponym* and *cultural* are more prominent. Generally, the distribution within the different classes is more evenly in text analysis than in image analysis. The two classes *sense of place* and *perceptual*

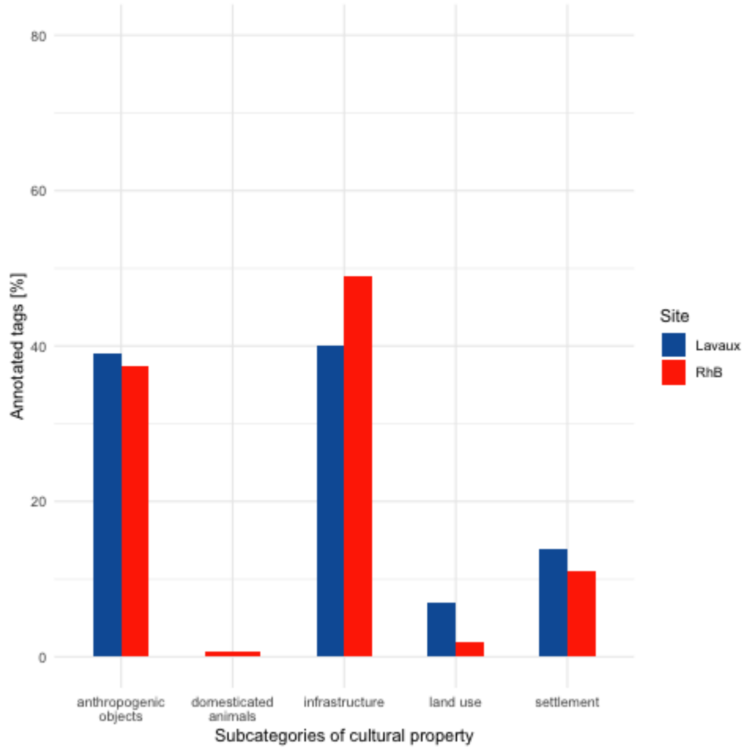

Figure 5.10: Fraction of subcategories composing the category *cultural* extracted by image analysis for each region.

contribute both around 10%, whereas they are less than 5% in the results of image analysis.

Clear differences can be seen when looking at the subcategories of the bio-physical category. The landscape in the region of the RhB shows a high proportion of words in the categories *landform* and *land use*, whereas hydrological as well as animal or plant (subcategories *flora* and *fauna*) related words are less

Furthermore, the text analysis shows a higher amount of returned words. When comparing the wordclouds of the image analysis, numerous of the most returned tags are visible in both regions. In the text analysis on the other hand, the extent of returned words is bigger.

Chapter 6

Discussion

In this chapter the obtained results are discussed. In a first section, the representativity of the data is explained showing possible bias which would influence the result. In a second part, the results of the text and image analysis are set in context and their similarities and differences are reviewed. Furthermore, the outcome is compared with the criteria of the UNESCO world heritage list to answer RQ3. In a final section, limitations of the analysis are explained and possibilities for future research is shown.

6.1 Representativity of the data

The resulting extracted landscape properties depend strongly on the input data. Only landscape properties which are present in the input data can be extracted. However, not all perceptions are reflected in the input data. Missing data leads to biased results, which is why it is important how the input data was created. By analysing the origins of the input data, it can be found whose perception is represented in the final results and what information might be missing.

6.1.1 Text data

When creating a corpus, temporal, spatial and thematic relevance should be addressed (Koblet and Purves, 2020a). In the following, the handling of the three types of relevance are elaborated.

Temporal relevance was given by filtering the data according to date. Only texts which were written before the area was considered to be world heritage are taken into account. Thereby, the influence of place writing, as described by Cooper and Gregory (2019), is limited. People writing about the landscape of Lavaux or the RhB do not know, that the landscape will be considered as UNESCO world heritage property. Nevertheless, text data transmit emotions of the author influencing the perception of the landscape (Cooper and Gregory, 2019). Moreover, it can be expected that the extracted properties transmit predominantly positive emotions due to the positive bias which exists in natural language (Dodds et al., 2015).

By using the place name of the area as query term on e-Periodica, the spatial relevance was ensured. Therefore, the place name was used as georeference (Hu, 2018). It is expected that most returned articles include the areas of interest, however it is not guaranteed that the query exclusively returns articles about the respective region. Nevertheless, the precision (see table 4.1) shows a high amount of relevant articles based on the place name for both areas of interest. Furthermore, word ambiguity was not found which increases the number of relevant articles.

Thematic relevance was given with the concording search term *landschaft*. Only texts and words which have a relation to landscape are used for the annotation. When using annotation analysis, it is assumed that an annotated word which is conocording with the query term, is relevant to the place (Gregory and Hardie, 2011). However, this assumption is not always met, as the term *landschaft* is not exclusively surrounded by descriptions of a landscape. Reason for this is that the word *landschaft* is ambiguous in Switzerland. On one hand *landschaft* is the German word for *landscape*. However, there is also a canton called *Basel Landschaft*. In listing 3 an example is shown of concording words with the term *landschaft* of the site of Lavaux. Due to the context of the sentences, the word ambiguity can be seen. The sentences are not complete as stop words are removed and the words are lemmatized. The first sentence in the example describes the landscape of Lavaux which is characterized by its vineyards. Furthermore, the uniqueness of the area is emphasized (*einzigartig*). On the other hand, the second example shows no relation to the vineyard terraces in Lavaux. The term *landschaft* is surrounded by toponyms of cantons in Switzerland, leading to the assumption that in this case the canton *Basel- Landschaft* was meant. This ambiguity is also shown when considering the precision (see table 4.2). Especially for the area of Lavaux, the precision has a rather low value indicating that the thematic relevance is not given in all articles.

Furthermore, the authors of the articles and the magazines in which they were published need to be discussed. By analysing the authors, it can be discussed whose opinions are represented (Cooper and Gregory, 2019). The magazines used for the analysis cover a broad range of interests. However, when looking at the magazine contributing several articles, magazines with an emphasis on tourism, traveling and protection of landscape identity dominate (see table 5.1). Is is assumed that magazines with a focus on tourism and travelling are fond of the landscape and region on which they choose to write about. Especially for travelling magazines the aim lies in promoting an

```
weinbaugebiet lavaux einzigartig landschaft formen
↪  weinbergen abstützen felsi

oni zürich zug fribourg solothurn landschaft schaffhausen
↪  appenzeu st gallen g
```

Listing 3: Example of words concording with *landschaft* in articles about the region of Lavaux.

area. The magazine HTR hotel revue for example discusses topics such as hotel business, gastronomy and tourism in Switzerland (Hotellerie Suisse, 2021). As a result, the articles will most likely transmit positive emotions about the landscapes. In the region of the RhB, the top returned magazines included magazines covering topics such as architecture or cultural identity (*Tec21, Hochparterre*). As a result, there is a strong focus on the technical and cultural aspect of the region, which is shown at the high amount of words annotated to the category *cultural* (see figure 5.1).

A final restriction can be found in the language in which the articles are written. The articles are filtered based on their language, such that only German articles were analysed. However, Lavaux is situated in the French speaking part of Switzerland and the region of the RhB covers not only the German speaking part of Grisons but also the Romansch. As shown by Regier et al. (2016), the extent of a language depends on the area it is spoken. By including only German articles, some properties could go missing due to translations. Furthermore, the possibility rises that the articles are written from an outside view and not by local people, which might have more knowledge about the area.

6.1.2 Image data

In general, when working with user generated content the user creating the images needs to be discussed. Although everybody is free to create user generated content, there are more men adding images to Flickr than women (Angradi et al., 2018). This creates a selection bias which describes that the sampled group is not representative for a larger population. On the contrary, one group of people is overrepresented (Seely-Gant and Frehill, 2015). This gender bias is the reason that the extracted properties of the two landscapes do not represent the perception of people, but especially how men perceive these regions.

Although the tags were generated by Google Vision and not by the user, the extracted landscape properties result from the images taken by the individual users and reflect their perception of the region. To limit the influence of a single user, the amount of images taken into account per user was restricted.

Furthermore, it can be assumed that contributors take pictures of landscapes which they perceive as beautiful (Gliozzo et al., 2016; Van Zanten et al., 2016). As a consequence, most images transmit positive emotions and beautiful landscapes whereas less pleasing landscapes are missing on platforms like Flickr. The UNESCO searches for landscapes which represent the interaction between humanity and nature showing the progress and technical achievements of people. Based on these criteria, the beauty of the landscape is not considered, but rather the cultural aspect is of importance. Therefore, less pleasing landscapes might also fulfill the criteria of UNESCO, but they are less likely to be found on photo-sharing platforms such as Flickr.

6.2 Extracting landscape properties

An important difference between the two regions lies in the use of the region. Although the RhB is listed as UNESCO world heritage property based on the difficulty it was to overcome the Alps by train and the resulting progress it brought to the local people to be better connected (UNESCO World Heritage Centre, 2008b), the main attraction of the region lies in its landscape, as the line passes valleys surrounded by the Alps. In addition, the region is not densely populated with only about 41 inhabitants / km^2 (Federal Statistical Office FSO, 2019). As a consequence, the main focus of the people passing through lies not on the train line but on the landscape itself. The view is often directed from inside the train outwards.

Lavaux on the other hand is situated in a settlement area. The area alongside the lake of Geneva includes residential and settlement area and also the defined world heritage zone of Lavaux contains a lot of houses. When including the buffer zone this number rises (UNESCO World Heritage Centre, 2007). Hence, the population density in the region is much higher than in the region of the RhB with about 722 inhabitants / km^2 (Federal Statistical Office FSO, 2019). Also the density of Flickr images taken is much lower in the region of the RhB than in Lavaux (see figures 4.3 & 4.4). People perceive the region not only as recreational area, but a great amount of people are spending time or even live in the region.

This difference can also be shown by the purpose of the world heritage site: the aim of a train is to move and get from one place to another whereas vineyards are permanently at the same place and need to be cultivated.

RQ1. What distinct landscape properties can be extracted using text and image processing that qualify to be UNESCO world heritage?

a) At the examples of Lavaux and RhB?

Lavaux
When taking into account the properties extracted by the image analysis, biophysical properties dominate the landscape. Especially, properties linked to water such as *lake* or *water* are returned often. In addition, the landform of the area is defining as it was extracted with both image and text analysis. The landscape is hilly, showing a view over the lake of Geneva. This setting corresponds to the region in which Lavaux is placed. The vineyard terraces are situated above the lake of Geneva. Furthermore, cultural properties are found when analysing the area of Lavaux. By analysing the proportion of the category cultural the purpose of the region to the local people can be extracted. The high proportion of infrastructure, settlement and land use suggests that the region is inhabited and the land is used for various purposes. This can also be seen at the proportion of the category activities which is much higher in this region than in the region of the RhB. People are using this region actively as they are living in it or are cultivating the vineyards. A large amount of the defined zone of cultural heritage contains settlement and residential areas. Moreover, the proportion of time depth shows that the area does have a history and developed over time, which is underlined with returned words such as *Jahrhundert* or *Bedeutung*. Interestingly, in the top returned tags no relation to grapes or vineyards can be found. Although words such as *Weinbau* or *vineyard* are returned in both the text and the image analysis, they were not annotated as their frequency is rather low. This underlines the assumption, that the perception of the region is not limited to the vineyard terraces, but is rather a region in which people are spending time.

RhB
The perception of the region of the RhB originates mostly from inside the train and is directed outwards. Therefore it is not surprising, that a high amount of annotated words are of biophysical nature. Especially the alpine character of the region can be seen when looking at often returned properties such as *snow* or *mountain*. The most common extracted biophysical properties include words related to the landform, the flora or the climate, which represent the mountains, the valleys but also the trees and the snow. However, cultural properties also define the region. Especially properties linked to infrastructure, which include properties linked to the train and the train tracks, can be extracted. Nevertheless, the development of the region can be seen with words linked to the time depth. Furthermore, properties indicating the sense of the place are extracted representing the identity of the landscape (*historisch, Geschichte*) as well as showing the meaning of the region (*Bedeutung, prägen, bestehen*).

b) What common properties can be extracted when comparing Lavaux to RhB?

Both regions show properties which show the close relation to nature. Biophysical features describing the nature present in the landscape such as *sky, cloud* or *plant* are present in both regions. Although both regions have a high proportion of biophysical properties, their composition differs (see figures 5.4 and 5.9) which shows the individual character of each site. However, not only the relation to nature but also the relation and importance to humanity is shown in both regions with the high proportion of cultural properties. Furthermore, the history and significance of the regions are extracted with properties connected to the sense of place. Extracted properties show the long history of the region (*Jahrhundert)* and the meaning of the regions to the people *Schutz, Bedeutung*.

6.3 Differences between text and image analysis

The distribution of the different classes shows a high similarity between the two sites for each data source (figures 5.1 & 5.6). However, in between the data sources, text and image, the distribution differs. This can be explained with the data structure (Wartmann et al., 2018). A journal article consists of text using natural language. In the text analysis, only words concording with *landschaft* are considered and not the text as a whole. The word's part of speech is not considered in the analysis, resulting in a wide range of words which are included in the analysis. By using natural language as data source, a lot of information can be transmitted (Van Zanten et al., 2016).

The information content of the returned tags on the other hand is more limited. Using the classification of returned tags by Jolicoeur et al. (1984), the level of detail which is obtained by Google Vision can be discussed. The most common returned tags all classify as basic or superordinate. On the other hand, tags on a subordinate level are less present. The level of detail of the extracted landscape properties is higher when using text data compared to image data. These differences can also be seen when comparing figures 4.7 & 4.9 resp. 4.8 & 4.10. Furthermore, the diversity of the data and amount of returned words is much higher in the text analysis than in the image analysis. The text analysis used articles written in natural language, which is constricted only in the extent of the language. The image analysis on the other hand used returned tags. Thereby only tags which were present in the training data are returned. Consequently, the training data defines the extent of the tags. This can be seen when comparing the total amount of returned words, which is much higher in the case of the text analysis than in the image analysis (5965 vs. 898 in the region of Lavaux, 4272 vs. 1314 in the region of the RhB). As a consequence, a higher percentage of the returned tags was annotated than it was the case for the text analysis. At the same time, the term frequencies of the annotated words are higher for the image analysis than it is in the text analysis. Using image analysis, common, less detailed properties are more likely extracted than specific, detailed ones.

RQ2. What differences in properties can be found when using text vs. image processing?

The properties obtained with text analysis are more detailed than the results for image processing. Capturing the perception and sense of place of a landscape shows better results when using text analysis. The majority of returned tags classify as basic or superordinate leading to less detail. Furthermore, only properties which are visible or can be derived from visual properties are captured when using image data. In addition, only tags which are present in the training data will be returned, limiting the extent further. The text data on the other hand allows descriptions which go further into detail. Additionally, feelings or sentiments associated with a place can be described allowing to add feeling of attachments or historical information, whereas the information content of a tag is much more limited. This is shown when looking at the extracted properties. Image analysis returns a high number of biophysical features, whereas properties classified as perceptual or sense of place are less present. The text analysis on the other hand shows an even distribution of all the categories, allowing to extract all aspects defining a landscape. Especially the categories perceptual and sense of place, which are strongly influenced by the people's perception and attachment to a place, are present. This tendency is also shown by Wartmann et al. (2018). However, different to Wartmann et al. (2018) information about activities are only found in image data and not in text data. Reason for this can be found in the types of magazines chosen. The text data in the case of Wartmann et al. (2018) hiking blogs are used, whereas in this study magazines covering cultural aspects and tourism are analysed.

Out of the three aspects of place by Agnew and Livingstone (2011) - location, locale and sense of place - mainly the aspect of locale is captured with image data. A small proportion of tags returned are annotated as sense of place. Using text data, all aspects of space are captured leading to a more complete image of the landscape.

6.4 Matching landscape properties to UNESCO

RQ3. To what extent match the extracted properties to the criteria of the UNESCO world heritage list?

The UNESCO defined five criteria to qualify as cultural heritage property. Most of these criteria can be seen when looking at the extracted properties. The property should be designed and created intentionally by humanity which can be seen at the amount of cultural properties extracted by text analysis. Furthermore, the amount of words connected to the history and identity of the regions are prove of the technical achievement made by humanity confirming another criterion of the UNESCO. Through the combination of text and image analysis, both the nature as well as the cultural aspect of the region were captured well. Including both biophysical as well as cultural properties the region fulfill the criteria of common works between nature and humanity. The authenticity and integrity cannot directly be shown at the extracted properties. However, some of the used text data date back to the beginning of the 20th century, showing the age of the place. Nevertheless, the integrity and authenticity cannot fully be captured using this approach.

When comparing the criteria to become part of the UNESCO world heritage list of cultural property with the definition of place proposed by Agnew and Livingstone (2011), the aspects of locale and sense of place are emphasized whereas the location is of lesser importance. Especially cultural properties which show human progress are desired. Also of importance are properties linked to the sense of place, which includes the history and attachment to the region and is part of the impact the region had on the local people. Finally, an important criteria to get on the list of cultural heritage is the absence of similar sites. The list aims to be as diverse as possible. Although the overall proportion of all categories between the two sites is similar, the sites are quite different when looking at the subcategories. Overall, the extracted properties match the five criteria of UNESCO quite well, showing a high agreement.

6.5 Limitations

Limitations regarding the representativity of the data set were elaborated in section 6.1. In this section, methodological limitations and challenges are discussed for the text analysis, the image analysis and the annotation of the data.

One general limitation exists due to the area of interests which were analysed. Both sites, Lavaux and the RhB, are classified as UNESCO world heritage property. The extracted properties match the UNESCO's criteria, nevertheless there is no reference region which is not classified as UNESCO world heritage. The extracted properties might therefore not be exclusive properties to the UNESCO world heritage but could be common properties of CES.

Oteros-Rozas et al. (2018) showed that water bodies are often related to recreational areas, whereas trees or urban features indicate recreation or spiritual and cultural heritage. These properties are often found in CES in Europe. Wartmann et al. (2018) highlight that especially biophysical properties are often extracted in CES, as they increase the connection and identity felt to a landscape. Furthermore, biophysical properties such as mountains or water increase the attractiveness and value of a CES (Tieskens et al., 2018; Van Zanten et al., 2016). These properties match well with the properties extracted by image analysis. Koblet and Purves (2020b) showed the importance of including text to capture the history and sense of place of a landscape. However, Cooper and Gregory (2019) elaborated on the emotions which are transported through text and ultimately influence the perception of a place. Different to recent studies, both data sources were combined to extract these different aspects of a landscape. To classify as UNESCO world heritage property both the naturality of the landscape as well as its history and connection to humanity is of importance. By combining text and image data, these two main aspects can be analysed.

6.5.1 Text analysis

The articles found on E-Periodica include various historical documents. It is shown by Butler et al. (2017) that historical documents are more difficult to process. Especially punctuation, case and hyphenation can be challenging for tokenization. In addition, the quality of the data differs. All the documents are digitized and converted into JSON format. However, digitizing text documents is a challenging task which can lead to incomplete text documents.

The text was handled as bag-of-words. As a result, the order and dependencies of the words were not taken into account, only concordance with the word *landschaft* were considered. Due to this approach, ambiguous words were taken out of context. The words were annotated into multiple categories, which would not be the case if word sense disambiguation had taken place. Rodd et al. (2004) showed a promising model of word sense disambiguation by capturing the sense of a word in a more flexible way.

6.5.2 Image analysis

Due to the use of Google Vision, some limitations need to be mentioned. Firstly, information about the people training the algorithm are missing. However, people training the algorithm have an influence on the outcome. The algorithm was trained based on their perception of landscapes which might be different to the definition of landscape used in this thesis (Mark et al., 2011), which creates a bias.

Furthermore, not only the trainers but also the training data is unknown. It is possible that no region in Switzerland was part of the training data. The algorithm was trained based on certain regions and returns labels based on these regions. However, this leads to tags that are returned which are not present in Switzerland. There are no hill stations in the Swiss Alps, nevertheless *hill station* was one of the most returned labels for the region of the RhB.

As all the returned tags are in English, two limitations arise. Firstly, the English language is not the mother tongue of the annotator. As a consequence, a language barrier exists when annotating the words.

Secondly, English is not the language spoken in the areas of interest. A language is adapted to the regions in which it is used and has often a specific vocabulary (Regier et al., 2016). As the tags are returned in English, it is possible that some properties which might be important to the identity of the landscape cannot be captured with the English language.

Finally, the data set used was limited to the number of images per user. However, this confines not only the amount of data taken into account, but also the diversity of the extracted terms. The extracted properties represent the most common perceived properties in the regions. To get a more diverse result, not the user frequency but the term frequency should be limited. By using metrics such as the term frequency–inverse document frequency (tf-idf) the importance of terms within a corpus can be determined. De Clercq et al. (2019) showed that the entire corpus can be analysed to extract properties related to food waste without the need to confine the data per user.

6.5.3 Annotation

When annotating words, the data is interpreted and a value is added. Thereby, the annotator's opinion is added to the corpus (Hunston, 2002). Although Cohen's Kappa showed a substantial agreement between the two annotators, most of the data was annotated by one person. The annotation results are influenced by the perception of the annotator.

Furthermore, annotating a text corpus removes the content and context of the data. The original text data should also be considered to enhance understanding (McEnery et al., 2006). Although a big amount of data can be processed by computational power, it is important to include both quantitative and qualitative techniques to analyse data (Biber, 1988).

A final limitation of the annotation and especially the comparison between the image and the text data is found in the language of the classified words. The words annotated from the text analysis are originally in German whereas the returned labels from the image analysis are in English. However, words have slightly different meanings in different languages. Regier et al. (2016) showed that in colder climates more differentiation exists in language between different types of snow and ice than in warmer regions. In addition, word ambiguity increases the challenge of comparing different languages to another. In a test by Regier et al. (2016) in which words were translated and back-translated into different languages, many languages failed.

6.6 Future research

In this analysis, only sites which are classified as UNESCO world heritage are taken into account. For further research, sites which are not part of the UNESCO world heritage list should be included to find differences between CES and UNESCO world heritage areas.

On a methodological level, the method for both text and image analysis could be improved further to gain more detailed information. The text analysis was solely based on concordance analysis. By including pattern detection, more context could be included in the analysis leading to a more detailed result.

The use of image recognition shows a high accuracy on the returned feature, however the results could be improved by training the machine learning algorithms with areas similar to the areas of interest.

This research concentrated on extracting landscape properties of value for the UNESCO world heritage list based on image and text data. For future analysis, the entire process could be reversed with the aim on finding potential UNESCO world heritage sites which are not yet classified as such based on image and text data.

Chapter 7

Conclusion

This master thesis was set out to find properties of cultural landscapes in Switzerland which are classified to be UNESCO world heritage properties. The objectives were to find characteristic properties of the two UNESCO world heritage properties in Switzerland, Lavaux and the RhB, by combining both text and image data. On a methodological level differences between the two types of analysis were looked at aiming to have the two methods complementing one another. Additionally, the extracted properties were compared to the criteria which UNESCO defined needed to be matched to qualify for the list of world heritage property.

With the rise of new technologies and Web 2.0, the availability and amount of data increased strongly. Combined with the easy access, the use of such data enables to include the perception of not only a few people but of a broad variety of people. As a consequence, different aspects in extracting a landscape can be captured. The resulting properties do not represent the objective view of only a few people but of a variety of people experiencing the landscape. It is shown, that the extracted properties depend strongly on the data source chosen. The data structure is crucial for the resulting distribution of the annotated words into the different categories. In addition, the level of detail of the extracted properties depends on the type of analysis. As in text analysis natural language is analysed, the level of detail in the extracted properties is much higher. In image analysis on the other hand, the returned tags are often on a basic level. However, this also shows the potential in combining several data sources for the extraction of landscape properties.

Using text analysis, it was possible to capture the history and identity of a region. Through the use of natural language the meaning and purpose of a place was extracted, which are relevant factors for the development and value of a region. On the other hand, the properties extracted with the image analysis include a high amount of biophysical properties. Thereby, the nature present in the landscape and landscape's character is captured. As the UNESCO world heritage list of cultural properties combines the idea of an outstanding technical progress made by humanity with a unique landscape, the combination of text and image analysis is promising in capturing both these aspects well.

Nevertheless, limitations of the two approaches were discovered. The advantage of text analysis lies in the level of detail which can be added. In case of the image analysis, a machine learning algorithm was used to extract properties. However, the resulting tags are only as exact and diverse as the original training data.

Furthermore, it has to be kept in mind, that all the data represent the view of the creator of the data and is not representative for the entire population. In general, the combination of different data sources is promising in receiving a more diverse impression of a landscape. However, the approaches of text and image analysis should be further elaborated using different methods for the extraction of properties in text analysis. In the case of the image analysis it is important that the machine learning algorithm was trained with data representative of the area in which it is applied.

Bibliography

Abdullah, Azizi, Remco C. Veltkamp and Marco A. Wiering (2009). "An ensemble of deep support vector machines for image categorization". In: *SoCPaR 2009 - Soft Computing and Pattern Recognition* 1, pp. 301–306. doi: 10.1109/SoCPaR.2009.67.

Agnew, John and David N. Livingstone (2011). *The SAGE handbook of geographical knowledge*. isbn: 9781446201091. doi: 10.4135/9781446201091.

Angradi, Ted R., Jonathon J. Launspach and Rick Debbout (2018). "Determining preferences for ecosystem benefits in Great Lakes Areas of Concern from photographs posted to social media". In: *Journal of Great Lakes Research* 44.2, pp. 340–351. issn: 03801330. doi: 10.1016/j.jglr.2017.12.007. url: https://doi.org/10.1016/j.jglr.2017.12.007.

Biber, Douglas (1988). *Variation across Speech and Writing*. doi: 10.1017/cbo9780511621024.

Bieling, Claudia (2014). "Cultural ecosystem services as revealed through short stories from residents of the Swabian Alb (Germany)". In: *Ecosystem Services* 8, pp. 207–215. issn: 22120416. doi: 10.1016/j.ecoser.2014.04.002. url: http://dx.doi.org/10.1016/j.ecoser.2014.04.002.

Bieling, Claudia et al. (2014). "Linkages between landscapes and human well-being: An empirical exploration with short interviews". In: *Ecological Economics* 105, pp. 19–30. issn: 09218009. doi: 10.1016/j.ecolecon.2014.05.013. url: http://dx.doi.org/10.1016/j.ecolecon.2014.05.013.

Britton, Bruce K. (1978). "Lexical ambiguity of words used in english text". In: *Behavior Research Methods & Instrumentation* 10.1, pp. 1–7. issn: 1554351X. doi: 10.3758/BF03205079.

Bubalo, Martina, Boris T. van Zanten and Peter H. Verburg (2019). "Crowdsourcing geo-information on landscape perceptions and preferences: A review". In: *Landscape and Urban Planning* 184.January, pp. 101–111. issn: 01692046. doi: 10.1016/j.landurbplan.2019.01.001. url: https://doi.org/10.1016/j.landurbplan.2019.01.001.

Bunce, R. G.H. et al. (1996). "Land classification for strategic ecological survey". In: *Journal of Environmental Management* 47.1, pp. 37–60. issn: 03014797. doi: 10.1006/jema.1996.0034.

Butler, Andrew (2016). "Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider". In: *Land-*

scape Research 41.2, pp. 239–252. issn: 14699710. doi: 10.1080/01426397.2015.1135315. url: http://dx.doi.org/10.1080/01426397.2015.1135315.

Butler, James O. et al. (2017). "Alts, Abbreviations, and AKAs: Historical Onomastic Variation and Automated Named Entity Recognition". In: *Journal of Map and Geography Libraries* 13.1. issn: 15420361. doi: 10.1080/15420353.2017.1307304.

Chan, Kai M.A., Terre Satterfield and Joshua Goldstein (2012). "Rethinking ecosystem services to better address and navigate cultural values". In: *Ecological Economics* 74, pp. 8–18. issn: 09218009. doi: 10.1016/j.ecolecon.2011.11.011.

Chen, Shih Hsin and Yi Hui Chen (2017). "A content-based image retrieval method based on the google cloud vision API and wordnet". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10191 LNAI.February, pp. 651–662. issn: 16113349. doi: 10.1007/978-3-319-54472-4{_}61.

Chesnokova, Olga et al. (2019). "Hearing the silence: finding the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District". In: *International Journal of Geographical Information Science* 33.12, pp. 2430–2454. issn: 13623087. doi: 10.1080/13658816.2018.1552789. url: https://doi.org/10.1080/13658816.2018.1552789.

Cooper, David (2019). "Digital Re-Enchantment: Place Writing, the Smartphone and Social Media". In: *Literary Geographies* 5.1, pp. 90–107. issn: ISSN 2397-1797. url: https://literarygeographies.net/index.php/LitGeogs/article/view/178.

Cooper, David and Ian Gregory (2019). "Mapping the English Lake District: A Literary GIS". In: 53.9, pp. 1689–1699. issn: 1098-6596. doi: 10.1017/CBO9781107415324.004.

Council of Europe (2000). "European landscape convention". In: Rep. Conv. Florence ETS No. 17, 8.

Das, Tanmoy (2017). "Machine Learning algorithms for Image Classification of hand digits and face recognition dataset". In: *International Research Journal of Engineering and Technology (IRJET)* 04.12, pp. 640–649.

De Clercq, Djavan, Zongguo Wen and Qingbin Song (2019). "Innovation hotspots in food waste treatment, biogas, and anaerobic digestion technology: A natural language processing approach". In: *Science of the Total Environment* 673, pp. 402–413. issn: 18791026. doi: 10.1016/j.scitotenv.2019.04.051. url: https://doi.org/10.1016/j.scitotenv.2019.04.051.

Dodds, Peter Sheridan et al. (2015). "Human language reveals a universal positivity bias". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.8. issn: 10916490. doi: 10.1073/pnas.1411678112.

E-Periodica (2020). *E-Periodica*. url: https://www.e-periodica.ch/digbib/about1 (visited on 15/03/2021).

Elwood, Sarah (2008). "Volunteered geographic information: Key questions, concepts and methods to guide emerging research and practice". In: *GeoJournal* 72.3-4, pp. 133–135. issn: 03432521. doi: 10.1007/s10708-008-9187-z.

Federal Office for the Environment FOEN (2008). *UNESCO-Welterbe Naturstätten*. url: https://opendata.swiss/de/dataset/unesco-welterbe-naturstatten (visited on 02/04/2021).

Federal Office of Culture FOC (2016). *UNESCO-Welterbe Kulturstätten*. url: https://opendata.swiss/de/dataset/unesco-welterbe-kulturstatten (visited on 02/04/2021).

Federal Statistical Office FSO (2019). *Statistik der Bevölkerung und Haushalte (STATPOP), Geodaten 2019*. url: https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/gebaeude-wohnungen-haushalte-personen/bevoelkerung-haushalte-ab-2010.assetdetail.14716365.html (visited on 05/04/2021).

Flickr (2021). *Work at Flickr*. url: https://flickr.com/jobs/ (visited on 05/04/2021).

Francis, N. and H. Kucera (1964). *Manual of Information to Accompany a Standard Corpus of Present-Day Edited American English for use with Digital Computers*. url: http://icame.uib.no/brown/ (visited on 20/05/2021).

Garside, R.G., Geoffrey Leech and Anthony Mark Mcenery (2020). "Introducing corpus annotation". In: *Corpus Annotation*. doi: 10.4324/9781315841366-7.

Gaskell, M. Gareth and William D. Marslen-Wilson (1997). "Integrating Form and Meaning: A Distributed Model of Speech Perception". In: *Language and Cognitive Processes* 12.5-6, pp. 613–656. issn: 01690965. doi: 10.1080/016909697386646.

Gerçek, Deniz, Vedat Toprak and Josef Stroblc (2011). "Object-based classification of landforms based on their local geometry and geomorphometric context". In: *International Journal of Geographical Information Science* 25.6, pp. 1011–1023. issn: 13658816. doi: 10.1080/13658816.2011.558845.

Girit, Hasan et al. (2012). "On the precision of search engines: Results from a controlled experiment". In: *Lecture Notes in Business Information Processing* 117 LNBIP.17102, pp. 201–212. issn: 18651348. doi: 10.1007/978-3-642-30359-3{_}18.

Gliozzo, Gianfranco, Nathalie Pettorelli and Mordechai Muki Haklay (2016). "Using crowdsourced imagery to detect cultural ecosystem services: A case study in South Wales, UK". In: *Ecology and Society* 21.3. issn: 17083087. doi: 10.5751/ES-08436-210306.

Goodchild, Michael F. (2007). "Citizens as sensors: The world of volunteered geography". In: *GeoJournal* 69.4, pp. 211–221. issn: 03432521. doi: 10.1007/s10708-007-9111-y.

Google Cloud (2020). *Vision AI*. url: https://cloud.google.com/vision/ (visited on 03/10/2020).

Gregory, Ian N. and Andrew Hardie (2011). "Visual GISting: Bringing together corpus linguistics and geographical information systems". In: *Literary and Linguistic Computing* 26.3, pp. 297–314. issn: 02681145. doi: 10.1093/llc/fqr022.

Hajishirzi, Hannaneh and N Goodwin Ave (2010). "Adaptive Near-Duplicate Detection via Similarity Learning Categories and Subject Descriptors". In: *Sigir*, pp. 419–426.

Haklay, Muki, Alex Singleton and Chris Parker (2008). "Web mapping 2.0: The neogeography of the GeoWeb". In: *Geography Compass* 2.6, pp. 2011–2039. issn: 17498198. doi: 10.1111/j.1749-8198.2008.00167.x.

Hotellerie Suisse (2021). *htr hotelrevue*. url: https://www.htr.ch/home.html (visited on 30/04/2021).

Hu, Yingjie (2018). "Geo-text data and data-driven geospatial semantics". In: *Geography Compass* 12.11. issn: 17498198. doi: 10.1111/gec3.12404.

Hunston, Susan (2002). *Corpora in Applied Linguistics*. doi: 10.1017/cbo9781139524773.

Jeawak, Shelan S., Christopher B. Jones and Steven Schockaert (2017). "Using flickr for characterizing the environment: An exploratory analysis". In: *Leibniz International Proceedings in Informatics, LIPIcs* 86.21. issn: 18688969. doi: 10.4230/LIPIcs.COSIT.2017.21.

Johnson, Leslie and Eugene Hunn (2010). *Landscape ethnoecology*. New York: Berghan Books.

Jolicoeur, Pierre, Mark A. Gluck and Stephen M. Kosslyn (1984). "Pictures and names: Making the connection". In: *Cognitive Psychology* 16.2, pp. 243–275. issn: 00100285. doi: 10.1016/0010-0285(84)90009-4.

Jones, M. and M. Stenseke, eds. (2011). *The European Landscape Convention: Challenges of Participation*. Vol. 37. 3. doi: 10.1080/01426397.2012.682388.

Kisilevich, Slava, Christian Rohrdantz and Daniel Keim (2010). ""Beautiful picture of an ugly place". Exploring photo collections using opinion and sentiment analysis of user comments". In: *Proceedings of the International Multiconference on Computer Science and Information Technology, IMCSIT 2010* 5, pp. 419–428. doi: 10.1109/imcsit.2010.5679726.

Koblet, Olga and Ross S. Purves (2020a). "Extracting Perceived Landscape Properties from Text Sources". PhD thesis. University of Zurich. doi: https://doi.org/10.5167/uzh-185108.

– (2020b). "From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally". In: *Landscape and Urban Planning*. issn: 01692046. doi: 10.1016/j.landurbplan.2020.103757.

Landis, J. Richard and Gary G. Koch (1977). "The Measurement of Observer Agreement for Categorical Data". In: *Biometrics* 33.1, p. 159. issn: 0006341X. doi: 10.2307/2529310.

Manning, C.D., P. Raghavan and H. Schütze (2008). *Introduction to Information Retrieval*. New ypork, NY, USA: Camdbridge University Press. doi: 10.1210/endo-38-3-156.

Manning, Christopher D., Hinrich Schütze and Gerhard Weikurn (2002). "Foundations of Statistical Natural Language Processing". In: *SIGMOD Record*. issn: 01635808. doi: 10.1145/601858.601867.

Mark, David M. et al. (2011). "Landscape in language". In: doi: 10.1075/clu.4.02mar.

McEnery, Anthony, Zhonghua Xiao and Yukio Tono (2006). *Corpus-based Language Studies: An Advanced Resource Book*. Vol. 86. 2.

Meiliana et al. (2017). "Mobile Smart Travelling Application for Indonesia Tourism". In: *Procedia Computer Science* 116, pp. 556–563. issn: 18770509. doi: 10.1016/j.procs.2017.10.059. url: https://doi.org/10.1016/j.procs.2017.10.059.

Millenium Ecosystem Assessment (2005). *Ecosystems and human well-being: synthesis*. doi: 10.1196/annals.1439.003. url: https://www.millenniumassessment.org/documents/document.356.aspx.pdf.

Miller, Harvey J. and Michael F. Goodchild (2015). "Data-driven geography". In: *GeoJournal* 80.4, pp. 449–461. issn: 15729893. doi: 10.1007/s10708-014-9602-6. url: http://dx.doi.org/10.1007/s10708-014-9602-6.

Mooney, Peter et al. (2017). "Considerations of Privacy, Ethics and Legal Issues in Volunteered Geographic Information". In: *Mapping and the Citizen Sensor*, pp. 119–135. doi: 10.5334/bbf.f.

Navigli, Roberto (2009). "Word sense disambiguation: A survey". In: *ACM Computing Surveys* 41.2. issn: 03600300. doi: 10.1145/1459352.1459355.

Nielsen, Jakob (2006). "Participation Inequality: Encouraging More Users to Contribute". In: 2008.1. Dezember, p. 2006.

NIST (2004). *Reuters Corpora (RCV1, RCV2, TRC2)*. url: https://trec.nist.gov/data/reuters/reuters.html (visited on 10/04/2021).

Oteros-Rozas, Elisa et al. (2018). "Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites". In: *Ecological Indicators* 94, pp. 74–86.
Perficient Digital Agency (2019). *Who Has the Best Image Recognition Engine?* (Visited on 08/05/2021).
Pleasant, Mary M. et al. (2014). "Managing cultural ecosystem services". In: *Ecosystem Services* 8, pp. 141–147. issn: 22120416. doi: 10.1016/j.ecoser.2014.03.006. url: http://dx.doi.org/10.1016/j.ecoser.2014.03.006.
Plieninger, Tobias et al. (2013). "Assessing, mapping, and quantifying cultural ecosystem services at community level". In: *Land Use Policy* 33, pp. 118–129. issn: 02648377. doi: 10.1016/j.landusepol.2012.12.013. url: http://dx.doi.org/10.1016/j.landusepol.2012.12.013.
Regier, Terry, Alexandra Carstensen and Charles Kemp (2016). "Languages Support Efficient Communication about the Environment: Words for Snow Revisited". In: *PLoS ONE* 11.4, pp. 1–17. issn: 19326203. doi: 10.1371/journal.pone.0151138.
Richards, Daniel R. and Bige Tunçer (2018). "Using image recognition to automate assessment of cultural ecosystem services from social media photographs". In: *Ecosystem Services* 31, pp. 318–325. issn: 22120416. doi: 10.1016/j.ecoser.2017.09.004. url: https://doi.org/10.1016/j.ecoser.2017.09.004.
Rodd, Jennifer M., M. Gareth Gaskell and William D. Marslen-Wilson (2004). "Modelling the effects of semantic ambiguity in word recognition". In: *Cognitive Science* 28.1, pp. 89–104. issn: 03640213. doi: 10.1016/j.cogsci.2003.08.002.
San Roque, Lila et al. (2015). "Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies". In: *Cognitive Linguistics* 26.1, pp. 31–60. issn: 16133641. doi: 10.1515/cog-2014-0089.
Schaich, Harald, Claudia Biding and Tobias Plieninger (2010). "Linking ecosystem services with cultural landscape research". In: *Gaia* 19.4, pp. 269–277. issn: 09405550. doi: 10.14512/gaia.19.4.9.
Seely-Gant, Katie and Lisa M Frehill (2015). "Exploring Bias and Error in Big Data Research". In: *Washington Academy of Sciences. Journal of the Washington Academy of Sciences* 101.3, p. 29.
Seresinhe, Chanuki Illushka, Tobias Preis and Helen Susannah Moat (2017). "Using deep learning to quantify the beauty of outdoor places". In: *Royal Society Open Science* 4.7. issn: 20545703. doi: 10.1098/rsos.170170.
Smith, Jos (2013). "An archipelagic literature: re-framing 'The New Nature Writing'". In: *Green Letters*. issn: 1468-8417. doi: 10.1080/14688417.2012.750840.
Stefanidis, Anthony, Andrew Crooks and Jacek Radzikowski (2013). "Harvesting ambient geospatial information from social media feeds". In: *GeoJournal* 78.2, pp. 319–338. issn: 03432521. doi: 10.1007/s10708-011-9438-2.
Tenerelli, Patrizia, Urška Demšar and Sandra Luque (2016). "Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes". In: *Ecological Indicators* 64, pp. 237–248. issn: 1470160X. doi: 10.1016/j.ecolind.2015.12.042.
Text + Berg (2021). *Text + Berg digital*. url: http://textberg.ch/site/de/willkommen/ (visited on 10/04/2021).
Thomee, Bart et al. (2016). "YFCC100M: The new data in multimedia research". In: *Communications of the ACM* 59.2, pp. 64–73. issn: 15577317. doi: 10.1145/2812802.

Tieskens, Koen F. et al. (2018). "Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape". In: *Landscape and Urban Planning* 177.September, pp. 128–137. issn: 01692046. doi: 10.1016/j.landurbplan.2018.05.002. url: https://doi.org/10.1016/j.landurbplan.2018.05.002.

Torra, Vicenç and Guillermo Navarro-Arribas (2014). "Data privacy". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4.4, pp. 269–280. issn: 19424795. doi: 10.1002/widm.1129.

Turner, Andrew J (2006). *Introduction to Neogeography*.

UNESCO World Heritage Centre (2007). *Lavaux, Vineyard Terraces*. url: https://whc.unesco.org/en/list/1243 (visited on 30/10/2020).

– (2008a). "Operational Guidelines for the Implementation of the World Heritage Convention". In: *Operational Guidelines for the Implementation of the World Heritage Convention* July, p. 167. url: http://whc.unesco.org/archive/opguide08-en.pdf.

– (2008b). *Rhaetian Railway in the Albula / Bernina Landscapes*. url: https://whc.unesco.org/en/list/1276 (visited on 30/10/2020).

Van Zanten, Boris T. et al. (2016). "Continental-scale quantification of landscape values using social media data". In: *Proceedings of the National Academy of Sciences of the United States of America* 113.46, pp. 12974–12979. issn: 10916490. doi: 10.1073/pnas.1614158113.

Vickery, G. and S. Wunsch-Vincent (Sept. 2007). *Participative Web and User-Created Content*. OECD. isbn: 9789264037465. doi: 10.1787/9789264037472-en. url: https://www.oecd-ilibrary.org/science-and-technology/participative-web-and-user-created-content_9789264037472-en.

Wang, Pin, En Fan and Peng Wang (2021). "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning". In: *Pattern Recognition Letters* 141, pp. 61–67. issn: 01678655. doi: 10.1016/j.patrec.2020.07.042. url: https://doi.org/10.1016/j.patrec.2020.07.042.

Warnock, Steven and Geoffrey Griffiths (2015). "Landscape Characterisation: The Living Landscapes Approach in the UK". In: *Landscape Research* 40.3. issn: 14699710. doi: 10.1080/01426397.2013.870541.

Warriner, Amy Beth and Victor Kuperman (2015). "Affective biases in English are bi-dimensional". In: *Cognition and Emotion* 29.7, pp. 1147–1167. issn: 14640600. doi: 10.1080/02699931.2014.968098. url: http://dx.doi.org/10.1080/02699931.2014.968098.

Wartmann, Flurina M., Elise Acheson and Ross S. Purves (2018). "Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach". In: *International Journal of Geographical Information Science* 32.8, pp. 1572–1592. issn: 13623087. doi: 10.1080/13658816.2018.1445257. url: https://doi.org/10.1080/13658816.2018.1445257.

Zachara, Marek and Dariusz Pałka (2016). "Comparison of text-similarity metrics for the purpose of identifying identical web pages during automated web application testing". In: *Advances in Intelligent Systems and Computing*. isbn: 9783319285597. doi: 10.1007/978-3-319-28561-0{_}3.

Zimmer, Michael (2018). "Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity". In: *Social Media and Society* 4.2. issn: 20563051. doi: 10.1177/2056305118768300.

Zucker, Daniel F. (2007). “What does Ajax mean for you?” In: *Interactions* 14.5, pp. 10–12. issn: 10725520. doi: 10.1145/1288515.1288523.

Appendix A

Appendix

A.1 Code

The entire code used for data extraction, data processing and data visualization are available at GitHub: https://github.com/fchriste/MA_Code.

A.2 Additional Tables

In the following tables the annotated words are listed.

category	subcategory	word
biophysical	geology	Grundgestein, geologisch, alpin
	soil	Terrain
	flora	Pflanze, Baum, Gras, Holz, Vegetation, Blume, Ast, Wald
	fauna	leben
	climate	Himmel, Wolke, Schnee, Sonnenlicht, Sommer, Klima
	landform	Berge, Tal, Landschaft, Val
	land cover	Natur, natürlich, Umwelt, Land, Erde
	hydrology	Wasser, See
cultural	land use	Kanton, Gemeinde, Region, Raumplanung, Landwirtschaft, Gebiet, Raum, Kulturlandschaft
	settlement	Architektur, Haus, Waldhaus, bauen, bilden, Dörfer, Kunstbauten, Bau, Bevölkerung, Mensch, Stadt, Siedlung, Ort, Gesellschaft, Platz, Bauten
	infrastructure	Abdeckung, Bundesamt, Tourismus, Abteilung, Institut, Bafu, Hotel, Strasse, Zug, Bahn, Kulturlandschaft, Brücke, Albulabahn, Morosani, fahren, Strecke, Projekt, Ingenieur, Tunnel, Verbindung, Schloss, Landrat, Anlagen, Technik Linienführung, Express, Franken, Adresse, planen, RhB, verbinden, führen, technisch, Meter, Kultur, wirtschaftlich
	anthropogenic objects	Wein, Buch, Karten, Telefon, Wissen, Seite, Liste, GmbH, Handelsregister, RhB
	time depth	immer, heute, Jahr, Zeit, seit, bereits, bisher, schon, Juli, meist, bald, morgen, Jahrhundert, erst, früh
perceptual	color / vision	Bild, gross, sehen, Blick, Aufnahmen, Sicht, zeigen, Schönheit, Attraktivität, Inszenierung, neu, gross, hoch, lang, Gestaltung, wirken, klein, alt, gestalten, darstellen, scheinen
	touch	stark
	feel	-
	weather	-
	atmospheric conditions	-
	pattern	rund, viel, wenig, zahlreich, beide, drei, letzter, mehr, kaum, vier, gemeinsam, verschieden, gerade, folgend, genau, folgen
	texture	Zustand, ganz
toponyms		Schweiz, Lavaux, Aargau, Alpen, Lausanne, Bern, Zürich, Basel, St. Gallen, Genfersee, Jura, Bellinzona, Davos, St. Moritz, Thusis, Zug, rhätisch
sense of place	meaning	Heimat, Beispiel
	feeling	lachen, Wunsch, verändern, wichtig, mögen
	associations	gut
	attachments	Heimatschutz, Stiftung, Schutz, erhalten, Verkehrsverein, Welterbe, Denkmalpflege, Bedeutung, unser, deren
	identity	Unesco, Sinn, Begriff, national, einzigartig, verstehen
	history	Geschichte, Entwicklung, entstehen, Wandel, bleiben, bestehen, gründen, prägen, historisch, erfahren
unclassified	-	liegen, sollen, werden, Beitrag, Nachhaltigkeit, lassen, sein, sowohl, stellen, sowie, geben, Aufgabe, Art, Teil, etwa, müssen, anhand, wer, ziehen, stehen, Geographie, Kunst, fast, Mal, zurück, dafür, gemeinsam, jedoch, aufnehmen, Ausstellung, Thema, Seite, setzen, Abteilung, sagen, bringen, bieten, denken, können, möglich, gelten, kommen, fragen, bestimmen, gehen, wohl, zustehen, ökologisch

Table A.1: Annotated words from text analysis

category	subcategory	tag
	geology	bedrock, geological phenomenon
	soil	terrain
	flora	plant, tree, grass, wood, larch, plant community, vegetation, flower, terrestrial plant, twig, branch, leaf, botany, evergreen, petal, flowering plant, shrub, biome
biophysical	fauna	bird, fawn, beak, biome
	climate	sky, cloud, snow, ecoregion, cumulus, atmosphere, biome, sunlight, light, biome
	landform	mountain, natural landscape, slope, landscape, ice cap, mountainous landform, glacial landform, fluvial landforms of streams, mountain range, formation, rolling, coastal and oceanic landforms, plain, highland, valley, nunatak, hill
	land cover	nature, natural environment, grassland, ground cover
	hydrology	water, lake, water resources, water courses, body of water, fluid, liquid
	land use	land lot, agriculture
	settlement	house, property, architecture, neighbourhood, residential area, rural area, cottage, hood, city, urban design
cultural	infrastructure	building, window, train, rolling stock, track, road surface, mode of transport, electricity, facade, railway, asphalt, infrastructure, road, gas, thoroughfare, locomotive, fence, wall, roof, flooring, tower, public address system, bridge, swimming pool
	domesticated animals	working animals
	anthropogenic objects	vehicles, wheel, motor vehicle, tire, car, font, bank, fixture, automotive lighting, table, automotive tire, trunk, chair, boat, cap, watercraft, hat, street light, shorts, tableware, headgear, eyewear, t-shirt, door, drinkware, flower pot, glass, automotive exterior, sunglasses, lighting, jacket, boats and boating equipment and supplies, sports equipment, ski equipment, helmet, string instrument, musical instrument, plucked string instruments, guitar accessory, musical instrument accessory, string instrument accessory, microphone, swim cap, bicycle equipment and gear, goggles, bicycle equipment and supplies
	time depth	-
	color / vision	azure, tints and shades, green, blue, black and white, shade, black, interior design, automotive design, urban design, evergreen, fawn
	touch	-
perceptual	feel	freezing
	sound	music, guitar, concert, musician, string instrument, musical instrument, plucked string instruments, music artist, band plays
	smell	gas
	weather	atmospheric phenomenon
	atmospheric conditions	horizon, dusk, daytime, afterglow, sunset
	pattern	rectangle, symmetry, line
	texture	liquid, fluid, glass, gas
activities		travel, sports equipment, people in nature, ski equipment, winter sport, goggles, ski, bicycle equipment and supplies, musician, musical instrument, swimming pool, swimmer, recreation, outdoor recreation, guitar, swim cap, flash photography, music artist, sport gear, band plays, medley swimming, sports, medley swimming
	meaning	gesture
sense of place	feeling	smile, laugh, happy
	associations	leisure, recreation, outdoor recreation
unclassified	-	food, composite material, ingredient, vision care, art, geography

Table A.2: Annotated tags from Flickr images

A.3 Software and Scripts

Python Programming Language and Environment		
Python	3.8.5	python.org
PyCharm	2019.2.3	www.jetbrains.com/pycharm/

Python libraries	*Version*	*URL*
beautifulsoup4	4.9.3	https://pypi.org/project/beautifulsoup4/
geopandas	0.8.1	https://geopandas.org/
geoplot	0.4.1	https://pypi.org/project/geoplot/
google-cloud-vision	2.0.0	pypi.org/project/google-cloud-vision/
json	2.0.9	docs.python.org/3/library/json.html
langdetect	1.0.8	pypi.org/project/langdetect/
matplotlib	3.3.3	pypi.org/project/matplotlib/
nltk	3.5	pypi.org/project/nltk/
os	3.8.5	docs.python.org/3/library/os.html
pandas	1.2.0	pandas.pydata.org/
re	2.2.1	docs.python.org/3/library/re.html
requests	2.25.1	pypi.org/project/requests/
seaborn	0.11.1	pypi.org/project/seaborn/
shapely	1.7.1	pypi.org/project/Shapely/
sklearn	0.24.1	https://sklearn.org/
spacey	3.0.6	https://spacy.io/
urllib3	1.26.2	pypi.org/project/urllib3/

R Programming Language and Environment		
R	4.0.2	www.r-project.org/
RStudio	1.2.5001	rstudio.com/

R Packages	*Version*	*URL*
dplyr	1.0.2	dplyr.tidyverse.org/.org
forcats	0.5.0	https://forcats.tidyverse.org/
ggplot2	3.3.2	ggplot2.tidyverse.org
jsonlite	1.7.1	rdocumentation.org/packages/jsonlite/versions/1.7.1
mapview	2.9.0	https://r-spatial.github.io/mapview/
sqldf	0.4-11	https://github.com/ggrothendieck/sqldf
wordcloud	2.6	http://blog.fellstat.com/?cat=11

Personal declaration

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

Fabienne Christen, 16.06.21